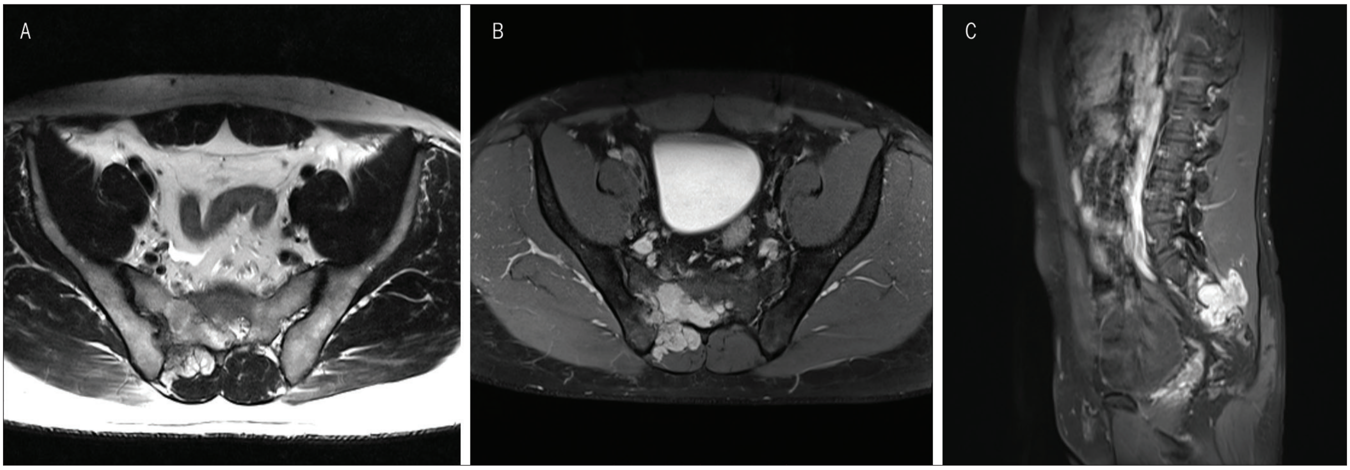


FIGURE. A chondroblastic osteosarcoma of the sacrum at the S2-3 level in a 36-year-old man. (A) Axial T2-weighted, (B) axial, contrast-enhanced, fat-saturated T1-weighted, and (C) sagittal, contrast-enhanced, fat-saturated T1-weighted magnetic resonance images of an ill-defined intraspinal lesion (approximately 2.8 [anteroposterior] × 2.8 [mediolateral] × 3.6 [craniocaudal] cm) at the S2-3 level, encasing the S2 and S3 roots and extending into the sacral foramina and paraspinal soft tissue (lumbar multifidi), with scalloping/erosion of the posterior surface of the sacral vertebrae (S2 and S3) on the right side. The lesion appears mildly hyperintense on the T2-weighted image (A). Postcontrast evaluation demonstrated moderate heterogeneous enhancement of the lesion (B and C). The right sacroiliac joint shows tumor invasion on its posterior aspect (A and B).

Sacral Osteosarcoma Masquerading as Posterior Thigh Pain

ASHOKAN ARUMUGAM, MPT, PhD, *Department of Community Medicine and Rehabilitation, Physiotherapy Section, Umeå University, Umeå, Sweden.*

A 36-YEAR-OLD MAN WITH INSIDIOUS onset of posterior right thigh pain that had started 1 month previously sought physical therapy consultation after his pain was nonresponsive to nonsteroidal anti-inflammatory drugs prescribed by his physician. Radiographs were noncontributory. He had constant unremitting pain of a gnawing type that caused difficulty in falling asleep and woke him at night. He reported constipation but was uncertain of altered sensation over the genital and perianal regions. On observation and palpation, no inflammatory signs or tenderness in the lower back and thigh were evident. Physical examination of the lumbosacral spine and bilateral lower-limb reflexes, sensation, and manual muscle testing were normal. He did not consent to genital/perianal sensory examination. The slump test was positive, and initial therapy with neurodynamic sliders in the slump position[1] aggravated his symptoms.

Given the atypical findings of unrelenting pain and positive neurodynamic tests with normal spinal mobility and lower extremity neurological screen, the physical therapist referred him to an orthopaedic surgeon, who ordered magnetic resonance imaging. Magnetic resonance imaging revealed a sacral tumor (**FIGURE**). Differential diagnosis included chordoma, giant cell tumor, chondrosarcoma, plexiform neurofibroma, and osteosarcoma of the sacrum. Subsequently, a computed tomography–guided biopsy by a radiologist confirmed a high-grade sacral chondroblastic osteosarcoma. The patient elected nonsurgical treatment and underwent 6 cycles (1 cycle every 3 weeks) of chemotherapy (methotrexate, cisplatin, and doxorubicin). Approximately 1 month after beginning the chemotherapy, he gradually developed right foot drop, urofecal incontinence, and diffuse edema in the right gluteal region. Repeat computed tomography at the end of the 18-week period of chemotherapy revealed tumor infiltration of the right gluteal muscles and metastasis to the lungs and liver, and palliative care began. He died 1 month later. This case highlights the importance of timely referral to a specialist and subsequent imaging in the setting of worsening pain unrelieved with a short trial of nonsurgical care. ◉ *J Orthop Sports Phys Ther 2018;48(8):665. doi:10.2519/jospt.2018.8032*

Reference

1. Shacklock MO. *Clinical Neurodynamics: A New System of Musculoskeletal Treatment*. Edinburgh, UK: Elsevier Health Sciences/Butterworth-Heinemann; 2005.

DAVID B. BERRY, PhD[1] • **BAHAR SHAHIDI,** PT, PhD[2,3] • **ANA E. RODRÍGUEZ-SOTO,** PhD[1]
JAN M. HUGHES-AUSTIN, PT, PhD[3] • **KAREN R. KELLY,** PhD[4] • **SAMUEL R. WARD,** PT, PhD[1-3]

Lumbar Muscle Structure Predicts Operational Postures in Active-Duty Marines

The muscles of the lumbar spine are crucial for stabilizing and supporting the upper trunk, especially during dynamic loading conditions. A muscle's force-generating capacity is directly related to its architectural and microstructural features, which are therefore variables of interest when trying to assess muscle health. Physiological cross-sectional area (PCSA) is a measure of muscle architecture that can be measured to estimate muscle force.[27] However, it is difficult to precisely measure PCSA in vivo, as it includes measures of muscle architecture, such as pennation angle and normalized fiber length. Volume is a dominant input variable to measure muscle PCSA and is commonly used as a proxy for muscle force-producing capacity.[7,9] However, muscle is a heterogeneous tissue, also consisting of fat and collagenous tissues, which can confound measures of muscle volume.

BACKGROUND: The relationship between lumbar spine posture and muscle structure is not well understood.

OBJECTIVES: To investigate the predictive capacity of muscle structure on lumbar spine posture in active-duty Marines.

METHODS: Forty-three Marines were scanned in this cross-sectional study, using an upright magnetic resonance imaging scanner while standing without load and standing, sitting, and prone on elbows with body armor. Cobb, horizontal, and sacral angles were measured. Marines were then scanned while unloaded in supine using a supine magnetic resonance imaging scanner. The imaging protocol consisted of T2 intervertebral disc mapping; high-resolution, anatomical, fat-water separation, and diffusion tensor imaging to quantify disc hydration and muscle volume, fat fraction, and restricted diffusion profiles in the lumbar muscles. A stepwise multiple linear regression model was used to identify physiological measures predictive of lumbar spine posture.

RESULTS: The multiple regression model demonstrated that fractional anisotropy of the erector spinae was a significant predictor of lumbar posture for 7 of 18 dependent variables measured, and explained 20% to 35% of the variance in each model. Decreased fractional anisotropy of the erector spinae predicted decreased lordosis, lumbosacral extension, and anterior pelvic tilt.

CONCLUSION: Fractional anisotropy is inversely related with muscle fiber size, which is associated with the isometric force-generating capacity of a muscle fiber. This suggests that stronger erector spinae muscles predict decreased lordosis, lumbosacral extension, and anterior pelvic tilt in a highly trained population. *J Orthop Sports Phys Ther 2018;48(8):613-621. Epub 17 May 2018. doi:10.2519/jospt.2018.7865*

KEY WORDS: *diffusion tensor imaging, lumbar spine, magnetic resonance imaging, military, posture, skeletal muscle*

Skeletal muscle exhibits a classic structure-function relationship, where its microstructural properties are closely related to whole muscle function. For example, muscle fiber isometric force-generating capacity is directly related to fiber cross-sectional area.[17,21,22] It is also difficult to measure muscle microstructure in vivo, although there is some evidence that diffusion-based imaging techniques are sensitive to different features of muscle microstructure, in particular fiber area.[5,8,12,35]

With injury and age, atrophy of muscle fibers and replacement of muscle tis-

[1]Department of Bioengineering, University of California, San Diego, La Jolla, CA. [2]Department of Radiology, University of California, San Diego, La Jolla, CA. [3]Department of Orthopaedic Surgery, University of California, San Diego, La Jolla, CA. [4]Department of Warfighter Performance, US Naval Health Research Center, San Diego, CA. The University of California, San Diego (UCSD) and US Naval Health Research Center (NHRC) Institutional Review Boards approved this study (UCSD protocol 110483; NHRC.2013.0023). The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or reflecting the views of the US Navy, US Army, US Air Force, the US Departments of Defense and Veterans Affairs, or the US Government. Any references to commercial organizations and trade names in this article do not constitute an official US Army endorsement or approval of the products or services of these organizations. Dr Kelly is a military service member. This work was prepared as part of her official duties. Title 17, USC, §105 provides that "Copyright protection under this title is not available for any work of the United States Government." Title 17, USC, §101 defines a US Government work as work prepared by a military service member or employee of the US Government as part of that person's official duties. The US Army Medical Research Acquisition Activity supported this manuscript through award number W81XWH-13-2-0043 (work unit number 1305). This work was also supported by US Department of Defense grant PR120576 (to Drs Ward and Kelly). Dr Ward is also supported by grants from the National Institutes of Health (R01AR070830, R01HD073180, R01HD088437). 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 Samuel R. Ward, University of California, San Diego, 9500 Gilman Drive (0863), La Jolla, CA 92093. E-mail: s1ward@ucsd.edu

sue with adipose and fibrotic tissue are typically observed compared to healthy muscle, further decreasing the overall volume of functional contractile tissue in the muscle.[7,36] As pathogenic (diseased) muscle becomes atrophied and fibrotic and contains more adipose tissue, the active and passive force-generating potential of the whole muscle changes, which can have a direct and negative effect on joint stability, range of motion, and posture.[20,26,39,48] The multifidus muscle is considered to be one of the primary muscular stabilizers of the lumbar spine, due to its ability to produce high forces over a narrow range of lengths, and often undergoes the pathogenic changes associated with injury, low back pain (LBP), or age.[46]

Changes to the orientation and position of bony structures of the spinal column are often observed simultaneously with these changes in muscle composition.[19,38,40] With age, gross changes in spinal posture, such as decreased lumbar lordosis, increased lumbar flexion, and increased pelvic tilt, are typically observed.[13,14,18,42] Decreased segmental range of motion has also been measured at vertebral levels with intervertebral disc (IVD) degeneration,[13,16] which is defined as decreased hydration of the nucleus pulposus with accompanying disc height loss.[25] However, changes in muscle structure, lumbar posture, and IVD health are not independent of one another, and their effects are confounded by age, sex, activity level, and the timing of disease progression.

In addition to associated changes with age and disease, external stimuli, such as carrying load or whole-body position, may affect posture.[3,29-31] Military members are highly active and often required to carry heavy loads in unusual positions. Studies investigating how Marines adapt to load carriage suggest that they routinely operate under conditions that put them at risk for developing lumbar musculoskeletal injury and that they exhibit higher rates of LBP than civilians.[33,34] This may be attributed to pathophysiologic changes of the lumbar spine structures as a result of the heavy loads and unusual postures experienced in training and combat.[15,28] A noninvasive tool that can correlate musculoskeletal health to posture under relevant loading conditions would allow clinicians to tailor rehabilitation protocols to target specific musculoskeletal components involved in regulating posture to mitigate an individual's risk of lumbar spine injury.

The purpose of this study was to investigate the predictive capacity of muscle structure, IVD health, and anthropometric measures on lumbar spine posture in active-duty Marines. We hypothesized that multifidus muscle volume would predict lumbar posture in different positions, because the multifidus provides intersegmental lumbar support and muscle volume is related to muscle strength.

METHODS

The University of California, San Diego and US Naval Health Research Center Institutional Review Boards approved this study, and all volunteers gave verbal and written consent to participate. Marines were included in this study if they were male, over 18 years of age, and healthy enough to perform their assigned duty. Marines were excluded from this study if they had undergone lumbar spine surgery or had the possibility of shrapnel in their bodies. Marines were not recruited based on LBP status or history. All Marines underwent standard magnetic resonance imaging (MRI) safety screening prior to scanning. All scans were performed early in the morning, between 4 am and 9 am.

Upright MRI

Marines were scanned using an upright 0.6-T MRI scanner (UPRIGHT Multi-Position MRI; Fonar Corporation, Melville, NY) and a planar coil. An elastic band was used to hold the coil against the volunteer's lumbar spine between the L1 and S1 levels while standing. The band was secured to hold the coil in place without altering the volunteer's natural position. A 3-plane localizer (repetition time [TR], 1254 milliseconds; echo time [TE], 100 milliseconds; field of view (FoV), 34 cm; matrix, 256 × 256; in-plane resolution, 1.33 × 1.33 mm; thickness, 9 mm; number of excitations, 1; time, 0:17) and sagittal T2-weighted images (TR, 1974 milliseconds; TE, 160 milliseconds; FoV, 35 cm; matrix, 224 × 224; in-plane resolution, 1.56 × 1.56 mm; thickness, 3 mm; gap, 0 mm; number of excitations, 1; time, 2:12) were acquired.[4]

Upright MRI: Load Carriage and Position Tasks

Marines were scanned in the following positions: standing without load, standing with body armor, sitting with body armor, and prone on elbows with body armor. Positions with external load were randomized to control for the cumulative effects of loading or time. The selected positions were static positions that Marines are often required to maintain for extended periods, depending on military occupational specialty, and are often reported as provoking LBP.[3] The load magnitude of 11.3 kg was chosen based on the use of body armor, which is the minimum protective equipment Marines are required to wear during military operations and training. Marines were not provided instruction on how to assume each position, but were asked to hold each position steady for the duration of the MRI acquisition. A previous study has shown no statistically significant difference in test-retest variation in posture within a subject, even after performing heavy-load and activity tasks.[31]

Upright MRI: Postural Measurements

Postural measurements were generated from upright MRI images in each position, using a previously validated algorithm.[3] Briefly, digital seed points were manually placed on the corners of the vertebral body and on the posterior elements of each vertebra using OsiriX Version 3.9.3 imaging software (Pixmeo

SARL, Bernex, Switzerland).[32] The locations of the seed points were imported into MATLAB (The MathWorks, Inc, Natick, MA) and used to define an end plate–based joint coordinate system applied to the superior and inferior end plate of each vertebra (L1-S1).

Global measurements of lumbar spine posture were calculated for each position to characterize the posture of the lumbar spine. Global measures included angle with respect to the horizontal to assess lumbosacral flexion/extension, sacral slope to assess sacral tilt, and sagittal Cobb angle to assess lumbar lordosis (**FIGURE 1**). Root-mean-square error values for global measurements were measured previously and are 0.28°, 0.95°, and 0.95°, respectively.[4,31] Global measurements between the standing unloaded and the standing loaded (delta load) positions, and between the sitting loaded and prone on elbows loaded (delta position) positions, were also calculated to determine lumbar kinematics in response to load and dynamic movement, respectively.

Supine MRI

Magnetic resonance images of the lumbar spine (L1-S1) were acquired using a 3-T MRI scanner (Discovery MR750; GE Healthcare, Waukesha, WI) and spine array coil. The imaging protocol consisted of (1) an anatomical scan, (2) fat-water separation scan, (3) diffusion tensor imaging (DTI) of the lumbar spine, and (4) T2 mapping of each lumbar IVD. Marines were scanned supine, with the lumbar muscles relaxed, to mitigate motion and breathing artifacts. The anatomical scan was an axial, fast spoiled-gradient echo with the following scanning parameters: TR, 5 milliseconds; TE, 2.3 milliseconds; flip angle, 20°; FoV, 32 cm; acquisition matrix, 512 × 512; pixel size, 0.625 × 0.625 mm^2; slice thickness, 1 mm; no gap; number of averages, 3. Fat-water separation images were acquired utilizing a 3-point iterative decomposition of water and fat, with echo asymmetry and a least-squares estimation sequence in the sagittal plane (TR, 1974 milliseconds; TE, 160 milliseconds; flip angle, 20°; FoV, 25.6 cm; 176 slices; acquisition matrix, 256 × 256; voxel size, 1 × 1 × 1 mm^3; no gap; number of averages, 1). Scanning parameters of the axial DTI sequence were as follows: TR, 10 seconds; TE, 46 milliseconds; FoV, 19.2 cm; 82 slices; acquisition matrix, 128 × 128; pixel size, 1.5 × 1.5 mm^2; slice thickness, 3 mm; no gap; B value, 400 mm^2/s; 45 diffusion directions. Last, multispin-echo data (8 echoes; TE, 8.6 to 68.8 milliseconds; TR, 800 milliseconds; FoV, 16 cm; 5 slices; acquisition matrix, 256 × 256; voxel size, 0.625 × 0.625 × 5 mm^3; no gap; number of averages, 1) were acquired and used to estimate the T2 of each lumbar IVD. The scanning plane was axial oblique, parallel to each lumbar IVD.

Supine MRI: Lumbar Physiology Measurements

Anatomical images were imported into the OsiriX imaging software for segmentation. Contours of the multifidus, erector spinae group, psoas, and quadratus lumborum muscles were manually traced from the L1 to S1 lumbar levels. The resulting segmentations were used to generate masks to quantify muscle volumes, fat fraction, and diffusion properties of Marines in the supine position.

Images acquired using the fat-water separation sequence yielded 2 sets of images: 1 where both fat and water MRI signals are in phase, and 1 where they are out of phase. This allows for isolating the independent contributions of water (S_W) and fat (S_F) to the total MRI signal. These data were then used to quantify the fat fraction (FF) of the multifidus and erector spinae group with the following relationship: $FF = S_F/(S_W + S_F)$.

The diffusion tensor was fitted using Analysis of Functional NeuroImages software (National Institutes of Health, Bethesda, MD) and function 3dDWItoDT.[6] Mean diffusivity, fractional anisotropy (FA), and the 3 eigenvalues (λ_{1-3}) of the diffusion tensor are reported. The quantitative relationship of diffusion variables to specific features of muscle microstructure is the focus of current work, although there is some evidence that they are related to muscle fiber

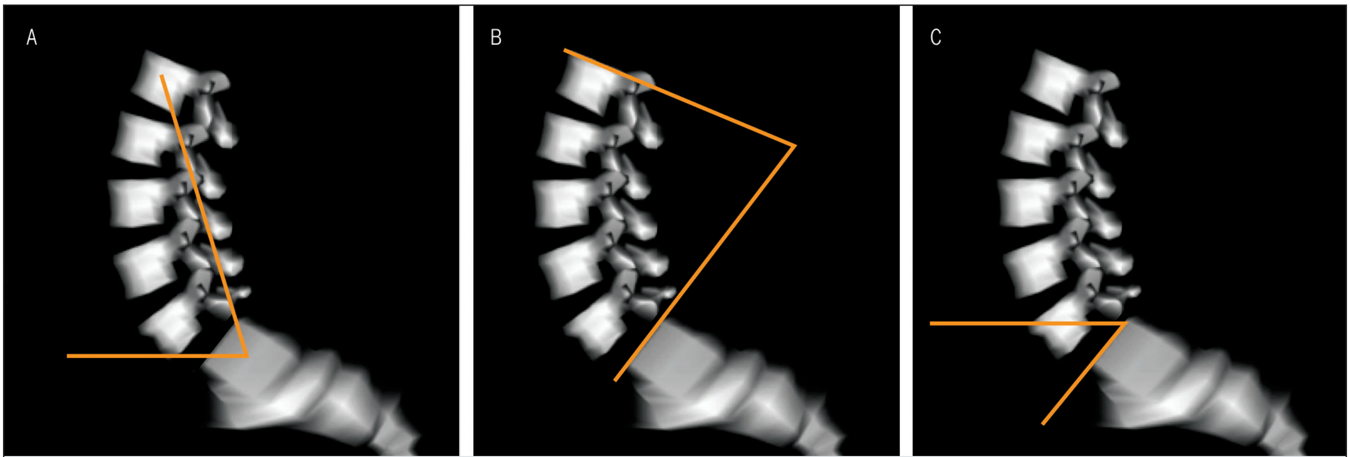


FIGURE 1. Schematic depicting lumbar spine postural measurements on a 3-dimensional model of the lumbar spine. Measurements include (A) angle with respect to the horizontal to assess lumbar flexion/extension, (B) sagittal Cobb angle to measure lumbar lordosis, and (C) sacral slope to assess rotation of the pelvis.

size.[5,8,12,35] Mean diffusivity describes the average restricted diffusion coefficient of $\lambda_{1\text{-}3}$ and is normally between 1×10^{-3} mm^2/s and 2×10^{-3} mm^2/s.[24] Fractional anisotropy is a unitless measurement from 0 to 1 that indicates the shape of the diffusion tensor. An FA value of 0 corresponds to isotropic diffusion (unrestricted), and an FA value of 1 corresponds to diffusion along a line (highly restricted). The eigenvalues ($\lambda_{1\text{-}3}$) define the magnitude of diffusion along (λ_1) and radial to ($\lambda_{2,3}$) the main direction of the muscle fiber.

The T2 values for each IVD were estimated by fitting the magnitude of the multiecho data to a monoexponential decay: $S_i = S_o e^{-t/T2}$.

Intervertebral disc health is often assessed by qualitatively assessing disc hydration from T2-weighted MRI scans. Quantitative T2 mapping provides a quantitative measurement of IVD hydration; T2 is inversely proportional to Pfirrmann grade, which is a common ordinal scale to assess IVD degeneration.[41]

Statistical Analysis

Dependent variables were global postural measurements (angle with respect to the horizontal, sagittal Cobb angle, and sacral angle) for all positions (standing unloaded and standing, sitting, and prone on elbows with load) and the change in load and flexion/extension positions (delta load, delta position). To assess variance, a coefficient of variation was calculated for each dependent and independent variable.

An a priori approach was used to minimize the number of independent variables input into each model (**FIGURE 2**). First, independent variables were empirically grouped into 3 separate domains: muscle structure (volume, FF, FA, mean diffusivity, and $\lambda_{1\text{-}3}$), IVD health (T2 relaxation of each disc), and anthropometric (age, weight, height, and body mass index [BMI][43]) measures. Hierarchical cluster analysis was used to verify domain groupings. Within each domain grouping, an additional hierarchical analysis was performed. Variables that did not cluster were entered into a stepwise multiple linear regression model for each dependent variable to identify physiologic measures predictive of lumbar spine posture.

Variables that did cluster were then sorted into like variables (eigenvectors), using principal-components analysis (PCA). Within each eigenvector, the Pearson correlation coefficient was used to remove collinear variables (r>0.80). For collinear variables, the variable with the smallest eigenvector value was removed to avoid redundancy of variance across variables. Collinearity was also verified at this point by the variance inflation factor; any variable that had a variance inflation factor greater than 10 was removed from the model. Remaining variables were then entered into the stepwise multiple linear regression model for each dependent variable. A stepwise multiple linear regression was run for each individual dependent variable (18 models: 6 positions by 3 postural measurements). Statistical analyses were performed using SPSS Version 20.0 (IBM Corporation, Armonk, NY).

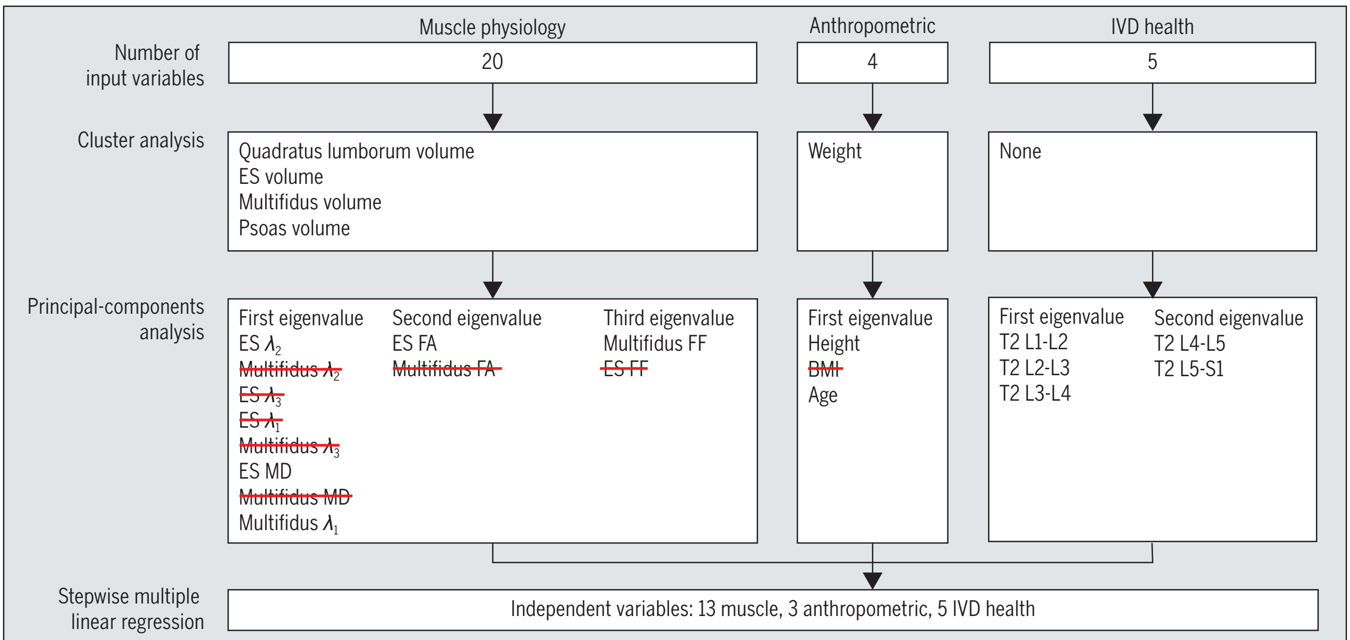


FIGURE 2. Schematic depicting the reduction of collinear independent variables for input into the stepwise multiple regression model. Initially, models were sorted into measures of muscle physiology, anthropometric measures, and IVD health. Cluster analysis was used to identify similar measures. For similar variables, principal-components analysis was used to separate like variables into groups (components). Within each component, Pearson correlations were used to identify collinear variables. If 2 variables were collinear (r>0.80 or variance inflation factor greater than 10), then the variable with the weaker contribution to the eigenvector was removed (crossed out). Abbreviations: BMI, body mass index; ES, erector spinae; FA, fractional anisotropy; FF, fat fraction; IVD, intervertebral disc; λ, eigenvalue; MD, mean diffusivity.

RESULTS

Volunteer Demographics

Forty-three male Marines (mean ± SD age, 26.8 ± 6.4 years; height, 1.8 ± 0.1 m; weight, 82.0 ± 9.9 kg) volunteered for this study. Two subjects dropped out during supine imaging due to claustrophobia in the MRI scanner. Additionally, DTI data sets of 10 subjects were deemed unusable due to breathing or motion artifact. Therefore, 31 Marines were included in this analysis (mean ± SD age, 27.3 ± 6.9 years; height, 1.8 ± 0.1 m; weight, 80.6 ± 8.7 kg). Marines excluded from the study had no differences in anthropometric measures compared with those included. Of these volunteers, 10 Marines self-reported experiencing LBP at the time of the scan.

Coefficients of variation were relatively low for dependent and independent variables (range, 0.04-10.61; median, 0.16) (**APPENDIX**, available at www.jospt.org). On average, the greatest variation was found for the IVD health measures.

Regression Model

After initial grouping of independent variables, collinearity resulted in the removal of 8 of the 29 independent variables from the model (**FIGURE 2**). Collinear variables that were removed included diffusion measurements from either the multifidus or erector spinae, erector spinae FF, and BMI. Surprisingly, 9 of 18 dependent variables were found from the stepwise multiple linear regressions to have a significant predictor. In fact, FA of the erector spinae was a significant predictor of lumbar posture for 7 of the 18 dependent variables measured, and explained 20% to 35% of the variance for each outcome (**TABLE**). In general, increased FA in the erector spinae was predictive of increased lumbar lordosis, lumbosacral extension, and pelvic tilt in each position. Additionally, decreased T2 relaxation of the L4-L5 IVD was a significant predictor of increased lumbosacral extension when standing unloaded (P = .025, R^2 = 0.192). When prone on elbows, increasing subject weight was a significant predictor of increased lumbar lordosis (P = .016, R^2 = 0.219). No muscle volume, muscle microstructure, IVD health, or anthropometric measures were significant predictors of posture when subjects were sitting loaded.

TABLE Results From Stepwise Multiple Linear Regression

Dependent Variable	Significant Independent Variable	β*	R^2	P Value
Cobb angle				
Standing unloaded	None			
Standing loaded	ES FA	0.453	0.205	.02
Sitting loaded	None			
Prone on elbows loaded	Weight	0.468	0.219	.016
Delta load†	None			
Delta position‡	None			
Angle with respect to horizontal				
Standing unloaded	T2 L4-L5	−0.439	0.192	.025
Standing loaded	ES FA	0.514	0.264	.007
Sitting loaded	None			
Prone on elbows loaded	ES FA	−0.480	0.23	.013
Delta load†	None			
Delta position‡	ES FA	0.455	0.207	.02
Sacral angle				
Standing unloaded	ES FA	0.442	0.195	.024
Standing loaded	ES FA	0.587	0.345	.002
Sitting loaded	None			
Prone on elbows loaded	ES FA	0.562	0.316	.003
Delta load†	None			
Delta position‡	None			

Abbreviations: ES, erector spinae; FA, fractional anisotropy.
**Standardized coefficient.*
†Standing unloaded to standing loaded.
‡Sitting loaded to prone on elbows loaded.

DISCUSSION

In this study, we evaluated the relationship between lumbar spine posture and muscle structure, IVD health, and anthropometric measures in 31 active-duty male Marines in simulated, relevant, operational positions and loading conditions. Fractional anisotropy of the erector spinae was a significant predictor in 7 of the 18 measures of lumbar spine posture across several different positions. For the standing loaded condition, FA of the erector spinae was a significant predictor of all 3 measures of lumbar posture; Marines with increased FA of the erector spinae had a more lordotic, extended lumbar posture with greater sacral tilt. Muscle volume was not a significant predictor of any postural measurements, despite being a commonly used proxy for muscle strength.[10,19] Together, the ability of FA to predict postural behavior in several positions and the absence of association between muscle volume and lumbar spine posture suggest that muscle microstructure, but not quantity—both measures associated with force-generating capacity of muscle—is an important predictor of lumbar spine posture.

Diffusion tensor imaging is an MRI technique that measures the restricted diffusion of water in tissues with anisotropic microstructure.[1] As the sarcolemma is considered to be the primary barrier to diffusion, DTI is believed to be most sensitive to changes in fiber size, because radial diffusion of water across a muscle fiber is more restricted (by the sarcolemma) than longitudinal diffusion within a muscle fiber.[44,45] While it has been shown that FA and fiber area are inversely related,[2,5,8,12,35] it is important to note that the exact relationship has not been validated. However, it is well established that muscle fiber area and isometric force are directly related.[17,21,22] Therefore, it appears that there is likely an inverse relationship between FA and isometric force-generating capacity of muscle. As such, it is inferred that when FA increases, the force-generating capacity of a muscle decreases (ie, the muscle is weaker). For example, if the multifidus muscles in 2 Marines were imaged using DTI and 1 had a larger FA (smaller fiber size), that muscle would be expected to generate less overall force.

Two unique relationships between posture and muscle structure were found in this study: (1) the erector spinae, not the multifidus, and (2) muscle microstructure, not volume, were found to be significant predictors of lumbar posture. First, FA of the multifidus and FA of the erector spinae were found to be collinear, with FA of the erector spinae being a stronger descriptor of the eigenvector from the PCA. Therefore, the multifidus was not included in the final statistical model. To verify that FA of the multifidus was not removed from the model because it had less variability than FA of the erector spinae, a coefficient of variation was calculated for both variables. Fractional anisotropy of the erector spinae had less variability relative to the mean than did FA of the multifidus (0.07 versus 0.08), further supporting the latter as a stronger descriptor of the eigenvector. While there is a small difference in variability of these measures, the variability values are both greater than the associated measurement error (0.03 and 0.04, respectively). This finding suggests that while the multifidus stabilizes the individual segments of the spinal column,[46,47] the erector spinae may play a role in determining gross lumbar posture.

Second, while muscle volume is proportional to muscle strength,[17,27] muscle microstructure has been shown to be a more accurate predictor of muscle force-generating capacity. Clinically, the findings from this study are important because they suggest that microstructural quality of the lumbar muscles is more important to whole lumbar posture in functionally loaded positions than the quantity or volume of muscle. This is not surprising given that measures of whole muscle size and volume are confounded by noncontractile tissue, such as fat and fibrosis. Importantly, FA may be a noninvasive composite measure of the functional contractile tissue present in a whole muscle, which seems to explain much of the variance in postural responses to body position.

In this study, T2 of the L4-L5 IVD was found to be inversely proportional to lumbosacral extension when Marines were standing without load. This suggests that Marines with decreased IVD T2 values (increased IVD degeneration) at L4-L5 have increased lumbosacral extension. Previously, using the Pfirrmann grading scale, the authors[4] reported no significant difference in lumbosacral extension in Marines when categorized by degeneration at L5-S1 (Pfirrmann grade greater than 2). As L5-S1 is the base of support of the lumbar spine, it was assumed that degeneration at this level would have whole lumbar postural consequences. However, our findings demonstrate that health of the L4-L5 IVD is related to whole lumbar posture and, therefore, should be considered an important structural level for whole lumbar stability. The finding that single-level disc health has the potential to influence lumbosacral flexion highlights the importance of the lower lumbar spine as a transition zone of load between the trunk and body. Changes to the health of this region have the potential to affect support of the torso.

Several studies have previously attempted to determine the relationship between lumbar lordosis and BMI. It appears that increased lumbar lordosis might be found in individuals with increased BMI[11,23]; however, other studies have shown no difference.[49] In this study, weight and BMI were found to be collinear, with weight being the stronger predictor of the eigenvector from PCA; therefore, BMI was dropped from the final statistical model. However, this is likely due to a larger variance in subject weight rather than in BMI in this relatively homogeneous population. If a more representative cross-section of the population were used, then these findings may have been different.

In this study, the researchers made several attempts to decrease the complexity of the model to decrease the amount of type I error that can be associated with making multiple comparisons. First, this study does not include individual vertebral-level measures of muscle structure or lumbar posture. Second, the authors removed collinear variables with clustering and PCA to minimize the number of independent variables representing similar constructs that were entered into the model. Third, this study evaluated forward, backward, and stepwise multiple linear regression models to determine which model was the most conservative approach. Results were the same with forward and stepwise elimination techniques, and backward elimination allowed for several more independent variables to be retained in the model, suggesting that it was the least conservative regression approach. Therefore, the authors chose to use a stepwise multiple linear regression technique, as it appeared to be the most conservative model.

The Marines in this study were not recruited based on history or presence of LBP at the time of the study, and approximately one third of the Marines who were included in this study reported LBP. It is important to note that no Marines had an episode of LBP so severe that they were relieved of duty. In a previous study, no difference in lumbar spine posture was found between Marines with and without

LBP at the time of data collection.[3] No differences have been observed between Marines with and without LBP at the time of data collection for muscle physiology, IVD health, or anthropometric measures (data not published). As LBP did not result in differences in the dependent or independent variables measured, it is unlikely that the inclusion of Marines with and without LBP affected the findings of this study.

There are several limitations to this study. First, the Marines had relatively normal muscle, with no underlying pathology observable. In patients with pathology or age-related atrophic changes in muscle, the volume or FF of muscle may be more important in predicting lumbar posture. Therefore, the results of this study may only extend to a highly active population. Second, the positions measured in this study place relatively small challenges on the muscles of the lumbar spine. A future direction of this research is to investigate whether muscle microstructure can predict posture, given the heavy loading conditions under which Marines routinely operate.

Finally, the model used in the present study incorporated 21 variables, with only 31 full data sets to include. This was a retrospective analysis of 2 studies investigating (1) the effect of operationally relevant positions on lumbar posture[3] and (2) normative paraspinal muscle composition in active-duty Marines. It was determined that 43 participants were needed to provide adequate power to these studies. However, to mitigate type I error associated with multiple comparisons, the authors used the most conservative statistical approach. While more participants may provide an increase in the amount of variance explained by the model, this study still reached significance with 31 complete data sets.

CONCLUSION

The authors believe that this study is the first to measure the predictive capacity of lumbar muscle structure, IVD health, and anthropometric measures on lumbar spine posture in different positions. It is surprising that any structural variable in muscle predicted any of the variance in posture, because many clinicians believe that short-term postural positions are more related to motor control than to strength or end organ–dependent behavior.

This study found that FA of the erector spinae was a significant predictor of several lumbar postural measures. In general, decreased FA of the erector spinae resulted in decreased lordosis, lumbosacral extension, and anterior pelvic tilt. This posture results in decreased shear stress at lower lumbar levels during hyperlordosis and may be considered a more protective posture for preventing injury and LBP when loading the lumbar spine.[37] Decreased FA of the erector spinae can be physiologically interpreted as larger muscle fibers with more capacity to generate force. Due to the intense training and demands of their jobs, the Marines in this study were extremely active and trained on how to adapt their posture in different positions, while wearing body armor, to minimize their risk of injury. Therefore, these findings may not translate to a civilian population.

The findings of this study support the idea that muscle strengthening/exercise may influence posture, although this cause-and-effect relationship needs to be substantiated in prospective clinical research. As this relationship was found in a healthy population with relatively little variance in muscle quality, it is likely that these relationships may be stronger in patients with LBP or injury. Understanding the influence of microstructural features of muscle on posture may allow clinicians to prognostically categorize patients into groups that may respond better to exercise-based treatments. Future studies should take a more controlled approach to determine whether targeted exercise of the erector spinae muscles increases muscle quality (measured with DTI) and can elicit a postural response. ◉

KEY POINTS

FINDINGS: Fractional anisotropy of the erector spinae was a significant predictor of lumbar lordosis, lumbar flexion, and sacral tilt in several different operationally relevant positions in active-duty Marines.

IMPLICATIONS: The finding that fractional anisotropy can predict postural responses in several positions, along with the absence of association between muscle volume and lumbar spine posture, suggests that muscle microstructure, but not quantity, is an important predictor of lumbar spine posture.

CAUTION: These findings were found in a group of highly active Marines and may not translate to a civilian population.

ACKNOWLEDGMENTS: *The authors thank the Marines from the 1st and 5th Regiments who supported this effort.*

REFERENCES

1. Basser PJ, Mattiello J, LeBihan D. MR diffusion *tensor* spectroscopy and imaging. *Biophys J*. 1994;66:259-267. https://doi.org/10.1016/S0006-3495(94)80775-1
2. Berry DB, Regner B, Galinsky V, Ward SR, Frank LR. Relationships between tissue microstructure and the diffusion tensor in simulated skeletal muscle. *Magn Reson Med*. 2018;80:317-329. https://doi.org/10.1002/mrm.26993
3. Berry DB, Rodríguez-Soto AE, Su J, et al. Lumbar spine postures in Marines during simulated operational positions. *J Orthop Res*. 2017;35:2145-2153. https://doi.org/10.1002/jor.23510
4. Berry DB, Rodríguez-Soto AE, Tokunaga JR, Gombatto SP, Ward SR. An endplate-based joint coordinate system for measuring kinematics in normal and abnormally-shaped lumbar vertebrae. *J Appl Biomech*. 2015;31:499-503. https://doi.org/10.1123/jab.2015-0008
5. Berry DB, You S, Warner J, Frank LR, Chen S, Ward SR. A 3D tissue-printing approach for validation of diffusion tensor imaging in skeletal muscle. *Tissue Eng Part A*. 2017;23:980-988. https://doi.org/10.1089/ten.tea.2016.0438
6. Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res*. 1996;29:162-173. https://doi.org/10.1006/cbmr.1996.0014
7. Demoulin C, Crielaard JM, Vanderthommen M. Spinal muscle evaluation in healthy individuals and low-back-pain patients: a literature review.

Joint Bone Spine. 2007;74:9-13. https://doi.org/10.1016/j.jbspin.2006.02.013

8. Deux JF, Malzy P, Paragios N, et al. Assessment of calf muscle contraction by diffusion tensor imaging. *Eur Radiol*. 2008;18:2303-2310. https://doi.org/10.1007/s00330-008-1012-z
9. Fortin M, Macedo LG. Multifidus and paraspinal muscle group cross-sectional areas of patients with low back pain and control patients: a systematic review with a focus on blinding. *Phys Ther*. 2013;93:873-888. https://doi.org/10.2522/ptj.20120457
10. Fukunaga T, Miyatani M, Tachi M, Kouzaki M, Kawakami Y, Kanehisa H. Muscle volume is a major determinant of joint torque in humans. *Acta Physiol Scand*. 2001;172:249-255. https://doi.org/10.1046/j.1365-201x.2001.00867.x
11. Guo JM, Zhang GQ, Alimujiang. [Effect of BMI and WHR on lumbar lordosis and sacrum slant angle in middle and elderly women]. *Zhongguo Gu Shang*. 2008;21:30-31.
12. Hatakenaka M, Yabuuchi H, Sunami S, et al. Joint position affects muscle proton diffusion: evaluation with a 3-T MR system. *AJR Am J Roentgenol*. 2010;194:W208-W211. https://doi.org/10.2214/AJR.09.3167
13. Keorochana G, Taghavi CE, Lee KB, et al. Effect of sagittal alignment on kinematic changes and degree of disc degeneration in the lumbar spine: an analysis using positional MRI. *Spine (Phila Pa 1976)*. 2011;36:893-898. https://doi.org/10.1097/BRS.0b013e3181f4d212
14. Kim YB, Kim YJ, Ahn YJ, et al. A comparative analysis of sagittal spinopelvic alignment between young and old men without localized disc degeneration. *Eur Spine J*. 2014;23:1400-1406. https://doi.org/10.1007/s00586-014-3236-8
15. Knapik JJ, Reynolds KL, Harman E. Soldier load carriage: historical, physiological, biomechanical, and medical aspects. *Mil Med*. 2004;169:45-56. https://doi.org/10.7205/MILMED.169.1.45
16. Lao L, Daubs MD, Scott TP, et al. Effect of disc degeneration on lumbar segmental mobility analyzed by kinetic magnetic resonance imaging. *Spine (Phila Pa 1976)*. 2015;40:316-322. https://doi.org/10.1097/BRS.0000000000000738
17. Lieber RL. *Skeletal Muscle Structure, Function, and Plasticity: The Physiological Basis of Rehabilitation*. 2nd ed. Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins; 2002.
18. Mac-Thiong JM, Roussouly P, Berthonnaud É, Guigui P. Age- and sex-related variations in sagittal sacropelvic morphology and balance in asymptomatic adults. *Eur Spine J*. 2011;20 suppl 5:572-577. https://doi.org/10.1007/s00586-011-1923-2
19. Meakin JR, Fulford J, Seymour R, Welsman JR, Knapp KM. The relationship between sagittal curvature and extensor muscle volume in the lumbar spine. *J Anat*. 2013;222:608-614. https://doi.org/10.1111/joa.12047
20. Meyer GA, Lieber RL. Elucidation of extracellular matrix mechanics from muscle fibers and fiber bundles. *J Biomech*. 2011;44:771-773. https://doi.org/10.1016/j.jbiomech.2010.10.044
21. Miller AE, MacDougall JD, Tarnopolsky MA, Sale DG. Gender differences in strength and muscle fiber characteristics. *Eur J Appl Physiol Occup Physiol*. 1993;66:254-262. https://doi.org/10.1007/BF00235103
22. Minamoto VB, Hulst JB, Lim M, et al. Increased efficacy and decreased systemic-effects of botulinum toxin A injection after active or passive muscle manipulation. *Dev Med Child Neurol*. 2007;49:907-914. https://doi.org/10.1111/j.1469-8749.2007.00907.x
23. Murrie VL, Dixon AK, Hollingworth W, Wilson H, Doyle TA. Lumbar lordosis: study of patients with and without low back pain. *Clin Anat*. 2003;16:144-147. https://doi.org/10.1002/ca.10114
24. Oudeman J, Nederveen AJ, Strijkers GJ, Maas M, Luijten PR, Froeling M. Techniques and applications of skeletal muscle diffusion tensor imaging: a review. *J Magn Reson Imaging*. 2016;43:773-788. https://doi.org/10.1002/jmri.25016
25. Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine (Phila Pa 1976)*. 2001;26:1873-1878. https://doi.org/10.1097/00007632-200109010-00011
26. Pontén E, Fridén J, Thornell LE, Lieber RL. Spastic wrist flexors are more severely affected than wrist extensors in children with cerebral palsy. *Dev Med Child Neurol*. 2005;47:384-389. https://doi.org/10.1111/j.1469-8749.2005.tb01158.x
27. Powell PL, Roy RR, Kanim P, Bello MA, Edgerton VR. Predictability of skeletal muscle tension from architectural determinations in guinea pig hindlimbs. *J Appl Physiol Respir Environ Exerc Physiol*. 1984;57:1715-1721. https://doi.org/10.1152/jappl.1984.57.6.1715
28. Reynolds KL, Kaszuba J, Mello RP, Patton JF. Prolonged Treadmill Load Carriage: Acute Injuries and Changes in Foot Anthropometry. Natick, MA: US Army Research Institute of Environmental Medicine; 1990.
29. Rodriguez-Soto AE, Berry DB, Jaworski R, et al. The effect of training on lumbar spine posture and intervertebral disc degeneration in active-duty Marines. *Ergonomics*. 2017;60:1055-1063. https://doi.org/10.1080/00140139.2016.1252858
30. Rodríguez-Soto AE, Berry DB, Palombo L, Valaik E, Kelly KR, Ward SR. Effect of load magnitude and distribution on lumbar spine posture in active-duty Marines. *Spine (Phila Pa 1976)*. 2017;42:345-351. https://doi.org/10.1097/BRS.0000000000001742
31. Rodríguez-Soto AE, Jaworski R, Jensen A, et al. Effect of load carriage on lumbar spine kinematics. *Spine (Phila Pa 1976)*. 2013;38:E783-E791. https://doi.org/10.1097/BRS.0b013e3182913e9f
32. Rosset A, Spadola L, Ratib O. OsiriX: an open-source software for navigating in multidimensional DICOM images. *J Digit Imaging*. 2004;17:205-216. https://doi.org/10.1007/s10278-004-1014-6
33. Roy TC, Knapik JJ, Ritland BM, Murphy N, Sharp MA. Risk factors for musculoskeletal injuries for soldiers deployed to Afghanistan. *Aviat Space Environ Med*. 2012;83:1060-1066. https://doi.org/10.3357/ASEM.3341.2012
34. Roy TC, Lopez HP, Piva SR. Loads worn by soldiers predict episodes of low back pain during deployment to Afghanistan. *Spine (Phila Pa 1976)*. 2013;38:1310-1317. https://doi.org/10.1097/BRS.0b013e31829265c4
35. Schwenzer NF, Steidle G, Martirosian P, et al. Diffusion tensor imaging of the human calf muscle: distinct changes in fractional anisotropy and mean diffusion due to passive muscle shortening and stretching. *NMR Biomed*. 2009;22:1047-1053. https://doi.org/10.1002/nbm.1409
36. Shahidi B, Parra CL, Berry DB, et al. Contribution of lumbar spine pathology and age to paraspinal muscle size and fatty infiltration. *Spine (Phila Pa 1976)*. 2017;42:616-623. https://doi.org/10.1097/BRS.0000000000001848
37. Shirazi-Adl A, El-Rich M, Pop DG, Parnianpour M. Spinal muscle forces, internal loads and stability in standing under various postures and loads—application of kinematics-based algorithm. *Eur Spine J*. 2005;14:381-392. https://doi.org/10.1007/s00586-004-0779-0
38. Sinaki M, Itoi E, Rogers JW, Bergstralh EJ, Wahner HW. Correlation of back extensor strength with thoracic kyphosis and lumbar lordosis in estrogen-deficient women. *Am J Phys Med Rehabil*. 1996;75:370-374. https://doi.org/10.1097/00002060-199609000-00013
39. Smith LR, Lee KS, Ward SR, Chambers HG, Lieber RL. Hamstring contractures in children with spastic cerebral palsy result from a stiffer extracellular matrix and increased *in vivo* sarcomere length. *J Physiol*. 2011;589:2625-2639. https://doi.org/10.1113/jphysiol.2010.203364
40. Sparrey CJ, Bailey JF, Safaee M, et al. Etiology of lumbar lordosis and its pathophysiology: a review of the evolution of lumbar lordosis, and the mechanics and biology of lumbar degeneration. *Neurosurg Focus*. 2014;36:E1. https://doi.org/10.3171/2014.1.FOCUS13551
41. Stelzeneder D, Welsch GH, Kovács BK, et al. Quantitative T2 evaluation at 3.0 T compared to morphological grading of the lumbar intervertebral disc: a standardized evaluation approach in patients with low back pain. *Eur J Radiol*. 2012;81:324-330. https://doi.org/10.1016/j.ejrad.2010.12.093
42. Sullivan MS, Dickinson CE, Troup JD. The influence of age and gender on lumbar spine sagittal plane range of motion. A study of 1126 healthy subjects. *Spine (Phila Pa 1976)*. 1994;19:682-686. https://doi.org/10.1097/00007632-199403001-00007
43. Takatalo J, Karppinen J, Taimela S, et al. Body mass index is associated with lumbar disc

degeneration in young Finnish males: subsample of Northern Finland birth cohort study 1986. *BMC Musculoskelet Disord*. 2013;14:87. https://doi.org/10.1186/1471-2474-14-87

44. Van Donkelaar CC, Kretzers LJ, Bovendeerd PH, et al. Diffusion tensor imaging in biomechanical studies of skeletal muscle function. *J Anat*. 1999;194 pt 1:79-88. https://doi.org/10.1046/j.1469-7580.1999.19410079.x

45. van Doorn A, Bovendeerd PH, Nicolay K, Drost MR, Janssen JD. Determination of muscle fibre orientation using Diffusion-Weighted MRI. *Eur J Morphol*. 1996;34:5-10.

46. Ward SR, Kim CW, Eng CM, et al. Architectural analysis and intraoperative measurements demonstrate the unique design of the multifidus muscle for lumbar spine stability. *J Bone Joint Surg Am*. 2009;91:176-185. https://doi.org/10.2106/JBJS.G.01311

47. Ward SR, Tomiya A, Regev GJ, et al. Passive mechanical properties of the lumbar multifidus muscle support its role as a stabilizer. *J Biomech*. 2009;42:1384-1389. https://doi.org/10.1016/j.jbiomech.2008.09.042

48. Wokke BH, van den Bergen JC, Versluis MJ, et al. Quantitative MRI and strength measurements in the assessment of muscle quality in Duchenne muscular dystrophy. *Neuromuscul Disord*. 2014;24:409-416. https://doi.org/10.1016/j.nmd.2014.01.015

49. Youdas JW, Hollman JH, Krause DA. The effects of gender, age, and body mass index on standing lumbar curvature in persons without current low back pain. *Physiother Theory Pract*. 2006;22:229-237. https://doi.org/10.1080/09593980600927864

APPENDIX

Coefficient of Variation Calculated for Each Dependent Variable

Position	Angle With Respect to Horizontal	Sacral Angle	Cobb Angle
Standing unloaded	0.05	0.17	0.18
Standing loaded	0.05	0.22	0.22
Sitting loaded	0.04	0.48	0.16
Prone on elbows loaded	0.06	0.26	0.16
Delta load	10.61	1.86	4.11
Delta position	1.18	1.30	0.37

Coefficient of Variation Calculated for Each Independent Variable

Independent Variable	Coefficient of Variation
Muscle measures	
Multifidus	
Volume	0.14
Fat fraction	0.41
Mean diffusivity	0.05
Fractional anisotropy	0.08
Lambda 1	0.04
Lambda 2	0.04
Lambda 3	0.06
Erector spinae	
Volume	0.22
Fat fraction	0.41
Mean diffusivity	0.05
Fractional anisotropy	0.07
Lambda 1	0.04
Lambda 2	0.04
Lambda 3	0.05
Psoas volume	0.13
Quadratus lumborum volume	0.19
IVD measures	
T2	
L1-L2	0.24
L2-L3	0.27
L3-L4	0.29
L4-L5	0.35
L5-S1	0.41
Anthropometric measures	
Age	0.24
Height	0.04
Weight	0.12
Body mass index	0.11

Abbreviation: IVD, intervertebral disc.

DAVID B. BERRY, PhD[1] • **BAHAR SHAHIDI,** PT, PhD[2,3] • **ANA E. RODRÍGUEZ-SOTO,** PhD[1]
JAN M. HUGHES-AUSTIN, PT, PhD[3] • **KAREN R. KELLY,** PhD[4] • **SAMUEL R. WARD,** PT, PhD[1-3]

Lumbar Muscle Structure Predicts Operational Postures in Active-Duty Marines

The muscles of the lumbar spine are crucial for stabilizing and supporting the upper trunk, especially during dynamic loading conditions. A muscle's force-generating capacity is directly related to its architectural and microstructural features, which are therefore variables of interest when trying to assess muscle health. Physiological cross-sectional area (PCSA) is a measure of

BACKGROUND: The relationship between lumbar spine posture and muscle structure is not well understood.

OBJECTIVES: To investigate the predictive capacity of muscle structure on lumbar spine posture in active-duty Marines.

METHODS: Forty-three Marines were scanned in this cross-sectional study, using an upright magnetic resonance imaging scanner while standing without load and standing, sitting, and prone on elbows with body armor. Cobb, horizontal, and sacral angles were measured. Marines were then scanned while unloaded in supine using a supine magnetic resonance imaging scanner. The imaging protocol consisted of T2 intervertebral disc mapping; high-resolution, anatomical, fat-water separation, and diffusion tensor imaging to quantify disc hydration and muscle volume, fat fraction, and restricted diffusion profiles in the lumbar muscles. A stepwise multiple linear regression model was used to identify physiological measures predictive of lumbar spine posture.

RESULTS: The multiple regression model demonstrated that fractional anisotropy of the erector spinae was a significant predictor of lumbar posture for 7 of 18 dependent variables measured, and explained 20% to 35% of the variance in each model. Decreased fractional anisotropy of the erector spinae predicted decreased lordosis, lumbosacral extension, and anterior pelvic tilt.

CONCLUSION: Fractional anisotropy is inversely related with muscle fiber size, which is associated with the isometric force-generating capacity of a muscle fiber. This suggests that stronger erector spinae muscles predict decreased lordosis, lumbosacral extension, and anterior pelvic tilt in a highly trained population. *J Orthop Sports Phys Ther 2018;48(8):613-621. Epub 17 May 2018. doi:10.2519/jospt.2018.7865*

KEY WORDS: *diffusion tensor imaging, lumbar spine, magnetic resonance imaging, military, posture, skeletal muscle*

muscle architecture that can be measured to estimate muscle force.[27] However, it is difficult to precisely measure PCSA in vivo, as it includes measures of muscle architecture, such as pennation angle and normalized fiber length. Volume is a dominant input variable to measure muscle PCSA and is commonly used as a proxy for muscle force-producing capacity.[7,9] However, muscle is a heterogeneous tissue, also consisting of fat and collagenous tissues, which can confound measures of muscle volume.

Skeletal muscle exhibits a classic structure-function relationship, where its microstructural properties are closely related to whole muscle function. For example, muscle fiber isometric force-generating capacity is directly related to fiber cross-sectional area.[17,21,22] It is also difficult to measure muscle microstructure in vivo, although there is some evidence that diffusion-based imaging techniques are sensitive to different features of muscle microstructure, in particular fiber area.[5,8,12,35]

With injury and age, atrophy of muscle fibers and replacement of muscle tis-

[1]Department of Bioengineering, University of California, San Diego, La Jolla, CA. [2]Department of Radiology, University of California, San Diego, La Jolla, CA. [3]Department of Orthopaedic Surgery, University of California, San Diego, La Jolla, CA. [4]Department of Warfighter Performance, US Naval Health Research Center, San Diego, CA. The University of California, San Diego (UCSD) and US Naval Health Research Center (NHRC) Institutional Review Boards approved this study (UCSD protocol 110483; NHRC.2013.0023). The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or reflecting the views of the US Navy, US Army, US Air Force, the US Departments of Defense and Veterans Affairs, or the US Government. Any references to commercial organizations and trade names in this article do not constitute an official US Army endorsement or approval of the products or services of these organizations. Dr Kelly is a military service member. This work was prepared as part of her official duties. Title 17, USC, §105 provides that "Copyright protection under this title is not available for any work of the United States Government." Title 17, USC, §101 defines a US Government work as work prepared by a military service member or employee of the US Government as part of that person's official duties. The US Army Medical Research Acquisition Activity supported this manuscript through award number W81XWH-13-2-0043 (work unit number 1305). This work was also supported by US Department of Defense grant PR120576 (to Drs Ward and Kelly). Dr Ward is also supported by grants from the National Institutes of Health (R01AR070830, R01HD073180, R01HD088437). 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 Samuel R. Ward, University of California, San Diego, 9500 Gilman Drive (0863), La Jolla, CA 92093. E-mail: s1ward@ucsd.edu

sue with adipose and fibrotic tissue are typically observed compared to healthy muscle, further decreasing the overall volume of functional contractile tissue in the muscle.[7,36] As pathogenic (diseased) muscle becomes atrophied and fibrotic and contains more adipose tissue, the active and passive force-generating potential of the whole muscle changes, which can have a direct and negative effect on joint stability, range of motion, and posture.[20,26,39,48] The multifidus muscle is considered to be one of the primary muscular stabilizers of the lumbar spine, due to its ability to produce high forces over a narrow range of lengths, and often undergoes the pathogenic changes associated with injury, low back pain (LBP), or age.[46]

Changes to the orientation and position of bony structures of the spinal column are often observed simultaneously with these changes in muscle composition.[19,38,40] With age, gross changes in spinal posture, such as decreased lumbar lordosis, increased lumbar flexion, and increased pelvic tilt, are typically observed.[13,14,18,42] Decreased segmental range of motion has also been measured at vertebral levels with intervertebral disc (IVD) degeneration,[13,16] which is defined as decreased hydration of the nucleus pulposus with accompanying disc height loss.[25] However, changes in muscle structure, lumbar posture, and IVD health are not independent of one another, and their effects are confounded by age, sex, activity level, and the timing of disease progression.

In addition to associated changes with age and disease, external stimuli, such as carrying load or whole-body position, may affect posture.[3,29-31] Military members are highly active and often required to carry heavy loads in unusual positions. Studies investigating how Marines adapt to load carriage suggest that they routinely operate under conditions that put them at risk for developing lumbar musculoskeletal injury and that they exhibit higher rates of LBP than civilians.[33,34] This may be attributed to pathophysiologic changes of the lumbar spine structures as a result of the heavy loads and unusual postures experienced in training and combat.[15,28] A noninvasive tool that can correlate musculoskeletal health to posture under relevant loading conditions would allow clinicians to tailor rehabilitation protocols to target specific musculoskeletal components involved in regulating posture to mitigate an individual's risk of lumbar spine injury.

The purpose of this study was to investigate the predictive capacity of muscle structure, IVD health, and anthropometric measures on lumbar spine posture in active-duty Marines. We hypothesized that multifidus muscle volume would predict lumbar posture in different positions, because the multifidus provides intersegmental lumbar support and muscle volume is related to muscle strength.

METHODS

The University of California, San Diego and US Naval Health Research Center Institutional Review Boards approved this study, and all volunteers gave verbal and written consent to participate. Marines were included in this study if they were male, over 18 years of age, and healthy enough to perform their assigned duty. Marines were excluded from this study if they had undergone lumbar spine surgery or had the possibility of shrapnel in their bodies. Marines were not recruited based on LBP status or history. All Marines underwent standard magnetic resonance imaging (MRI) safety screening prior to scanning. All scans were performed early in the morning, between 4 am and 9 am.

Upright MRI

Marines were scanned using an upright 0.6-T MRI scanner (UPRIGHT Multi-Position MRI; Fonar Corporation, Melville, NY) and a planar coil. An elastic band was used to hold the coil against the volunteer's lumbar spine between the L1 and S1 levels while standing. The band was secured to hold the coil in place without altering the volunteer's natural position. A 3-plane localizer (repetition time [TR], 1254 milliseconds; echo time [TE], 100 milliseconds; field of view (FoV), 34 cm; matrix, 256 × 256; in-plane resolution, 1.33 × 1.33 mm; thickness, 9 mm; number of excitations, 1; time, 0:17) and sagittal T2-weighted images (TR, 1974 milliseconds; TE, 160 milliseconds; FoV, 35 cm; matrix, 224 × 224; in-plane resolution, 1.56 × 1.56 mm; thickness, 3 mm; gap, 0 mm; number of excitations, 1; time, 2:12) were acquired.[4]

Upright MRI: Load Carriage and Position Tasks

Marines were scanned in the following positions: standing without load, standing with body armor, sitting with body armor, and prone on elbows with body armor. Positions with external load were randomized to control for the cumulative effects of loading or time. The selected positions were static positions that Marines are often required to maintain for extended periods, depending on military occupational specialty, and are often reported as provoking LBP.[3] The load magnitude of 11.3 kg was chosen based on the use of body armor, which is the minimum protective equipment Marines are required to wear during military operations and training. Marines were not provided instruction on how to assume each position, but were asked to hold each position steady for the duration of the MRI acquisition. A previous study has shown no statistically significant difference in test-retest variation in posture within a subject, even after performing heavy-load and activity tasks.[31]

Upright MRI: Postural Measurements

Postural measurements were generated from upright MRI images in each position, using a previously validated algorithm.[3] Briefly, digital seed points were manually placed on the corners of the vertebral body and on the posterior elements of each vertebra using OsiriX Version 3.9.3 imaging software (Pixmeo

SARL, Bernex, Switzerland).[32] The locations of the seed points were imported into MATLAB (The MathWorks, Inc, Natick, MA) and used to define an end plate–based joint coordinate system applied to the superior and inferior end plate of each vertebra (L1-S1).

Global measurements of lumbar spine posture were calculated for each position to characterize the posture of the lumbar spine. Global measures included angle with respect to the horizontal to assess lumbosacral flexion/extension, sacral slope to assess sacral tilt, and sagittal Cobb angle to assess lumbar lordosis (**FIGURE 1**). Root-mean-square error values for global measurements were measured previously and are 0.28°, 0.95°, and 0.95°, respectively.[4,31] Global measurements between the standing unloaded and the standing loaded (delta load) positions, and between the sitting loaded and prone on elbows loaded (delta position) positions, were also calculated to determine lumbar kinematics in response to load and dynamic movement, respectively.

Supine MRI

Magnetic resonance images of the lumbar spine (L1-S1) were acquired using a 3-T MRI scanner (Discovery MR750; GE Healthcare, Waukesha, WI) and spine array coil. The imaging protocol consisted of (1) an anatomical scan, (2) fat-water separation scan, (3) diffusion tensor imaging (DTI) of the lumbar spine, and (4) T2 mapping of each lumbar IVD. Marines were scanned supine, with the lumbar muscles relaxed, to mitigate motion and breathing artifacts. The anatomical scan was an axial, fast spoiled-gradient echo with the following scanning parameters: TR, 5 milliseconds; TE, 2.3 milliseconds; flip angle, 20°; FoV, 32 cm; acquisition matrix, 512 × 512; pixel size, 0.625 × 0.625 mm^2; slice thickness, 1 mm; no gap; number of averages, 3. Fat-water separation images were acquired utilizing a 3-point iterative decomposition of water and fat, with echo asymmetry and a least-squares estimation sequence in the sagittal plane (TR, 1974 milliseconds; TE, 160 milliseconds; flip angle, 20°; FoV, 25.6 cm; 176 slices; acquisition matrix, 256 × 256; voxel size, 1 × 1 × 1 mm^3; no gap; number of averages, 1). Scanning parameters of the axial DTI sequence were as follows: TR, 10 seconds; TE, 46 milliseconds; FoV, 19.2 cm; 82 slices; acquisition matrix, 128 × 128; pixel size, 1.5 × 1.5 mm^2; slice thickness, 3 mm; no gap; B value, 400 mm^2/s; 45 diffusion directions. Last, multispin-echo data (8 echoes; TE, 8.6 to 68.8 milliseconds; TR, 800 milliseconds; FoV, 16 cm; 5 slices; acquisition matrix, 256 × 256; voxel size, 0.625 × 0.625 × 5 mm^3; no gap; number of averages, 1) were acquired and used to estimate the T2 of each lumbar IVD. The scanning plane was axial oblique, parallel to each lumbar IVD.

Supine MRI: Lumbar Physiology Measurements

Anatomical images were imported into the OsiriX imaging software for segmentation. Contours of the multifidus, erector spinae group, psoas, and quadratus lumborum muscles were manually traced from the L1 to S1 lumbar levels. The resulting segmentations were used to generate masks to quantify muscle volumes, fat fraction, and diffusion properties of Marines in the supine position.

Images acquired using the fat-water separation sequence yielded 2 sets of images: 1 where both fat and water MRI signals are in phase, and 1 where they are out of phase. This allows for isolating the independent contributions of water (S_W) and fat (S_F) to the total MRI signal. These data were then used to quantify the fat fraction (FF) of the multifidus and erector spinae group with the following relationship: $FF = S_F/(S_W + S_F)$.

The diffusion tensor was fitted using Analysis of Functional NeuroImages software (National Institutes of Health, Bethesda, MD) and function 3dDWItoDT.[6] Mean diffusivity, fractional anisotropy (FA), and the 3 eigenvalues (λ_{1-3}) of the diffusion tensor are reported. The quantitative relationship of diffusion variables to specific features of muscle microstructure is the focus of current work, although there is some evidence that they are related to muscle fiber

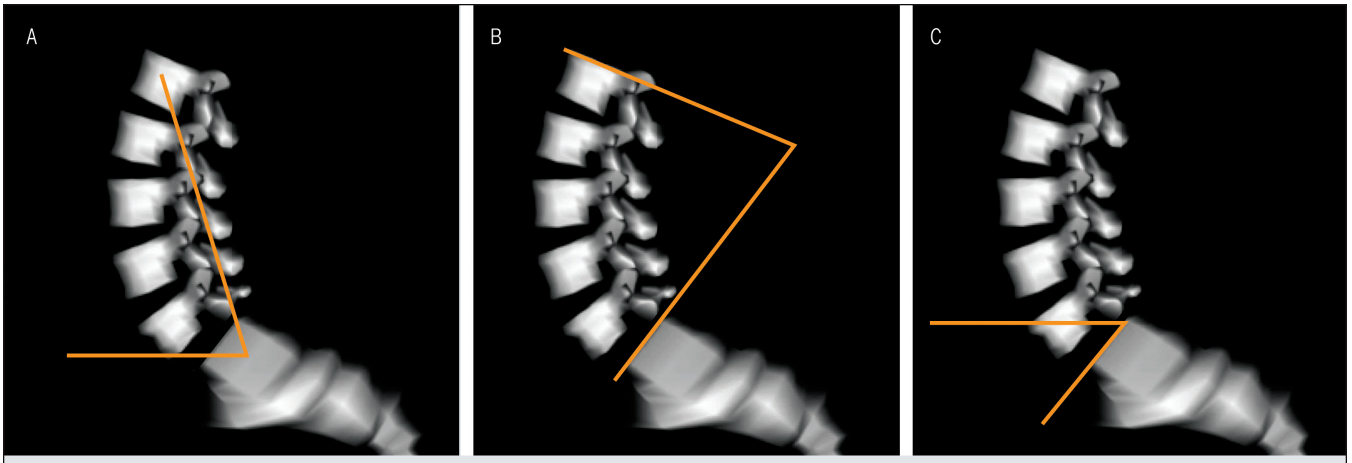


FIGURE 1. Schematic depicting lumbar spine postural measurements on a 3-dimensional model of the lumbar spine. Measurements include (A) angle with respect to the horizontal to assess lumbar flexion/extension, (B) sagittal Cobb angle to measure lumbar lordosis, and (C) sacral slope to assess rotation of the pelvis.

size.[5,8,12,35] Mean diffusivity describes the average restricted diffusion coefficient of $\lambda_{1\text{-}3}$ and is normally between 1×10^{-3} mm^2/s and 2×10^{-3} mm^2/s.[24] Fractional anisotropy is a unitless measurement from 0 to 1 that indicates the shape of the diffusion tensor. An FA value of 0 corresponds to isotropic diffusion (unrestricted), and an FA value of 1 corresponds to diffusion along a line (highly restricted). The eigenvalues ($\lambda_{1\text{-}3}$) define the magnitude of diffusion along (λ_1) and radial to ($\lambda_{2,3}$) the main direction of the muscle fiber.

The T2 values for each IVD were estimated by fitting the magnitude of the multiecho data to a monoexponential decay: $S_i = S_o e^{-t/T2}$.

Intervertebral disc health is often assessed by qualitatively assessing disc hydration from T2-weighted MRI scans. Quantitative T2 mapping provides a quantitative measurement of IVD hydration; T2 is inversely proportional to Pfirrmann grade, which is a common ordinal scale to assess IVD degeneration.[41]

Statistical Analysis

Dependent variables were global postural measurements (angle with respect to the horizontal, sagittal Cobb angle, and sacral angle) for all positions (standing unloaded and standing, sitting, and prone on elbows with load) and the change in load and flexion/extension positions (delta load, delta position). To assess variance, a coefficient of variation was calculated for each dependent and independent variable.

An a priori approach was used to minimize the number of independent variables input into each model (**FIGURE 2**). First, independent variables were empirically grouped into 3 separate domains: muscle structure (volume, FF, FA, mean diffusivity, and $\lambda_{1\text{-}3}$), IVD health (T2 relaxation of each disc), and anthropometric (age, weight, height, and body mass index [BMI][43]) measures. Hierarchical cluster analysis was used to verify domain groupings. Within each domain grouping, an additional hierarchical analysis was performed. Variables that did not cluster were entered into a stepwise multiple linear regression model for each dependent variable to identify physiologic measures predictive of lumbar spine posture.

Variables that did cluster were then sorted into like variables (eigenvectors), using principal-components analysis (PCA). Within each eigenvector, the Pearson correlation coefficient was used to remove collinear variables (r>0.80). For collinear variables, the variable with the smallest eigenvector value was removed to avoid redundancy of variance across variables. Collinearity was also verified at this point by the variance inflation factor; any variable that had a variance inflation factor greater than 10 was removed from the model. Remaining variables were then entered into the stepwise multiple linear regression model for each dependent variable. A stepwise multiple linear regression was run for each individual dependent variable (18 models: 6 positions by 3 postural measurements). Statistical analyses were performed using SPSS Version 20.0 (IBM Corporation, Armonk, NY).

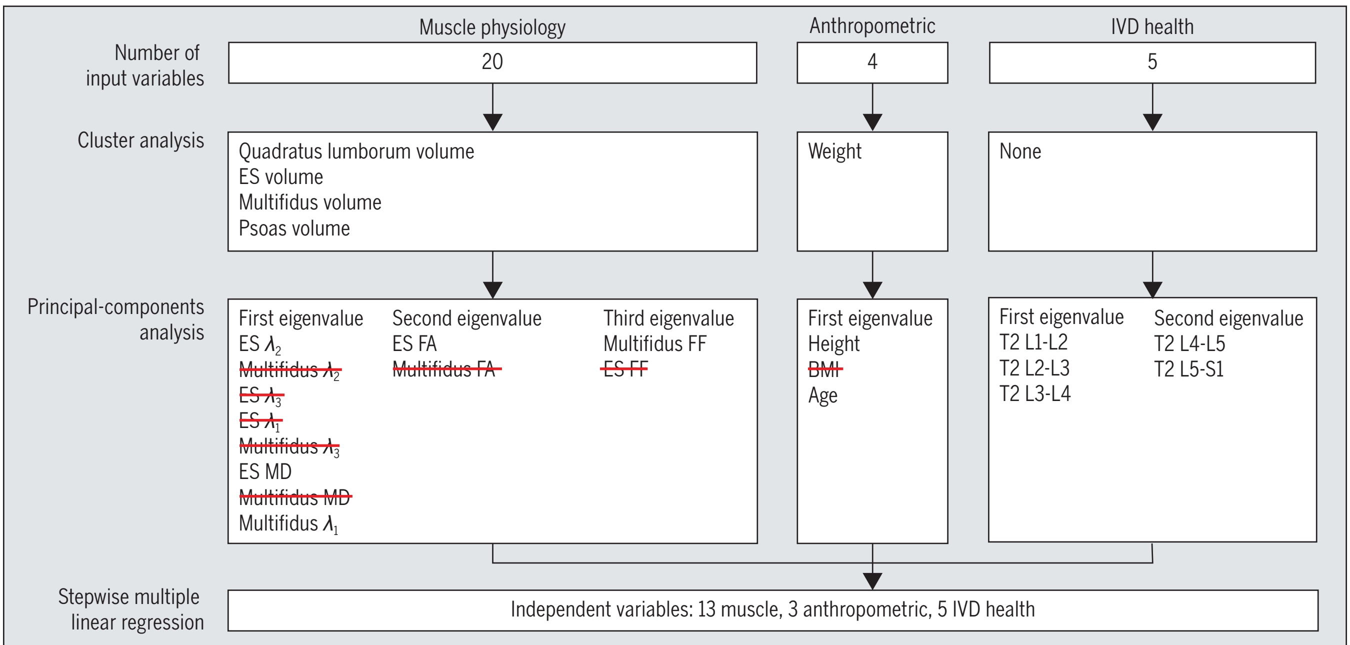


FIGURE 2. Schematic depicting the reduction of collinear independent variables for input into the stepwise multiple regression model. Initially, models were sorted into measures of muscle physiology, anthropometric measures, and IVD health. Cluster analysis was used to identify similar measures. For similar variables, principal-components analysis was used to separate like variables into groups (components). Within each component, Pearson correlations were used to identify collinear variables. If 2 variables were collinear (r>0.80 or variance inflation factor greater than 10), then the variable with the weaker contribution to the eigenvector was removed (crossed out). Abbreviations: BMI, body mass index; ES, erector spinae; FA, fractional anisotropy; FF, fat fraction; IVD, intervertebral disc; λ, eigenvalue; MD, mean diffusivity.

RESULTS

Volunteer Demographics

Forty-three male Marines (mean ± SD age, 26.8 ± 6.4 years; height, 1.8 ± 0.1 m; weight, 82.0 ± 9.9 kg) volunteered for this study. Two subjects dropped out during supine imaging due to claustrophobia in the MRI scanner. Additionally, DTI data sets of 10 subjects were deemed unusable due to breathing or motion artifact. Therefore, 31 Marines were included in this analysis (mean ± SD age, 27.3 ± 6.9 years; height, 1.8 ± 0.1 m; weight, 80.6 ± 8.7 kg). Marines excluded from the study had no differences in anthropometric measures compared with those included. Of these volunteers, 10 Marines self-reported experiencing LBP at the time of the scan.

Coefficients of variation were relatively low for dependent and independent variables (range, 0.04-10.61; median, 0.16) (**APPENDIX**, available at www.jospt.org). On average, the greatest variation was found for the IVD health measures.

Regression Model

After initial grouping of independent variables, collinearity resulted in the removal of 8 of the 29 independent variables from the model (**FIGURE 2**). Collinear variables that were removed included diffusion measurements from either the multifidus or erector spinae, erector spinae FF, and BMI. Surprisingly, 9 of 18 dependent variables were found from the stepwise multiple linear regressions to have a significant predictor. In fact, FA of the erector spinae was a significant predictor of lumbar posture for 7 of the 18 dependent variables measured, and explained 20% to 35% of the variance for each outcome (**TABLE**). In general, increased FA in the erector spinae was predictive of increased lumbar lordosis, lumbosacral extension, and pelvic tilt in each position. Additionally, decreased T2 relaxation of the L4-L5 IVD was a significant predictor of increased lumbosacral extension when standing unloaded (P = .025, R^2 = 0.192). When prone on elbows, increasing subject weight was a significant predictor of increased lumbar lordosis (P = .016, R^2 = 0.219). No muscle volume, muscle microstructure, IVD health, or anthropometric measures were significant predictors of posture when subjects were sitting loaded.

TABLE Results From Stepwise Multiple Linear Regression

Dependent Variable	Significant Independent Variable	β*	R^2	P Value
Cobb angle				
Standing unloaded	None			
Standing loaded	ES FA	0.453	0.205	.02
Sitting loaded	None			
Prone on elbows loaded	Weight	0.468	0.219	.016
Delta load†	None			
Delta position‡	None			
Angle with respect to horizontal				
Standing unloaded	T2 L4-L5	−0.439	0.192	.025
Standing loaded	ES FA	0.514	0.264	.007
Sitting loaded	None			
Prone on elbows loaded	ES FA	−0.480	0.23	.013
Delta load†	None			
Delta position‡	ES FA	0.455	0.207	.02
Sacral angle				
Standing unloaded	ES FA	0.442	0.195	.024
Standing loaded	ES FA	0.587	0.345	.002
Sitting loaded	None			
Prone on elbows loaded	ES FA	0.562	0.316	.003
Delta load†	None			
Delta position‡	None			

Abbreviations: ES, erector spinae; FA, fractional anisotropy.
**Standardized coefficient.*
†Standing unloaded to standing loaded.
‡Sitting loaded to prone on elbows loaded.

DISCUSSION

In this study, we evaluated the relationship between lumbar spine posture and muscle structure, IVD health, and anthropometric measures in 31 active-duty male Marines in simulated, relevant, operational positions and loading conditions. Fractional anisotropy of the erector spinae was a significant predictor in 7 of the 18 measures of lumbar spine posture across several different positions. For the standing loaded condition, FA of the erector spinae was a significant predictor of all 3 measures of lumbar posture; Marines with increased FA of the erector spinae had a more lordotic, extended lumbar posture with greater sacral tilt. Muscle volume was not a significant predictor of any postural measurements, despite being a commonly used proxy for muscle strength.[10,19] Together, the ability of FA to predict postural behavior in several positions and the absence of association between muscle volume and lumbar spine posture suggest that muscle microstructure, but not quantity—both measures associated with force-generating capacity of muscle—is an important predictor of lumbar spine posture.

Diffusion tensor imaging is an MRI technique that measures the restricted diffusion of water in tissues with anisotropic microstructure.[1] As the sarcolemma is considered to be the primary barrier to diffusion, DTI is believed to be most sensitive to changes in fiber size, because radial diffusion of water across a muscle fiber is more restricted (by the sarcolemma) than longitudinal diffusion within a muscle fiber.[44,45] While it has been shown that FA and fiber area are inversely related,[2,5,8,12,35] it is important to note that the exact relationship has not been validated. However, it is well established that muscle fiber area and isometric force are directly related.[17,21,22] Therefore, it appears that there is likely an inverse relationship between FA and isometric force-generating capacity of muscle. As such, it is inferred that when FA increases, the force-generating capacity of a muscle decreases (ie, the muscle is weaker). For example, if the multifidus muscles in 2 Marines were imaged using DTI and 1 had a larger FA (smaller fiber size), that muscle would be expected to generate less overall force.

Two unique relationships between posture and muscle structure were found in this study: (1) the erector spinae, not the multifidus, and (2) muscle microstructure, not volume, were found to be significant predictors of lumbar posture. First, FA of the multifidus and FA of the erector spinae were found to be collinear, with FA of the erector spinae being a stronger descriptor of the eigenvector from the PCA. Therefore, the multifidus was not included in the final statistical model. To verify that FA of the multifidus was not removed from the model because it had less variability than FA of the erector spinae, a coefficient of variation was calculated for both variables. Fractional anisotropy of the erector spinae had less variability relative to the mean than did FA of the multifidus (0.07 versus 0.08), further supporting the latter as a stronger descriptor of the eigenvector. While there is a small difference in variability of these measures, the variability values are both greater than the associated measurement error (0.03 and 0.04, respectively). This finding suggests that while the multifidus stabilizes the individual segments of the spinal column,[46,47] the erector spinae may play a role in determining gross lumbar posture.

Second, while muscle volume is proportional to muscle strength,[17,27] muscle microstructure has been shown to be a more accurate predictor of muscle force-generating capacity. Clinically, the findings from this study are important because they suggest that microstructural quality of the lumbar muscles is more important to whole lumbar posture in functionally loaded positions than the quantity or volume of muscle. This is not surprising given that measures of whole muscle size and volume are confounded by noncontractile tissue, such as fat and fibrosis. Importantly, FA may be a noninvasive composite measure of the functional contractile tissue present in a whole muscle, which seems to explain much of the variance in postural responses to body position.

In this study, T2 of the L4-L5 IVD was found to be inversely proportional to lumbosacral extension when Marines were standing without load. This suggests that Marines with decreased IVD T2 values (increased IVD degeneration) at L4-L5 have increased lumbosacral extension. Previously, using the Pfirrmann grading scale, the authors[4] reported no significant difference in lumbosacral extension in Marines when categorized by degeneration at L5-S1 (Pfirrmann grade greater than 2). As L5-S1 is the base of support of the lumbar spine, it was assumed that degeneration at this level would have whole lumbar postural consequences. However, our findings demonstrate that health of the L4-L5 IVD is related to whole lumbar posture and, therefore, should be considered an important structural level for whole lumbar stability. The finding that single-level disc health has the potential to influence lumbosacral flexion highlights the importance of the lower lumbar spine as a transition zone of load between the trunk and body. Changes to the health of this region have the potential to affect support of the torso.

Several studies have previously attempted to determine the relationship between lumbar lordosis and BMI. It appears that increased lumbar lordosis might be found in individuals with increased BMI[11,23]; however, other studies have shown no difference.[49] In this study, weight and BMI were found to be collinear, with weight being the stronger predictor of the eigenvector from PCA; therefore, BMI was dropped from the final statistical model. However, this is likely due to a larger variance in subject weight rather than in BMI in this relatively homogeneous population. If a more representative cross-section of the population were used, then these findings may have been different.

In this study, the researchers made several attempts to decrease the complexity of the model to decrease the amount of type I error that can be associated with making multiple comparisons. First, this study does not include individual vertebral-level measures of muscle structure or lumbar posture. Second, the authors removed collinear variables with clustering and PCA to minimize the number of independent variables representing similar constructs that were entered into the model. Third, this study evaluated forward, backward, and stepwise multiple linear regression models to determine which model was the most conservative approach. Results were the same with forward and stepwise elimination techniques, and backward elimination allowed for several more independent variables to be retained in the model, suggesting that it was the least conservative regression approach. Therefore, the authors chose to use a stepwise multiple linear regression technique, as it appeared to be the most conservative model.

The Marines in this study were not recruited based on history or presence of LBP at the time of the study, and approximately one third of the Marines who were included in this study reported LBP. It is important to note that no Marines had an episode of LBP so severe that they were relieved of duty. In a previous study, no difference in lumbar spine posture was found between Marines with and without

LBP at the time of data collection.[3] No differences have been observed between Marines with and without LBP at the time of data collection for muscle physiology, IVD health, or anthropometric measures (data not published). As LBP did not result in differences in the dependent or independent variables measured, it is unlikely that the inclusion of Marines with and without LBP affected the findings of this study.

There are several limitations to this study. First, the Marines had relatively normal muscle, with no underlying pathology observable. In patients with pathology or age-related atrophic changes in muscle, the volume or FF of muscle may be more important in predicting lumbar posture. Therefore, the results of this study may only extend to a highly active population. Second, the positions measured in this study place relatively small challenges on the muscles of the lumbar spine. A future direction of this research is to investigate whether muscle microstructure can predict posture, given the heavy loading conditions under which Marines routinely operate.

Finally, the model used in the present study incorporated 21 variables, with only 31 full data sets to include. This was a retrospective analysis of 2 studies investigating (1) the effect of operationally relevant positions on lumbar posture[3] and (2) normative paraspinal muscle composition in active-duty Marines. It was determined that 43 participants were needed to provide adequate power to these studies. However, to mitigate type I error associated with multiple comparisons, the authors used the most conservative statistical approach. While more participants may provide an increase in the amount of variance explained by the model, this study still reached significance with 31 complete data sets.

CONCLUSION

The authors believe that this study is the first to measure the predictive capacity of lumbar muscle structure, IVD health, and anthropometric measures on lumbar spine posture in different positions. It is surprising that any structural variable in muscle predicted any of the variance in posture, because many clinicians believe that short-term postural positions are more related to motor control than to strength or end organ–dependent behavior.

This study found that FA of the erector spinae was a significant predictor of several lumbar postural measures. In general, decreased FA of the erector spinae resulted in decreased lordosis, lumbosacral extension, and anterior pelvic tilt. This posture results in decreased shear stress at lower lumbar levels during hyperlordosis and may be considered a more protective posture for preventing injury and LBP when loading the lumbar spine.[37] Decreased FA of the erector spinae can be physiologically interpreted as larger muscle fibers with more capacity to generate force. Due to the intense training and demands of their jobs, the Marines in this study were extremely active and trained on how to adapt their posture in different positions, while wearing body armor, to minimize their risk of injury. Therefore, these findings may not translate to a civilian population.

The findings of this study support the idea that muscle strengthening/exercise may influence posture, although this cause-and-effect relationship needs to be substantiated in prospective clinical research. As this relationship was found in a healthy population with relatively little variance in muscle quality, it is likely that these relationships may be stronger in patients with LBP or injury. Understanding the influence of microstructural features of muscle on posture may allow clinicians to prognostically categorize patients into groups that may respond better to exercise-based treatments. Future studies should take a more controlled approach to determine whether targeted exercise of the erector spinae muscles increases muscle quality (measured with DTI) and can elicit a postural response.

KEY POINTS

FINDINGS: Fractional anisotropy of the erector spinae was a significant predictor of lumbar lordosis, lumbar flexion, and sacral tilt in several different operationally relevant positions in active-duty Marines.

IMPLICATIONS: The finding that fractional anisotropy can predict postural responses in several positions, along with the absence of association between muscle volume and lumbar spine posture, suggests that muscle microstructure, but not quantity, is an important predictor of lumbar spine posture.

CAUTION: These findings were found in a group of highly active Marines and may not translate to a civilian population.

ACKNOWLEDGMENTS: *The authors thank the Marines from the 1st and 5th Regiments who supported this effort.*

REFERENCES

1. Basser PJ, Mattiello J, LeBihan D. MR diffusion tensor spectroscopy and imaging. *Biophys J*. 1994;66:259-267. https://doi.org/10.1016/S0006-3495(94)80775-1
2. Berry DB, Regner B, Galinsky V, Ward SR, Frank LR. Relationships between tissue microstructure and the diffusion tensor in simulated skeletal muscle. *Magn Reson Med*. 2018;80:317-329. https://doi.org/10.1002/mrm.26993
3. Berry DB, Rodríguez-Soto AE, Su J, et al. Lumbar spine postures in Marines during simulated operational positions. *J Orthop Res*. 2017;35:2145-2153. https://doi.org/10.1002/jor.23510
4. Berry DB, Rodríguez-Soto AE, Tokunaga JR, Gombatto SP, Ward SR. An endplate-based joint coordinate system for measuring kinematics in normal and abnormally-shaped lumbar vertebrae. *J Appl Biomech*. 2015;31:499-503. https://doi.org/10.1123/jab.2015-0008
5. Berry DB, You S, Warner J, Frank LR, Chen S, Ward SR. A 3D tissue-printing approach for validation of diffusion tensor imaging in skeletal muscle. *Tissue Eng Part A*. 2017;23:980-988. https://doi.org/10.1089/ten.tea.2016.0438
6. Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res*. 1996;29:162-173. https://doi.org/10.1006/cbmr.1996.0014
7. Demoulin C, Crielaard JM, Vanderthommen M. Spinal muscle evaluation in healthy individuals and low-back-pain patients: a literature review.

Joint Bone Spine. 2007;74:9-13. https://doi.org/10.1016/j.jbspin.2006.02.013
8. Deux JF, Malzy P, Paragios N, et al. Assessment of calf muscle contraction by diffusion tensor imaging. *Eur Radiol*. 2008;18:2303-2310. https://doi.org/10.1007/s00330-008-1012-z
9. Fortin M, Macedo LG. Multifidus and paraspinal muscle group cross-sectional areas of patients with low back pain and control patients: a systematic review with a focus on blinding. *Phys Ther*. 2013;93:873-888. https://doi.org/10.2522/ptj.20120457
10. Fukunaga T, Miyatani M, Tachi M, Kouzaki M, Kawakami Y, Kanehisa H. Muscle volume is a major determinant of joint torque in humans. *Acta Physiol Scand*. 2001;172:249-255. https://doi.org/10.1046/j.1365-201x.2001.00867.x
11. Guo JM, Zhang GQ, Alimujiang. [Effect of BMI and WHR on lumbar lordosis and sacrum slant angle in middle and elderly women]. *Zhongguo Gu Shang*. 2008;21:30-31.
12. Hatakenaka M, Yabuuchi H, Sunami S, et al. Joint position affects muscle proton diffusion: evaluation with a 3-T MR system. *AJR Am J Roentgenol*. 2010;194:W208-W211. https://doi.org/10.2214/AJR.09.3167
13. Keorochana G, Taghavi CE, Lee KB, et al. Effect of sagittal alignment on kinematic changes and degree of disc degeneration in the lumbar spine: an analysis using positional MRI. *Spine (Phila Pa 1976)*. 2011;36:893-898. https://doi.org/10.1097/BRS.0b013e3181f4d212
14. Kim YB, Kim YJ, Ahn YJ, et al. A comparative analysis of sagittal spinopelvic alignment between young and old men without localized disc degeneration. *Eur Spine J*. 2014;23:1400-1406. https://doi.org/10.1007/s00586-014-3236-8
15. Knapik JJ, Reynolds KL, Harman E. Soldier load carriage: historical, physiological, biomechanical, and medical aspects. *Mil Med*. 2004;169:45-56. https://doi.org/10.7205/MILMED.169.1.45
16. Lao L, Daubs MD, Scott TP, et al. Effect of disc degeneration on lumbar segmental mobility analyzed by kinetic magnetic resonance imaging. *Spine (Phila Pa 1976)*. 2015;40:316-322. https://doi.org/10.1097/BRS.0000000000000738
17. Lieber RL. *Skeletal Muscle Structure, Function, and Plasticity: The Physiological Basis of Rehabilitation*. 2nd ed. Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins; 2002.
18. Mac-Thiong JM, Roussouly P, Berthonnaud É, Guigui P. Age- and sex-related variations in sagittal sacropelvic morphology and balance in asymptomatic adults. *Eur Spine J*. 2011;20 suppl 5:572-577. https://doi.org/10.1007/s00586-011-1923-2
19. Meakin JR, Fulford J, Seymour R, Welsman JR, Knapp KM. The relationship between sagittal curvature and extensor muscle volume in the lumbar spine. *J Anat*. 2013;222:608-614. https://doi.org/10.1111/joa.12047
20. Meyer GA, Lieber RL. Elucidation of extracellular matrix mechanics from muscle fibers and fiber bundles. *J Biomech*. 2011;44:771-773. https://doi.org/10.1016/j.jbiomech.2010.10.044
21. Miller AE, MacDougall JD, Tarnopolsky MA, Sale DG. Gender differences in strength and muscle fiber characteristics. *Eur J Appl Physiol Occup Physiol*. 1993;66:254-262. https://doi.org/10.1007/BF00235103
22. Minamoto VB, Hulst JB, Lim M, et al. Increased efficacy and decreased systemic-effects of botulinum toxin A injection after active or passive muscle manipulation. *Dev Med Child Neurol*. 2007;49:907-914. https://doi.org/10.1111/j.1469-8749.2007.00907.x
23. Murrie VL, Dixon AK, Hollingworth W, Wilson H, Doyle TA. Lumbar lordosis: study of patients with and without low back pain. *Clin Anat*. 2003;16:144-147. https://doi.org/10.1002/ca.10114
24. Oudeman J, Nederveen AJ, Strijkers GJ, Maas M, Luijten PR, Froeling M. Techniques and applications of skeletal muscle diffusion tensor imaging: a review. *J Magn Reson Imaging*. 2016;43:773-788. https://doi.org/10.1002/jmri.25016
25. Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine (Phila Pa 1976)*. 2001;26:1873-1878. https://doi.org/10.1097/00007632-200109010-00011
26. Pontén E, Fridén J, Thornell LE, Lieber RL. Spastic wrist flexors are more severely affected than wrist extensors in children with cerebral palsy. *Dev Med Child Neurol*. 2005;47:384-389. https://doi.org/10.1111/j.1469-8749.2005.tb01158.x
27. Powell PL, Roy RR, Kanim P, Bello MA, Edgerton VR. Predictability of skeletal muscle tension from architectural determinations in guinea pig hindlimbs. *J Appl Physiol Respir Environ Exerc Physiol*. 1984;57:1715-1721. https://doi.org/10.1152/jappl.1984.57.6.1715
28. Reynolds KL, Kaszuba J, Mello RP, Patton JF. Prolonged Treadmill Load Carriage: Acute Injuries and Changes in Foot Anthropometry. Natick, MA: US Army Research Institute of Environmental Medicine; 1990.
29. Rodriguez-Soto AE, Berry DB, Jaworski R, et al. The effect of training on lumbar spine posture and intervertebral disc degeneration in active-duty Marines. *Ergonomics*. 2017;60:1055-1063. https://doi.org/10.1080/00140139.2016.1252858
30. Rodríguez-Soto AE, Berry DB, Palombo L, Valaik E, Kelly KR, Ward SR. Effect of load magnitude and distribution on lumbar spine posture in active-duty Marines. *Spine (Phila Pa 1976)*. 2017;42:345-351. https://doi.org/10.1097/BRS.0000000000001742
31. Rodríguez-Soto AE, Jaworski R, Jensen A, et al. Effect of load carriage on lumbar spine kinematics. *Spine (Phila Pa 1976)*. 2013;38:E783-E791. https://doi.org/10.1097/BRS.0b013e3182913e9f
32. Rosset A, Spadola L, Ratib O. OsiriX: an open-source software for navigating in multidimensional DICOM images. *J Digit Imaging*. 2004;17:205-216. https://doi.org/10.1007/s10278-004-1014-6
33. Roy TC, Knapik JJ, Ritland BM, Murphy N, Sharp MA. Risk factors for musculoskeletal injuries for soldiers deployed to Afghanistan. *Aviat Space Environ Med*. 2012;83:1060-1066. https://doi.org/10.3357/ASEM.3341.2012
34. Roy TC, Lopez HP, Piva SR. Loads worn by soldiers predict episodes of low back pain during deployment to Afghanistan. *Spine (Phila Pa 1976)*. 2013;38:1310-1317. https://doi.org/10.1097/BRS.0b013e31829265c4
35. Schwenzer NF, Steidle G, Martirosian P, et al. Diffusion tensor imaging of the human calf muscle: distinct changes in fractional anisotropy and mean diffusion due to passive muscle shortening and stretching. *NMR Biomed*. 2009;22:1047-1053. https://doi.org/10.1002/nbm.1409
36. Shahidi B, Parra CL, Berry DB, et al. Contribution of lumbar spine pathology and age to paraspinal muscle size and fatty infiltration. *Spine (Phila Pa 1976)*. 2017;42:616-623. https://doi.org/10.1097/BRS.0000000000001848
37. Shirazi-Adl A, El-Rich M, Pop DG, Parnianpour M. Spinal muscle forces, internal loads and stability in standing under various postures and loads—application of kinematics-based algorithm. *Eur Spine J*. 2005;14:381-392. https://doi.org/10.1007/s00586-004-0779-0
38. Sinaki M, Itoi E, Rogers JW, Bergstralh EJ, Wahner HW. Correlation of back extensor strength with thoracic kyphosis and lumbar lordosis in estrogen-deficient women. *Am J Phys Med Rehabil*. 1996;75:370-374. https://doi.org/10.1097/00002060-199609000-00013
39. Smith LR, Lee KS, Ward SR, Chambers HG, Lieber RL. Hamstring contractures in children with spastic cerebral palsy result from a stiffer extracellular matrix and increased *in vivo* sarcomere length. *J Physiol*. 2011;589:2625-2639. https://doi.org/10.1113/jphysiol.2010.203364
40. Sparrey CJ, Bailey JF, Safaee M, et al. Etiology of lumbar lordosis and its pathophysiology: a review of the evolution of lumbar lordosis, and the mechanics and biology of lumbar degeneration. *Neurosurg Focus*. 2014;36:E1. https://doi.org/10.3171/2014.1.FOCUS13551
41. Stelzeneder D, Welsch GH, Kovács BK, et al. Quantitative T2 evaluation at 3.0 T compared to morphological grading of the lumbar intervertebral disc: a standardized evaluation approach in patients with low back pain. *Eur J Radiol*. 2012;81:324-330. https://doi.org/10.1016/j.ejrad.2010.12.093
42. Sullivan MS, Dickinson CE, Troup JD. The influence of age and gender on lumbar spine sagittal plane range of motion. A study of 1126 healthy subjects. *Spine (Phila Pa 1976)*. 1994;19:682-686. https://doi.org/10.1097/00007632-199403001-00007
43. Takatalo J, Karppinen J, Taimela S, et al. Body mass index is associated with lumbar disc

degeneration in young Finnish males: subsample of Northern Finland birth cohort study 1986. *BMC Musculoskelet Disord*. 2013;14:87. https://doi.org/10.1186/1471-2474-14-87

44. Van Donkelaar CC, Kretzers LJ, Bovendeerd PH, et al. Diffusion tensor imaging in biomechanical studies of skeletal muscle function. *J Anat*. 1999;194 pt 1:79-88. https://doi.org/10.1046/j.1469-7580.1999.19410079.x

45. van Doorn A, Bovendeerd PH, Nicolay K, Drost MR, Janssen JD. Determination of muscle fibre orientation using Diffusion-Weighted MRI. *Eur J Morphol*. 1996;34:5-10.

46. Ward SR, Kim CW, Eng CM, et al. Architectural analysis and intraoperative measurements demonstrate the unique design of the multifidus muscle for lumbar spine stability. *J Bone Joint Surg Am*. 2009;91:176-185. https://doi.org/10.2106/JBJS.G.01311

47. Ward SR, Tomiya A, Regev GJ, et al. Passive mechanical properties of the lumbar multifidus muscle support its role as a stabilizer. *J Biomech*. 2009;42:1384-1389. https://doi.org/10.1016/j.jbiomech.2008.09.042

48. Wokke BH, van den Bergen JC, Versluis MJ, et al. Quantitative MRI and strength measurements in the assessment of muscle quality in Duchenne muscular dystrophy. *Neuromuscul Disord*. 2014;24:409-416. https://doi.org/10.1016/j.nmd.2014.01.015

49. Youdas JW, Hollman JH, Krause DA. The effects of gender, age, and body mass index on standing lumbar curvature in persons without current low back pain. *Physiother Theory Pract*. 2006;22:229-237. https://doi.org/10.1080/09593980600927864

APPENDIX

Coefficient of Variation Calculated for Each Dependent Variable

Position	Angle With Respect to Horizontal	Sacral Angle	Cobb Angle
Standing unloaded	0.05	0.17	0.18
Standing loaded	0.05	0.22	0.22
Sitting loaded	0.04	0.48	0.16
Prone on elbows loaded	0.06	0.26	0.16
Delta load	10.61	1.86	4.11
Delta position	1.18	1.30	0.37

Coefficient of Variation Calculated for Each Independent Variable

Independent Variable	Coefficient of Variation
Muscle measures	
Multifidus	
Volume	0.14
Fat fraction	0.41
Mean diffusivity	0.05
Fractional anisotropy	0.08
Lambda 1	0.04
Lambda 2	0.04
Lambda 3	0.06
Erector spinae	
Volume	0.22
Fat fraction	0.41
Mean diffusivity	0.05
Fractional anisotropy	0.07
Lambda 1	0.04
Lambda 2	0.04
Lambda 3	0.05
Psoas volume	0.13
Quadratus lumborum volume	0.19
IVD measures	
T2	
L1-L2	0.24
L2-L3	0.27
L3-L4	0.29
L4-L5	0.35
L5-S1	0.41
Anthropometric measures	
Age	0.24
Height	0.04
Weight	0.12
Body mass index	0.11

Abbreviation: IVD, intervertebral disc.

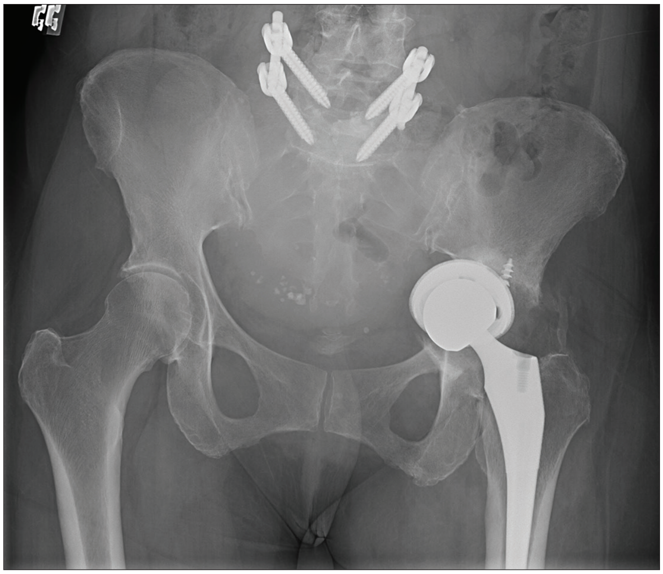


FIGURE 2. Anteroposterior radiograph of the pelvis taken 8 months post total hip arthroplasty, showing medial migration of the acetabular component, indicating loosening. Additionally, orthopaedic hardware is seen in the lumbar spine from spinal fusion surgery performed 10 weeks prior.

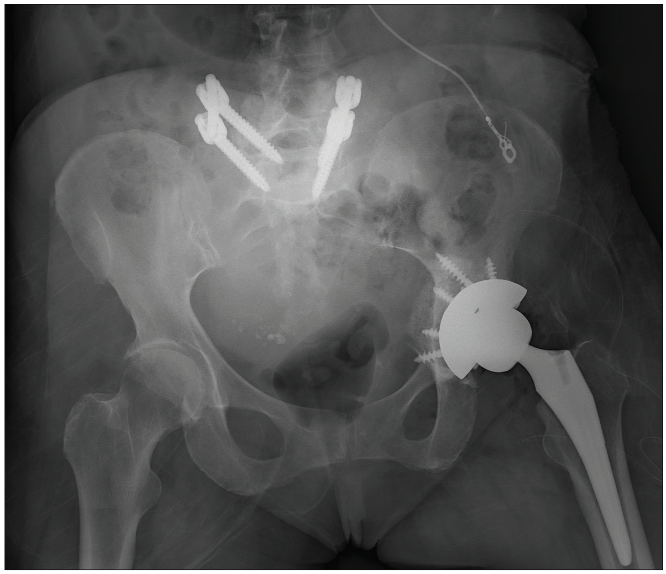


FIGURE 3. Anteroposterior radiograph of the pelvis taken post total hip arthroplasty revision, showing near anatomic alignment of the revision total hip arthroplasty.

Delayed Infection in a Patient After Total Hip Arthroplasty

MANASI BOHRA, PT, MS, OCS, *Department of Therapy Services, University of Chicago Medicine, Chicago, IL.*
DEIDRE CWIAN, PT, DPT, *Department of Therapy Services, University of Chicago Medicine, Chicago, IL.*
COLLEEN PEYTON, PT, DPT, PCS, *Department of Therapy Services, University of Chicago Medicine, Chicago, IL.*

A 67-YEAR-OLD WOMAN PRESENTED 3 weeks following left total hip arthroplasty (THA) with a direct anterior approach (**FIGURE 1**, available at www.jospt.org). She received 12 physical therapy sessions over 2 months, and then returned to work as a nurse. One week after returning to work, the patient experienced severe left buttock pain and was diagnosed with degenerative L5-S1 spondylolisthesis and foraminal stenosis. Nonsurgical treatment over a 2-month period failed, and she subsequently underwent L5-S1 laminectomy and fusion 5 months after the THA.

The patient returned to physical therapy 2 weeks after her laminectomy, ambulating with a cane. She slowly progressed in gait stability and strength over an 8-week period. Then, over a 2-week period, her gait deteriorated, left buttock pain worsened, and left hip extensor and flexor strength decreased from 4/5 to 2−/5. She was afebrile, with no warmth to palpation. The patient was referred back to the physician, who ordered radiographs, which showed medial migration of the hip components (**FIGURE 2**).

The patient had a revision THA the day after imaging, which consisted of bone grafting to the acetabular defect and replacement of the acetabular component (**FIGURE 3**). Joint cultures were positive for the bacteria *Parvimonas micra*. The patient was treated with intravenous vancomycin for 2 weeks. She then completed 5 months of rehabilitation and returned to work.

Infection following a THA occurs in less than 1% of patients.[1] Infections are classified as early (less than 3 months), delayed (3-24 months), or late (greater than 24 months).[2] Complications of infection include prosthetic-component loosening and failure[2] and should be considered when establishing differential diagnoses in patients presenting with joint pain who have had a THA. *J Orthop Sports Phys Ther 2018;48(8):666. doi:10.2519/jospt.2018.7727*

References
1. American Academy of Orthopaedic Surgeons. Joint replacement infection. Available at: http://orthoinfo.aaos.org/topic.cfm?topic=A00629. Accessed March 15, 2017.
2. Zimmerli W, Trampuz A, Ochsner PE. Prosthetic-joint infections. *N Engl J Med*. 2004;351:1645-1654. https://doi.org/10.1056/NEJMra040181

DANIEL DEUTSCHER, PT, PhD[1] • **MARK W. WERNEKE**, PT, MS, Dip MDT[2] • **DEANNA HAYES**, PT, DPT, MS[2] • **JEROME E. MIODUSKI**, MS[2]
KARON F. COOK, PhD[3] • **JULIE M. FRITZ**, PT, PhD[4,5] • **LINDA J. WOODHOUSE**, PT, PhD[6] • **PAUL W. STRATFORD**, PT, MS[7]

Impact of Risk Adjustment on Provider Ranking for Patients With Low Back Pain Receiving Physical Therapy

Quality of care, as measured by the outcomes that are most relevant to the patient, has become a central focus of the effort to improve today's health care system.[28] The 2001 publication *Crossing the Quality Chasm: A New Health System for the 21st Century*, by the Institute of Medicine,[22] shifted the focus of health system improvement from volume of services to an emphasis on the quality of care. Patient-reported outcome measures are a key component for understanding the quality of care.[29] Risk-adjusted patient-reported outcome measures and clinic ranking systems have been recommended by the Centers for Medicare and Medicaid Services[4] to meaningfully assess the quality of physical therapy services and help consumers choose the best provider for their medical needs. Thus, statistical methods for developing clinic ranking have become an important area for research.[15,33,34]

BACKGROUND: The impact of risk adjustment on clinic quality ranking for patients treated in physical therapy outpatient clinics is unknown.

OBJECTIVES: To compare clinic ranking, based on unadjusted versus risk-adjusted outcomes for patients with low back pain (LBP) who are treated in physical therapy outpatient clinics.

METHODS: This retrospective cohort study involved a secondary analysis of data from adult patients with LBP treated in outpatient physical therapy clinics from 2014 to 2016. Patients with complete outcomes data at admission and discharge were included to develop the risk-adjustment model. Clinics with complete outcomes data for at least 50% of patients and at least 10 complete episodes of care per clinician per year were included for ranking assessment. The R^2 shrinkage and predictive ratio were used to assess overfitting. Agreement between unadjusted and adjusted rankings was assessed with percentile ranking by deciles or 3 distinct quality ranks based on uncertainty assessment.

RESULTS: The primary sample included 414 125 patients (mean ± SD age, 57 ± 17 years; 60% women) treated by 12 569 clinicians from 3048 clinics from all US states; 82% of patients from 2107 clinics were included in the ranking assessment. The R^2 shrinkage was less than 1%, with a predictive ratio of 1. Risk adjustment impacted ranking for 70% or 31% of clinics, based on deciles or 3 distinct quality levels, respectively.

CONCLUSION: Important changes in ranking were found after adjusting for basic patient characteristics of those admitted to physical therapy for treatment of LBP. Risk-adjustment profiling is necessary to more accurately reflect quality of care when treating patients with LBP.

LEVEL OF EVIDENCE: Therapy, level 2b. *J Orthop Sports Phys Ther 2018;48(8):637-648. Epub 22 May 2018. doi:10.2519/jospt.2018.7981*

KEY WORDS: *functional status, patient-reported outcome measures, physical therapy, provider ranking, risk adjustment*

Comprehensive and robust methods of risk adjustment are essential to achieve objective comparisons of patient-reported outcome measures,[45] including functional status, for the purpose of clinic ranking.[20] Clinic ranking allows for benchmark reporting[15] and is integral to quality payment initiatives, also known as *value-based purchasing*[32,41] or *pay for performance*.[16] The use of patient-reported outcomes data for provider ranking has also been promoted by the National Quality Forum to help reduce variation in health care quality.[27]

Risk adjustment is required when examining observational data to rank or compare outcomes across patients and providers. The risk-adjusted data allow for fair comparison by taking into account lower outcomes due to patients' prognoses or medical complexities that are beyond the influence of the provider. Thus, risk adjustment aims to mitigate threats to internal validity by controlling for potential confounding of results that may be attributed to differences in case mix characteristics. Therefore, risk adjustment provides

[1]Department of Physical Therapy and the Maccabi Institute for Research & Innovation, Maccabi Healthcare Services, Tel-Aviv, Israel. [2]Focus On Therapeutic Outcomes Inc, Knoxville, TN. [3]Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL. [4]Intermountain Healthcare, Salt Lake City, UT. [5]Department of Physical Therapy and Athletic Training, University of Utah, Salt Lake City, UT. [6]Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, Canada. [7]School of Rehabilitation Science, McMaster University, Hamilton, Canada. The University of Utah Institutional Review Board approved this study. Drs Cook and Deutscher acknowledge that they are consultants for Focus On Therapeutic Outcomes Inc, the database management company that manages the data analyzed in the manuscript. Mr Werneke, Dr Hayes, and Mr Mioduski are employees of Focus On Therapeutic Outcomes Inc. Drs Julie Fritz and Linda Woodhouse serve on the Research Advisory Board for Focus on Therapeutic Outcomes Inc. The authors certify that they have no other 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 Daniel Deutscher, 27 Hamered Street, Tel Aviv 68125 Israel. E-mail: deutsch_d@mac.org.il

a more accurate reflection of a clinic's or clinician's quality of care. The selection of factors for which adjustments need to be made remains under debate.[43] However, it is generally accepted that quality should be determined based on outcomes that reflect the complexity of each provider's patient mix.[20]

Risk adjustment has been used in physical therapy for many years.[9,15,21,33-35] For example, in 2006, Hart and Connolly[16] described the need for risk adjustment when developing a pay-for-performance model in outpatient physical and occupational therapy. More recently, Resnik et al[35] and Gozalo et al[15] described risk-adjustment methods to benchmark the performance of physical therapy clinics, using observational patient-reported outcome measures of functional status. Because functional status is a major goal of rehabilitation treatment, it is commonly targeted in performance outcome measurement.[2,3,9,15,17,33,36] Other target outcomes include health care costs,[12] return to work,[13,14] and value of care.[24,30] Patient demographic and health characteristics are known to impact intended outcomes and, thus, are frequently adjusted for in rehabilitation studies. Demographic factors often include age, sex, race, payer type, and other sociodemographic indices. Frequently, adjustments for health factors, other than the condition being treated, include acuity of the condition (days from onset), comorbidities, and chronic medication use.[9,15,16,35]

Benchmarking without risk adjustment raises a number of notable concerns. Use of unadjusted patient-reported outcome measure data can lead to misclassification of provider performance, misalignment of payer reimbursement for provider services, and obscured relationships between nonmodifiable patient factors and outcomes of interest. Further, the use of such unadjusted data disincentivizes providers from treating the most complex patients. To our knowledge, no studies have compared the rankings of physical therapy outpatient clinics, with or without risk adjustment of patient-reported outcome data. Thus, the purpose of this study was to compare clinic ranking results, based on raw (unadjusted) versus risk-adjusted patient-reported outcomes of functional status from patients with low back pain (LBP), which accounts for some of the highest health care spending of all medical conditions[10] and accounts for the largest group of patients who seek outpatient physical therapy treatment.[9,16] Although identification of an optimal set of adjustment factors was not the study's primary aim, the authors examined the available patient factors known to be associated with functional status outcomes to establish an optimal risk-adjustment model for this data set.[9,15,16] The researchers hypothesized that clinic rankings would change substantially after risk adjustment.

METHODS

Design and Sample Selection

The authors conducted a secondary analysis of prospectively collected data from adult patients (aged 18 years or older) with LBP treated in outpatient physical therapy clinics in the United States from 2014 to 2016. Patients were identified as having LBP by their selection of the lumbar spine region on the functional status survey. Routinely, patients are instructed to select the body area most affected and to identify the main cause for seeking treatment. Because normal treatment was not altered, patient informed consent was not required.

All participating clinics routinely assessed patient-reported outcome measures of functional status using the Patient Inquiry software (Focus On Therapeutic Outcomes Inc, Knoxville, TN).[40] The majority of clinics (96%) that utilize Patient Inquiry for outcome measurement are private practice or hospital-based outpatient clinics.[5] Patients who completed the self-reported functional status assessment both at admission and discharge were included in the development of the risk-adjustment model.

To assess the potential for a systematic patient selection bias, the authors compared characteristics of patients with complete and incomplete outcomes data. To increase generalizability, this study used 2 selection criteria to assess the impact of risk adjustment on provider ranking. First, only clinics with a completion rate equal to or greater than 50% were included, as recommended previously.[8] Completion rate was defined as the percentage of patients whose self-reported functional status was assessed both at admission and again at discharge.[7] Second, to increase representativeness of patients included for clinic ranking, only clinics with at least 10 complete patient episodes of care per clinician per year were included in the ranking assessment, as previously described.[6,8]

Data Collection

Focus On Therapeutic Outcomes Inc collects a standardized set of data that includes patient-reported outcome measures, patient demographics, and health characteristics, providing a wide range of variables to examine for associations with functional status outcomes.[40] To decrease the possibility for a systematic bias in the collection of patient-reported outcomes, Focus On Therapeutic Outcomes Inc routinely implements educational modules instructing providers on how to administer patient-reported outcome measures in a neutral manner to their patients. The patient-reported outcome measure of functional status was measured at admission and at discharge from therapy using the lumbar computerized adaptive test (LCAT). There is substantial empirical evidence for the LCAT's responsiveness, construct validity, and clinical interpretability.[18,19,44] The data included the following patient factors that could be evaluated for inclusion in a model for risk adjustment: functional status at admission (continuous), age (continuous), sex (male/female), acuity as number of days from onset of the treated condition (6 categories), type of payer (10 categories), number of related surgeries (4 categories), exercise history (3 categories), use of medication at intake for the treatment of LBP (yes/no), previous treatment for LBP (yes/no), treatment post surgery (lumbar fusion, laminectomy, or other), and 31 comorbidities, excluding only the comorbidity of back pain, a condition expected to exist in the target population of this study (**TABLE 1**).

Assessment of Patient Selection Bias

To assess possible patient selection bias and the impact of selection criteria on the ranking results, the authors compared the characteristics of 3 sets of pa-

TABLE 1 HEALTH AND DEMOGRAPHIC PATIENT CHARACTERISTICS*

Patient Characteristic	Total	Incomplete Outcomes	Complete Outcomes Selected for Ranking	Complete Outcomes Not Selected for Ranking
n	618199	204074	341642	72483
Mean ± SD (minimum-maximum) FS score at admission	48.6 ± 12.8 (0-98)	48.2 ± 13.3 (0-98)	48.8 ± 12.5 (0-98)	48.9 ± 12.6 (0-98)
Mean ± SD (minimum-maximum) FS LCAT score at discharge	...	...	62.9 ± 16.3 (0-99)	62.0 ± 16.3 (0-98)
Mean ± SD (minimum-maximum) age, y	55.6 ± 16.9 (18-116)	52.7 ± 16.7 (18-116)	57.0 ± 16.8 (18-116)	57.5 ± 17.0 (18-116)
Sex: female	60.0	60.3	59.8	60.2
Acuity				
0-7 d	4.0	4.1	4.0	3.9
8-14 d	6.4	6.2	6.5	6.4
15-21 d	7.8	7.6	7.8	8.0
22-90 d	23.3	22.5	23.6	23.8
91 d to 6 mo	12.5	12.4	12.6	12.8
>6 mo	46.0	47.2	45.5	45.2
Payer				
Indemnity insurance	3.7	5.0	2.6	5.4
Medicaid	6.0	8.4	4.8	4.9
Medicare A	1.4	1.2	1.3	1.9
Medicare B, under age 65	4.2	4.8	3.8	4.6
Medicare B, age 65 or above	24.6	17.4	27.8	30.2
Patient	0.6	0.7	0.5	0.4
Workers' compensation	5.7	4.9	5.9	6.6
Other (litigation, Medicare C, school, no charge, early intervention, commercial insurance)	8.7	10.0	8.0	8.8
No fault, auto insurance	1.4	1.2	1.6	1.5
HMO, preferred provider	43.6	46.5	43.5	35.6
Surgical history				
No related surgery	81.8	83.5	80.9	81.0
1 related surgery	11.7	10.3	12.4	12.3
2 related surgeries	3.7	3.5	3.9	3.8
3 or more related surgeries	2.8	2.7	2.8	2.9
Exercise history				
At least 3 times per week	38.6	37.3	39.0	40.1
1-2 times per week	24.2	24.5	24.1	24.2
Seldom or never	37.2	38.2	36.9	35.8
Medication use at intake	56.0	55.2	55.1	54.6
Previous treatment	49.0	49.3	49.8	48.9
Lumbar surgery procedure				
Fusion	1.3	0.9	1.6	1.4
Laminectomy/foramenectomy/discectomy	1.4	1.0	1.6	1.6
Other surgical codes	0.1	0.1	0.1	0.1

Table continues on page 640.

TABLE 1 HEALTH AND DEMOGRAPHIC PATIENT CHARACTERISTICS* (CONTINUED)

Patient Characteristic	Total	Incomplete Outcomes	Complete Outcomes Selected for Ranking	Complete Outcomes Not Selected for Ranking
Mean ± SD comorbidities, n[†]	4.9 ± 3.3 (4, 5)	4.9 ± 3.3 (4, 5)	5.0 ± 3.2 (4, 4)	4.9 ± 3.3 (4, 5)
Specific comorbidities				
Allergy	26.7	25.6	27.1	26.7
Angina	1.6	1.6	1.6	1.6
Anxiety or panic disorders	15.4	20.2	15.7	15.4
Arthritis	48.3	43.4	48.3	48.3
Asthma	11.2	12.1	11.0	11.2
Back pain (neck pain, low back pain, degenerative disc disease)[‡]	79.8	79.5	80.2	79.8
Cancer	8.7	6.8	8.4	8.7
Chronic obstructive pulmonary disease	4.2	4.5	4.2	4.2
Congestive heart failure	5.4	4.8	5.5	5.4
Depression	17.9	22.0	17.9	17.9
Diabetes type I or II	13.9	12.8	13.9	13.9
Gastrointestinal	18.6	18.0	18.9	18.6
Headaches	22.1	26.6	22.2	22.1
Hearing	7.0	5.7	6.9	7.0
Hepatitis/HIV/AIDS	1.1	1.3	1.0	1.1
High blood pressure	37.9	33.4	38.1	37.9
Heart attack (myocardial infarction)	3.2	3.0	3.2	3.2
Incontinence	6.9	5.9	6.6	6.9
Kidney, bladder, prostate, or urination problems	11.5	10.2	11.3	11.5
Neurological disease	2.0	1.8	1.9	2.0
Obesity (BMI ≥30 kg/m^2)	39.1	41.0	40.0	39.1
Osteoporosis	10.9	8.7	10.4	10.9
Other disorders	4.8	5.4	5.2	4.8
Peripheral vascular disease (or claudication)	1.8	1.7	1.8	1.8
Previous accidents (motor vehicle, work, or other accident)	13.5	13.8	13.3	13.5
Previous surgery	37.6	34.4	37.6	37.6
Prosthesis/implants	7.2	5.9	7.3	7.2
Sleep dysfunction	19.7	21.6	19.9	19.7
Stroke or transient ischemic attack	3.5	3.1	3.3	3.5
Visual impairment	11.7	9.0	11.2	11.7
Pacemaker	0.8	0.7	0.8	0.8
Seizures	0.6	0.9	0.7	0.6

Abbreviations: AIDS, acquired immune deficiency syndrome; BMI, body mass index; FS, functional status; HIV, human immunodeficiency virus; HMO, health maintenance organization; LCAT, lumbar computerized adaptive test.
**Patient characteristics at admission to physical therapy for the sample used to develop the risk-adjusted model (total), the sample used for the ranking analyses (selected), and the sample excluded from the ranking analyses (not selected). Values are percent unless otherwise indicated.*
[†]Values in parentheses are median, interquartile range, reported for number of comorbidities due to the skewed distribution.
[‡]Back pain was not allowed to enter the risk-adjusted model.

tients: those with incomplete outcomes data, and those with complete outcomes data who were either selected or not selected for ranking (**TABLE 1**). Additionally, the researchers assessed the impact of adjusting for patient censoring, using inverse probability weighting on the results. In this method, complete cases are weighted by the inverse of their probability of being a complete case.[38] Hence, patients less likely to have complete functional status data were given more weight in the risk-adjusted model than those who were likely to have complete data.[37] The authors compared predictions created by the unweighted and weighted models.

Risk-Adjustment Modeling

Risk-adjustment models were constructed and assessed for predictive validity in 3 steps. First, the researchers used a backward, stepwise, linear ordinary-least-square regression to identify patient factors that significantly contributed to the prediction of functional status outcomes at discharge. The backward stepwise procedure allows variables to be removed and entered in a sequential manner to create the most parsimonious final model. To adjust for the large data set available and to reduce the risk of getting statistically significant results with minimal deviations from the null hypothesis, variables were entered if the significance of their *t* value was less than 0.005 (entry level) and removed if the significance was greater than 0.01 (removal level).

Categorical variables were tested in comparison to a reference category represented by the largest category for nominal data (eg, payer categories) or the largest of the extreme (minimal or maximal) category for ordinal variables (eg, acuity). Multiple regression models in general, and stepwise procedures specifically, have a risk of overinterpretation based on the particular characteristics of the sample at hand, a phenomenon known as overfitting.[1] Because of the large sample size examined and the generous ratio of cases per number of predictors tested, the authors expected the risk of overfitting to be minimal, even when adopting strict criteria for the ratio between sample size and number of predictors.[26] Nonetheless, assessing for model overfitting—yielding findings that will not replicate in a different sample—is necessary.

Second, to assess for overfitting, the authors examined results from 3 cross-validation analyses using 2 randomly and evenly split samples: a development sample and a test sample. The researchers fit the stepwise regression model separately for the development and test samples. Variables that were significant in both samples were identified as being "stable" and tested in the final model. Next, the authors calculated the R^2 shrinkage[1] and the predictive ratio.[16] The R^2 shrinkage was assessed using several approaches. The authors compared the adjusted R^2 to the unadjusted R^2 results from the stepwise regression. The adjusted R^2 is an estimate of what the fit of the regression model would be if it were fitted against a new data set, assuming that all the degrees of freedom have been accounted for.[1] The authors then used the development sample to estimate the predicted functional status at discharge for the full sample (development and test samples). The predicted estimate was then fitted against the functional status scores at discharge using only the test sample. This study compared the predictive power (R^2) of the test sample, using a prediction model created using the development sample, to the R^2 of the development sample. Shrinkage is defined as the decrease in R^2 between the development sample and the test sample.

Although there are no clear standards for acceptable levels of shrinkage, the authors considered shrinkage of less than 10% to be sufficient to support the generalizability of the model's coefficients. As a confirmation analysis, a previously recommended bootstrap procedure[39] was applied using the "regvalidate" Stata (StataCorp LLC, College Station, TX) program.[11] To estimate the predictive ratio, the mean predicted discharge functional status scores of the test sample, estimated using the development sample, were divided by mean actual discharge functional status scores obtained from the test sample.[23] When the average predicted discharge functional status was close to the average actual discharge functional status, that is, the predictive ratio was close to 1, the predictive validity of the regression model was considered to be supported.[16,23]

Third, the final model's error terms (residuals) for the test sample were visually inspected to assess for normality and homoscedasticity—that is, deviations of the residuals are constant across the predicted outcome. Normality and homoscedasticity are assumptions of linear regression. The residual was the difference between the actual and predicted outcomes, with positive and negative residuals representing higher and lower outcomes, respectively. The authors preferred the visual inspection over statistical testing, because large data sets tend to have substantial power and can yield statistically significant results when there are only trivial deviations from normality and homoscedasticity. Normality was inspected by plotting a normal distribution line against the distribution of the residuals. Homoscedasticity was inspected by fitting a regression line to the squared residuals across the predicted outcome. A horizontal fitted line supports homoscedasticity.

Impact of Risk Adjustment on Clinic Ranking

To assess impact of risk adjustment on clinic ranking, this study compared unadjusted to risk-adjusted outcomes. For unadjusted outcomes, the authors used each clinic's mean raw functional status score at discharge. For risk-adjusted outcomes, they used each clinic's mean residual score, because the residual reflects the difference between the predicted discharge score and the actual discharge score.

Unadjusted and risk-adjusted outcomes were compared using 2 ranking methods. First, clinics were ranked by percentiles and further divided into 10 equal groups (deciles). The decision to examine by deciles was arbitrary and based on an assumption that 10 ranking groups would represent a categorization that was easy to interpret and meaningful to clinicians, managers, and payers.

Second, uncertainty assessments, as recommended previously,[25] were used to rank clinics into 3 significantly different quality levels (high, average, or low) based on the 95% confidence interval (CI) of each clinic's average unadjusted or risk-adjusted outcome. For clarity, the authors refer to this ranking method as quality ranking, implying that each rank represents a statistically unique quality level. Percentile-based ranking does not assume that ranks are statistically unique. A quality ranking of high or low, respectively, was achieved when the clinic's entire 95% CI range fell above or below that of the average.[15] All remaining clinics were ranked as average. For each of these 2 ranking methods, the researchers assessed percent agreement and chance-corrected agreement (using Cohen's kappa) of unadjusted and risk-adjusted rankings.

All analyses were conducted using Stata Version 14 (StataCorp LLC). The University of Utah Institutional Review Board approved the study protocol.

RESULTS

Patient Sample

The primary sample of patients who completed the patient-reported outcome measure data at intake included 618 199 episodes of care. From these patients, 414 125 (mean ± SD age, 57 ± 17 years; 60% female) completed the LCAT at admission and discharge, representing a completion rate of 67%. After applying the 2 criteria for inclusion in the ranking analyses, the remaining sample included 341 642 patients treated by 6934 clinicians in 2107 clinics from all the states and the District of Columbia. A diagram showing the progression of the study sample is presented in **FIGURE 1**.

Patients selected (n = 341 642) or not selected (n = 72 483) for ranking had practically identical functional status intake scores, age, and acuity levels. Other differences identified between selected and excluded samples were identified but were interpreted as having negligible clinical importance. Patients selected for ranking who had complete outcomes data (n = 341 642), compared to patients with incomplete data (n = 204 074), had similar but slightly higher functional status scores at admission, were older and slightly less chronic, and had a higher rate of workers' compensation and auto insurance payer types, more surgeries related to their LBP, a higher rate of exercise history, a lower rate of depression, and a slightly higher rate of diabetes (**TABLE 1**). The comparison of predictions created by the unweighted and weighted models using inverse probability weighting resulted in practically identical mean and median predictions, with no impact of inverse probability weighting on ranking

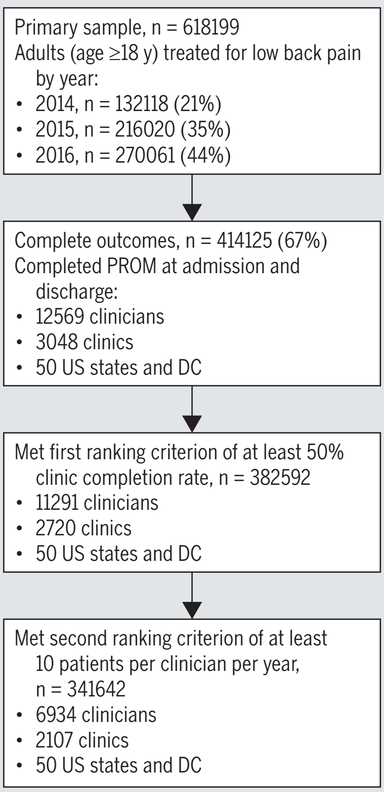


FIGURE 1. Patient sampling method. Abbreviation: PROM, patient-reported outcome measure.

TABLE 2 Risk-Adjusted Model: Associations Between Patient Characteristics at Admission and FS at Discharge*

Significant Predictors of FS at Discharge	β[†]	t[‡]
Intercept	42.4 (42.1, 42.7)	280.9
FS score at admission[§]	0.6 (0.6, 0.6)	320.6
Age	−0.1 (−0.1, −0.1)	−69.3
Sex: female	−0.3 (−0.4, −0.3)	−8.0
Acuity (>6 mo)		
0-7 d	12.5 (12.3, 12.7)	116.4
8-14 d	9.2 (9.0, 9.3)	105.8
15-21 d	7.0 (6.8, 7.1)	88.0
22-90 d	4.2 (4.1, 4.3)	78.7
91 d to 6 mo	1.8 (1.7, 1.9)	27.7
Payer (HMO, preferred provider)		
Indemnity insurance	−2.6 (−2.9, −2.4)	−22.5
Medicaid	−4.7 (−4.9, −4.5)	−47.7
Medicare A	−1.4 (−1.7, −1.1)	−8.4
Medicare B under age 65 y	−3.0 (−3.2, −2.8)	−28.2
No fault, auto insurance	−4.2 (−4.5, −3.8)	−25.3
Workers' compensation	−5.7 (−5.9, −5.5)	−64.0
Other (litigation, Medicare C, school, no charge, early intervention, commercial insurance)	−1.1 (−1.3, −1.0)	−15.0
Surgical history (no related surgery)		
1 related surgery	−1.8 (−1.9, −1.7)	−27.4
2 related surgeries	−2.9 (−3.1, −2.6)	−26.3
3 or more related surgeries	−3.7 (−4.0, −3.5)	−29.9
Exercise history (seldom or never)		
At least 3 times per week	1.3 (1.2, 1.4)	27.0
1-2 times per week	0.6 (0.5, 0.7)	12.1

Table continues on page 643.

change after risk adjustment (data available on request).

Risk-Adjustment Modeling

The risk-adjusted model, developed using all patients completing the LCAT at admission and discharge (n = 414 125), is described in **TABLE 2**. The dependent variable was functional status at discharge. The model identified 11 constructs that explained 37% of the variance in discharge functional status, with functional status at admission, acuity, payer type, and age being the most important predictors. Results from different approaches used to estimate R^2 shrinkage ranged between 0.0% and 0.1%. The researchers are not aware of an agreed-on value for an acceptable level of shrinkage. However, the authors considered shrinkage of less than 1% to strongly support the model's external validity. The average predicted discharge functional status of the test sample (n/2 = 207 063), estimated using the development sample, was practically identical to the average actual discharge functional status obtained by the test sample (62.743 and 62.737, respectively). The authors interpreted this as an almost perfect predictive ratio of 1.0, providing additional support for the predictive validity of the final model.

Plots of the model's residuals for normality and homoscedasticity are presented in **FIGURE 2** and **FIGURE 3**, respectively. The results supported normality, with only slight deviations. Residuals were consistent across the predicted functional status scores, supporting homoscedasticity.

Impact of Risk Adjustment on Clinic Ranking

The comparison of clinic ranking by deciles of unadjusted (raw) and risk-adjusted outcomes is presented in **TABLE 3**. Higher rankings represent higher outcomes, and cells represent number of clinics. The clinics along the diagonal (marked in bold) represent the agreement between unadjusted and risk-adjusted ranks, which was 30% (Cohen's κ = 0.22; 95% CI: 0.21,

TABLE 2 RISK-ADJUSTED MODEL: ASSOCIATIONS BETWEEN PATIENT CHARACTERISTICS AT ADMISSION AND FS AT DISCHARGE (CONTINUED)*

Significant Predictors of FS at Discharge	β†	t‡
Medication use at intake	−1.3 (−1.4, −1.2)	−29.9
Previous treatment	−1.5 (−1.6, −1.5)	−36.4
Lumbar surgery procedure (no surgical codes)		
Fusion	1.5 (1.2, 1.9)	9.2
Laminectomy/foramenectomy/discectomy	2.3 (1.9, 2.6)	13.5
Specific comorbidities		
Angina	−0.6 (−1.0, −0.3)	−3.9
Anxiety	−0.9 (−1.1, −0.8)	−14.7
Arthritis	−1.1 (−1.2, −1.0)	−23.2
Asthma	−0.3 (−0.4, −0.1)	−3.9
Chronic obstructive pulmonary disease	−1.0 (−1.2, −0.8)	−9.5
Depression	−1.1 (−1.2, −1.0)	−18.1
Diabetes type I or II	−0.6 (−0.7, −0.5)	−10.0
Headaches	−1.2 (−1.3, −1.1)	−23.3
Incontinence	−0.8 (−1.0, −0.6)	−9.1
Kidney, bladder, prostate, or urination problems	−0.4 (−0.5, −0.2)	−5.6
Neurological disease	−1.3 (−1.6, −1.0)	−8.7
Obesity (BMI ≥30 kg/m^2)	−0.6 (−0.7, −0.5)	−14.5
Osteoporosis	−0.5 (−0.6, −0.4)	−7.3
Previous accidents	−0.5 (−0.6, −0.4)	−8.7
Sleep dysfunction	−1.2 (−1.3, −1.1)	−22.3
Stroke	−0.5 (−0.7, −0.3)	−4.6

Abbreviations: BMI, body mass index; FS, functional status; HMO, health maintenance organization.
**Patients, n = 414 125. Adjusted R^2 = 37.3%. Reference group for categorical variables is in parentheses.*
†Coefficient indicating the amount of expected change in discharge FS given a 1-unit change in the value of the variable, given that all other variables in the model are held constant. Values in parentheses are 95% confidence interval.
‡Values indicate the importance of each independent variable for predicting discharge FS (dependent variable). All t values were significant at the .001 level.
§Higher FS scores represent higher levels of functioning.

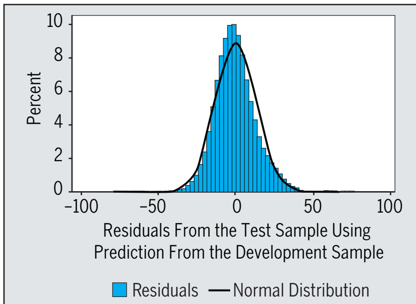


FIGURE 2. Visual inspection of normality of residuals. Distribution of the error term (residuals) from the risk-adjusted model, compared to the normal distribution. A distribution of residuals that is close to normal supports the normality assumption of linear regression. The mean was 0.09 and the median was 1.03.

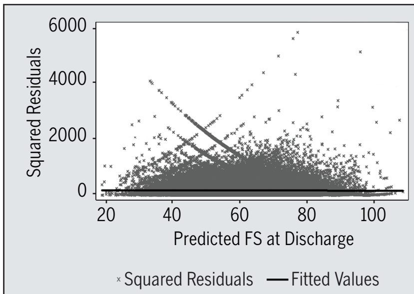


FIGURE 3. Visual inspection of homoscedasticity. Distribution of residuals (squared) across the range of the predicted FS scores at discharge. The fitted line represents fitted values for the squared residuals. A horizontal fitted line supports the homoscedasticity assumption of linear regression; that is, deviations of residuals are constant across the predicted outcome. Abbreviation: FS, functional status.

0.23; $P<.001$). Clinics above the diagonal had an increase in their decile rank in the risk-adjusted model, and clinics below the diagonal had a decrease in their decile rank in the risk-adjusted model. Ranking changed for 70% of clinics. The percent of clinics moving to higher or lower decile ranking categories by 1, 2, 3, or 4 or more ranks was 32%, 20%, 10%, and 8%, respectively (**FIGURE 4A**).

The comparison of clinic quality ranking for unadjusted and risk-adjusted outcomes is presented in **TABLE 4**. The agreement between unadjusted and risk-adjusted ranks by the 3 quality levels was 69.0% (Cohen's $\kappa = 0.5$; 95% CI: 0.47, 0.53; $P<.001$), with 31% of clinics moving up or down by 1 rank (**FIGURE 4B**). **FIGURE 5** illustrates how clinics were grouped into the 3 quality levels and were impacted by risk adjustment, with **FIGURE 5A** showing risk-adjusted average outcome and 95% CI, and **FIGURE 5B** showing the average risk-adjusted and unadjusted clinic outcomes. For comparison, both risk-adjusted and unadjusted outcomes were centered to zero, representing the overall average clinic outcome.

DISCUSSION

The authors compared clinic ranking based on the raw data versus the risk-adjusted data from patient-reported outcomes of functional status for patients with LBP treated in physical therapy. This study used 2 ranking methods, 10 equal groups (deciles) or 3 distinct quality groups (high, average, and low), to assess the impact of risk adjustment on clinic ranking. The researchers found that 70% of clinics changed decile rank and 31% changed quality grouping rank following risk adjustment. This supports the hypothesis that the ranks would be substantially different after risk adjustment. This study demonstrates the impact of risk adjustment on ranking for outpatient physical therapy clinics managing patients with LBP. Ranking that is based on unadjusted (raw) data would generate erroneous results and obscure meaningful interpretation of the quality of services by patients, payers, policy makers, and providers.

Ranking of performance is fundamentally different from ranking for other purposes, where examining unadjusted aggregated measures might be appropriate. For the purpose of identifying clinics that have the highest percentage of patients with a particular characteristic (eg, patients above 75 years of age), the unadjusted measure would suffice. In such a case, the purpose of ranking providers could be to identify their need for education on geriatric rehabilitation; thus, understanding the reasons older patients seek treatments in specific clinics would not be important. However, as shown by these results, it is essential to understand characteristics associated with the intended outcome and to control for those that are outside of the provider's influence when ranking clinics by their average clinical outcomes.

This risk-adjusted model included 11 constructs (functional status at admission, age, sex, acuity level, payer, surgical and exercise histories, medication use at intake, previous treatment, lumbar surgery procedures, and specific comorbidities) found to be significantly associated with functional status at discharge. The t values for the different coefficients allowed identification of the importance of the different predictors, with higher absolute

TABLE 3 RANKING COMPARISON BY DECILE (RAW/OLS)*

Unadjusted Rank (Raw)‡	Risk-Adjusted Rank (OLS)† 1	2	3	4	5	6	7	8	9	10
1	**131**	46	18	6	5	0	2	2	1	0
2	45	**54**	44	30	15	9	6	4	3	1
3	18	39	**51**	33	32	18	13	3	2	2
4	10	31	37	**33**	32	32	17	11	5	2
5	3	24	30	35	**32**	26	36	11	10	4
6	2	12	12	27	39	**44**	29	24	18	4
7	0	2	11	25	22	32	**44**	37	23	14
8	0	2	3	13	22	33	33	**43**	42	20
9	2	1	4	7	9	10	21	54	**68**	35
10	0	0	1	1	3	7	9	22	39	**128**

Abbreviation: OLS, ordinary least square.
**Values are number of clinics (total n = 2107). Ranks are assigned by decile, with higher ranks representing higher outcomes. Clinics marked in bold are those that did not change rank following risk adjustment.*
†Based on the average clinic residual from the OLS model.
‡Based on the raw score of functional status at discharge.

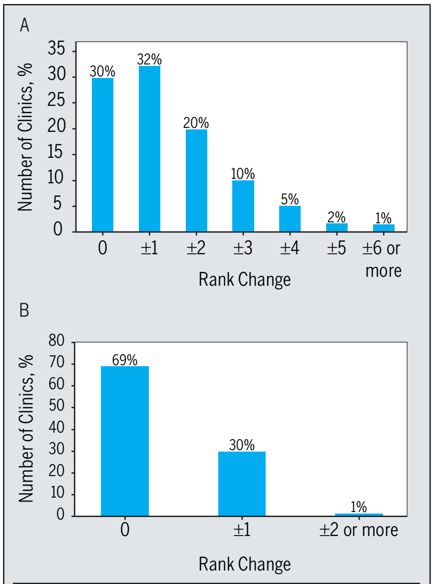


FIGURE 4. Percentage of clinics that changed rank following risk adjustment for (A) the decile ranking method and (B) a ranking method using 3 quality levels (high, average, and low).

values representing higher importance (**TABLE 2**). After functional status score at admission, the most important predictors were acuity, payer type, and age.

From a statistical perspective, a model including only these 4 constructs would retain most of the model's power while reducing the data-collection burden. However, retaining additional constructs affects the predicted outcomes of some patients. For example, depression was not 1 of the 4 constructs that had the greatest predictive power. With depression included and assuming all other modeled constructs were constant, the predicted functional status score at discharge for a patient with depression who had previous treatment for the condition would be 2.6 less than it would have been had these factors been left out of the model. A similar predicted outcome (–2.5 points) was found for patients with headaches and medication use at intake. Thus, inclusion of additional predictors had some impact on prediction and, importantly, may lead to greater acceptance by practicing clinicians.

From a clinician's perspective, a model that only includes predictors that contribute most to its predictive power may be deemed inadequate and unfair when caring for patients with a wide variety of complexities that impact the clinical presentation and expectations for treatment results. Successful translation of research into clinical practice becomes more feasible when front-line clinicians are assured that multiple factors are accounted for in a model used to assess their performance. This study does not provide a definitive list of factors to be risk adjusted, and other factors may be relevant, depending on the study population, clinicians, the outcome measure, or specific study purposes.

As illustrated by this study, different ranking methods can yield different results. Percentile-based ranking approaches seem intuitive and easy to comprehend and implement, but they are highly susceptible to the number of ranking categories used. For example, had the clinics been categorized by their actual percentile (100 equal groups) instead of 10 groups (deciles), then the impact of risk adjustment on clinic ranking would have been larger, because less change would be needed to change rank. Therefore, this study's choice of deciles should be considered arbitrary. Percentile ranking does not consider the amount of error in the estimates, so adjacent ranks cannot be assumed to be significantly different from each other. Thus, as previously recommended,[25] the authors also examined provider ranking that included uncertainty assessments (95% CIs of clinics' average outcomes). This ensured that differences in rank represented significant differences

TABLE 4 RANKING COMPARISON BY 3 OUTCOME LEVELS (RAW/OLS)*

Unadjusted Rank (Raw)‡	Risk-Adjusted Rank (OLS)† Low	Average	High
Low	**440**	196	11
Average	164	**746**	88
High	14	180	**268**

Abbreviation: OLS, ordinary least square.
**Values are number of clinics (total n = 2107). Ranks are assigned by 3 outcome levels. Clinics marked in bold are those that did not change rank following risk adjustment.*
†Based on the average clinic residual from the OLS model.
‡Based on the raw score of functional status at discharge.

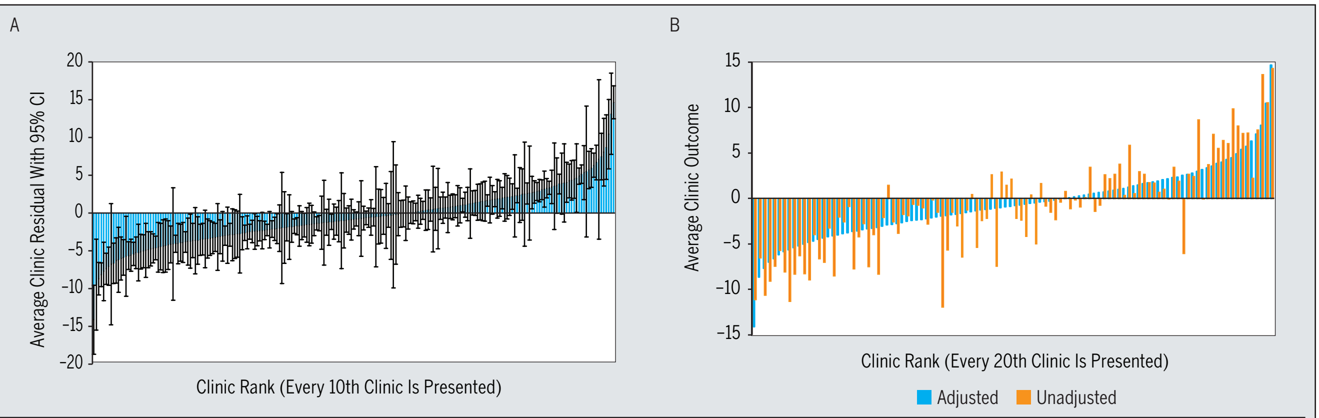


FIGURE 5. Clinic ranking using average clinic outcomes used to rank clinics into 3 quality levels (high, average, and low). Clinics were sorted in ascending order of their risk-adjusted outcomes. Outcomes were centered to zero to allow comparison of risk-adjusted and unadjusted outcomes. (A) The risk-adjusted clinic outcome with 95% CIs. Each bar represents a clinic. A high- or low-quality rank was achieved when the clinic's entire 95% CI range fell above or below zero, respectively. All remaining clinics were ranked as average. For clarity, only every 10th clinic was included. (B) Every clinic's risk-adjusted (blue bar) and unadjusted (orange bar) average outcome. Large deviations between adjusted and unadjusted outcomes (blue or orange bars per clinic) represent large impact of risk adjustment on clinic ranking. For clarity, only every 20th clinic was included. Abbreviation: CI, confidence interval.

in providers' quality, as defined by functional status score improvement.

The impact of risk adjustment on clinic ranking can also be studied using hierarchical multilevel modeling to account for nonrandom variation in the distribution of patients within clinicians and/or clinics.[15,42] Multilevel models aim to make up for unavailable data that could account for patients being treated by specific clinicians within specific clinics (patient nesting). Significant variance in outcomes attributed to patient nesting suggests that nesting is not random and, therefore, needs to be included in risk adjustment.[31]

The challenge in applying multilevel models for patient nesting is that such models adjust both for factors that are and are not influenced by providers. Some patient sociodemographic factors (eg, education, income level) may be associated both with functional outcomes and with which provider is seen. These factors are not within providers' control, and it would be appropriate to adjust for them in quality rankings. However, other variables are within the clinician's control and should not be "adjusted out" of the rankings. For example, clinicians' abilities to personally connect and communicate with their patients, clinical examination and intervention skills, and ease of scheduling convenient appointments can be influenced by providers. The use of multilevel models that account for nesting within clinics or providers could adjust for, and thus mask, quality differences among providers that are the target of the ranking. Because the data available in the current study did not allow for differentiation between these different types of nonrandom patient nesting, the authors do not present the multilevel modeling results here (results available on request). Studies comparing the impact of single-level and multilevel modeling on provider ranking are needed to improve the profession's ability to select the most appropriate modeling methods.

This study had some limitations. First, the authors assessed ranking based on patient-reported outcomes of functional status without incorporating cost data needed to implement value-based purchasing or pay-for-performance initiatives.[28,30,32,41] Studies examining the best methods to adjust for rehabilitation cost, incorporated within a value-based model, are warranted.[24]

Second, to minimize potential for patient selection bias, this study used strict inclusion criteria. This resulted in 18% of the complete outcomes data being excluded from the ranking analyses. Similar inclusion criteria have been proposed and used previously to strengthen the external validity of the study's sample by excluding clinics and clinicians that tend to include a minority of their patients in complete outcomes collection.[8] However, any patient censoring could introduce selection bias. Patients selected or not selected for ranking had practically identical functional status intake scores and acuity levels (**TABLE 1**), which reduced the potential for a systematic selection bias, as both factors are known to be among the stronger predictors of functional status outcomes.[9] Other differences between the included and excluded samples were judged to be trivial and without clinical relevance.

Another potential source of patient selection bias was patient censoring due to missing functional status at discharge, which precluded ranking based on functional status outcomes. The comparison of patients with incomplete and complete outcomes data (**TABLE 1**) showed some differences supporting and some not supporting the chance for a potential patient selection bias. For example, patients with complete outcomes data were slightly less chronic, had a higher rate of exercise history, and a lower rate of depression—characteristics found to be associated with higher functional status outcomes. However, these patients were also older, had a higher rate of workers' compensation and auto insurance payer types, and had more related surgeries and a slightly higher rate of diabetes—characteristics associated with lower functional status outcomes. Additionally, the lack of impact of inverse probability weighting on ranking change after risk adjustment supported mostly random patient censoring, reducing the potential for a systematic patient selection bias. However, inverse probability weighting might not have been able to correct for nonrandom patient censoring using the available data. In this case, selection bias might still exist and is, therefore, acknowledged here as a limitation.

CONCLUSION

Our study demonstrated important changes in provider ranking when a risk-adjustment method was used to account for basic patient characteristics at admission to physical therapy for patients treated for LBP. The results support the need for risk adjustment when profiling providers based on their patients' outcomes. Failing to do so could discourage clinicians from treating sicker patients with characteristics known to be associated with lower predicted functional status outcomes. Additional risk-adjustment studies should be conducted using sophisticated scientific approaches that ensure minimal bias when evaluating the performance of clinical care providers. ◉

KEY POINTS

FINDINGS: Important changes in provider ranking were identified when adjusting for basic patient characteristics at admission to physical therapy for patients treated for low back pain.
IMPLICATIONS: Results emphasize the need for risk adjustment when profiling providers based on their patients' outcomes to avoid discouraging clinicians from treating patients with characteristics known to be associated with lower predicted functional status outcomes.
CAUTION: The use of ranking methods other than those applied in this study might have generated different results.

ACKNOWLEDGMENTS: *The authors thank the thousands of rehabilitation therapists and their facilities across the United States engaged in ongoing outcomes data collection* *for the benefit of their patients. The researchers specifically thank Dr Philip B. Ender from the Social Research Methodology Division, University of California at Los Angeles Department of Education, and Dr Pedro L. Gozalo from the Department of Health Services, Policy, and Practice, School of Public Health, Brown University, for their guidance with statistical methods described in this study.*

REFERENCES

1. Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom Med*. 2004;66:411-421.
2. Beattie P, Maher C. The role of functional status questionnaires for low back pain. *Aust J Physiother*. 1997;43:29-38. https://doi.org/10.1016/S0004-9514(14)60400-1
3. Beurskens AJ, de Vet HC, Köke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain*. 1996;65:71-76. https://doi.org/10.1016/0304-3959(95)00149-2
4. Centers for Medicare and Medicaid Services. Roadmap for Implementing Value Driven Healthcare in the Traditional Medicare Fee-for-Service Program. Baltimore, MD: Centers for Medicare and Medicaid Services; 2013.
5. Childs JD, Harman JS, Rodeghero JR, Horn M, George SZ. Implications of practice setting on clinical outcomes and efficiency of care in the delivery of physical therapy services. *J Orthop Sports Phys Ther*. 2014;44:955-963. https://doi.org/10.2519/jospt.2014.5224
6. Cleland JA, Fritz JM, Brennan GP, Magel J. Does continuing education improve physical therapists' effectiveness in treating neck pain? A randomized clinical trial. *Phys Ther*. 2009;89:38-47. https://doi.org/10.2522/ptj.20080033
7. Deutscher D, Hart DL, Dickstein R, Horn SD, Gutvirtz M. Implementing an integrated electronic outcomes and electronic health record process to create a foundation for clinical practice improvement. *Phys Ther*. 2008;88:270-285. https://doi.org/10.2522/ptj.20060280
8. Deutscher D, Hart DL, Stratford PW, Dickstein R. Construct validation of a knee-specific functional status measure: a comparative study between the United States and Israel. *Phys Ther*. 2011;91:1072-1084. https://doi.org/10.2522/ptj.20100175
9. Deutscher D, Horn SD, Dickstein R, et al. Associations between treatment processes, patient characteristics, and outcomes in outpatient physical therapy practice. *Arch Phys Med Rehabil*. 2009;90:1349-1363. https://doi.org/10.1016/j.apmr.2009.02.005
10. Dieleman JL, Baral R, Birger M, et al. US spending on personal health care and public health, 1996-2013. *JAMA*. 2016;316:2627-2646. https://doi.org/10.1001/jama.2016.16885
11. Ender P. Cross validation. Available at: http://www.philender.com/courses/linearmodels/notes2/cross.html. Accessed December 31, 2010.
12. Fritz JM, Childs JD, Wainner RS, Flynn TW. Primary care referral of patients with low back pain to physical therapy: impact on future health care utilization and costs. *Spine (Phila Pa 1976)*. 2012;37:2114-2121. https://doi.org/10.1097/BRS.0b013e31825d32f5
13. Fritz JM, Wainner RS, Hicks GE. The use of nonorganic signs and symptoms as a screening tool for return-to-work in patients with acute low back pain. *Spine (Phila Pa 1976)*. 2000;25:1925-1931.
14. Godges JJ, Anger MA, Zimmerman G, Delitto A. Effects of education on return-to-work status for people with fear-avoidance beliefs and acute low back pain. *Phys Ther*. 2008;88:231-239. https://doi.org/10.2522/ptj.20050121
15. Gozalo PL, Resnik LJ, Silver B. Benchmarking outpatient rehabilitation clinics using functional status outcomes. *Health Serv Res*. 2016;51:768-789. https://doi.org/10.1111/1475-6773.12344
16. Hart DL, Connolly JB. Pay-for-Performance for Physical Therapy and Occupational Therapy: Medicare Part B Services. Knoxville, TN: Focus On Therapeutic Outcomes Inc; June 1, 2006.
17. Hart DL, Deutscher D, Werneke MW, Holder J, Wang YC. Implementing computerized adaptive tests in routine clinical practice: experience implementing CATs. *J Appl Meas*. 2010;11:288-303.
18. Hart DL, Mioduski JE, Werneke MW, Stratford PW. Simulated computerized adaptive test for patients with lumbar spine impairments was efficient and produced valid measures of function. *J Clin Epidemiol*. 2006;59:947-956. https://doi.org/10.1016/j.jclinepi.2005.10.017
19. Hart DL, Werneke MW, Wang YC, Stratford PW, Mioduski JE. Computerized adaptive test for patients with lumbar spine impairments produced valid and responsive measures of function. *Spine (Phila Pa 1976)*. 2010;35:2157-2164. https://doi.org/10.1097/BRS.0b013e3181cbc17f
20. Iezzoni LI. Risk adjusting rehabilitation outcomes: an overview of methodologic issues. *Am J Phys Med Rehabil*. 2004;83:316-326. https://doi.org/10.1097/01.PHM.0000118041.17739.BB
21. Iezzoni LI. Using risk-adjusted outcomes to assess clinical practice: an overview of issues pertaining to risk adjustment. *Ann Thorac Surg*. 1994;58:1822-1826. https://doi.org/10.1016/0003-4975(94)91721-3
22. Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academies Press; 2001.
23. Kautter J, Ingber M, Pope GC, Freeman S. Improvements in Medicare Part D risk adjustment: beneficiary access and payment accuracy. *Med Care*. 2012;50:1102-1108. https://doi.org/10.1097/MLR.0b013e318269eb20
24. Lentz TA, Harman JS, Marlow NM, George SZ. Application of a value model for the prevention and management of chronic musculoskeletal pain by physical therapists. *Phys Ther*. 2017;97:354-364.
25. Lin R, Louis TA, Paddock SM, Ridgeway G. Ranking USRDS provider specific SMRs from 1998-2001. *Health Serv Outcomes Res Methodol*. 2009;9:22-38. https://doi.org/10.1007/s10742-008-0040-0
26. Maxwell SE. Sample size and multiple regression analysis. *Psychol Methods*. 2000;5:434-458. https://doi.org/10.1037/1082-989X.5.4.434
27. National Quality Forum. Patient Reported Outcomes (PROs) in Performance Measurement. Washington, DC: National Quality Forum; January 10, 2013.
28. Porter ME. What is value in health care? *N Engl J Med*. 2010;363:2477-2481. https://doi.org/10.1056/NEJMp1011024
29. Porter ME, Larsson S, Lee TH. Standardizing patient outcomes measurement. *N Engl J Med*. 2016;374:504-506. https://doi.org/10.1056/NEJMp1511701
30. Porter ME, Teisberg EO. *Redefining Health Care: Creating Value-Based Competition on Results*. Boston, MA: Harvard Business School Press; 2006.
31. Rabe-Hesketh S, Skrondal A. *Multilevel and Longitudinal Modeling Using Stata, Volumes I and II*. 3rd ed. College Station, TX: Stata Press; 2012.
32. Raso R. Value-based purchasing: are you part of the shift? *Nurs Manage*. 2015;46:24-31; quiz 31-32. https://doi.org/10.1097/01.NUMA.0000463882.54082.64
33. Resnik L, Gozalo P, Hart DL. Weighted index explained more variance in physical function than an additively scored functional comorbidity scale. *J Clin Epidemiol*. 2011;64:320-330. https://doi.org/10.1016/j.jclinepi.2010.02.019
34. Resnik L, Hart DL. Using clinical outcomes to identify expert physical therapists. *Phys Ther*. 2003;83:990-1002. https://doi.org/10.1093/ptj/83.11.990
35. Resnik L, Liu D, Hart DL, Mor V. Benchmarking physical therapy clinic performance: statistical methods to enhance internal validity when using observational data. *Phys Ther*. 2008;88:1078-1087. https://doi.org/10.2522/ptj.20070327
36. Riddle DL, Stratford PW. Use of generic versus region-specific functional status measures on patients with cervical spine disorders. *Phys Ther*. 1998;78:951-963. https://doi.org/10.1093/ptj/78.9.951
37. Robins JM, Hernán MÁ, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11:550-560.
38. Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. *Stat Methods Med Res*. 2013;22:278-295. https://doi.org/10.1177/0962280210395740

39. Steyerberg EW, Harrell FE, Jr., Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001;54:774-781. https://doi.org/10.1016/S0895-4356(01)00341-9
40. Swinkels IC, van den Ende CH, de Bakker D, et al. Clinical databases in physical therapy. *Physiother Theory Pract*. 2007;23:153-167. https://doi.org/10.1080/09593980701209097
41. Tanenbaum SJ. What is the value of value-based purchasing? *J Health Polit Policy Law*. 2016;41:1033-1045. https://doi.org/10.1215/03616878-3632254
42. Teppala S, Ottenbacher KJ, Eschbach K, et al. Variation in functional status after hip fracture: facility and regional influence on mobility and self-care. *J Gerontol A Biol Sci Med Sci*. 2017;72:1376-1382. https://doi.org/10.1093/gerona/glw249
43. Van Der Wees PJ, Nijhuis-Van Der Sanden MW, Ayanian JZ, Black N, Westert GP, Schneider EC. Integrating the use of patient-reported outcomes for both clinical practice and performance measurement: views of experts from 3 countries. *Milbank Q*. 2014;92:754-775. https://doi.org/10.1111/1468-0009.12091
44. Wang YC, Hart DL, Werneke M, Stratford PW, Mioduski JE. Clinical interpretation of outcome measures generated from a lumbar computerized adaptive test. *Phys Ther*. 2010;90:1323-1335. https://doi.org/10.2522/ptj.20090371
45. Weldring T, Smith SM. Patient-reported outcomes (PROs) and patient-reported outcome measures (PROMs). *Health Serv Insights*. 2013;6:61-68. https://doi.org/10.4137/HSI.S11093

CARA L. LEWIS, PT, PhD[1] • ANNE KHUU, PhD[1] • KARI L. LOVERRO, PhD[1]

Gait Alterations in Femoroacetabular Impingement Syndrome Differ by Sex

Femoroacetabular impingement (FAI) syndrome is an increasingly recognized clinical diagnosis for hip pain in young and middle-aged adults. Individuals who present with hip pain in combination with structural hip morphology thought to contribute to premature contact between the proximal femur and acetabulum are classified as having FAI syndrome.[17] While there is agreement that movement contributes to FAI syndrome,[17] a very small percentage of the current research evaluates factors beyond available hip joint range of motion. Among the few studies that examine functional movement, gait has been evaluated more than any other task[5,13,19,22,24,36]; however, the picture of how gait is altered in the presence of FAI syndrome remains unclear. Some of the variability in findings for hip and pelvic kinematics among studies could be due to individual differences in walking speeds, as studies to date have used a self-selected speed for testing.

A substantial limitation of the current gait studies in individuals with FAI syndrome is that the majority of the participants in these studies are males, even though females comprise an equal or greater percentage of the surgical population.[9] Additionally, these studies report only on the involved or painful limb during gait, despite the frequent presence bilaterally of structural morphology consistent with FAI syndrome.

Given these limitations, a study that contributes to the understanding of gait in both males and females with FAI syndrome is warranted. Therefore, the purpose of this study was to evaluate sex-specific differences in individuals with FAI syndrome compared to individuals without hip pain, walking at both their preferred speed as well as a prescribed speed. The authors anticipated that some gait alterations would be consistent across the sexes, and that some differences would be unilateral (limb specific) and others would be bilateral (person specific).

BACKGROUND: Femoroacetabular impingement (FAI) syndrome may affect gait kinematics differently between males and females.

OBJECTIVES: To investigate whether individuals with FAI syndrome have different hip and pelvic motion during gait, at their preferred speed and a prescribed speed, compared to individuals of the same sex without pain.

METHODS: Twenty-one participants (11 males and 10 females) with FAI syndrome and 41 participants (19 males and 22 females) without hip pain were included in this case-control laboratory study. There were no differences between the 2 groups in age, body mass index, and activity score. Kinematic data for all participants were collected while walking at a preferred speed and at 1.25 m/s. For sex and walking speed, linear regression analyses were used to examine the effect of group and the interaction of group by limb.

RESULTS: At both speeds, males with FAI syndrome walked with more than 6° less peak hip extension ($P \leq .018$), 5° greater anterior pelvic tilt ($P \leq .020$), and 5° less posterior pelvic tilt ($P \leq .018$) compared to males without hip pain. Females with FAI syndrome walked with 2° less hip extension ($P \leq .012$) and at least 3° more hip adduction ($P < .001$) in the more painful hip than in the less painful hip at both speeds.

CONCLUSION: Males and females with FAI syndrome have different gait alterations when compared to a same-sex comparison group. In males, differences were between groups. In females with FAI syndrome, differences were between the more painful and the less painful limb. *J Orthop Sports Phys Ther 2018;48(8):649-658. Epub 22 May 2018. doi:10.2519/jospt.2018.7913*

KEY WORDS: *biomechanics, FAIS, gait, hip pain, movement system*

METHODS

Participants

Using an a priori power analysis on peak hip extension, a group mean ± SD difference of 4.8° ± 3.2° in hip extension angles[21] during natural treadmill walking at 1.25 m/s was anticipated.[26] Accordingly, to achieve statistical power of 0.80 with an alpha of .05, a minimum of 8 participants of each sex for the FAI syndrome group and for the healthy comparison group were needed.

To be a participant in either group, individuals had to be between 14 and 50

[1]Department of Physical Therapy and Athletic Training, PhD Program in Rehabilitation Sciences, College of Health and Rehabilitation Sciences, Sargent College, Boston University, Boston, MA. This study was approved by the Institutional Review Boards of Boston University and Boston Children's Hospital. Research reported in this manuscript was supported by the Peter Paul Career Development Professorship and the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under award numbers R21 AR061690, K23 AR063235, and P60 AR047785. 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 Cara L. Lewis, 635 Commonwealth Avenue, Boston, MA 02215. E-mail: lewisc@bu.edu

years of age and report being able to walk safely for at least 10 minutes without an assistive device. Individuals with a history of neurological disorder or back surgery or with current back, knee, or ankle pain were excluded from participation in either group.

Individuals with FAI syndrome were recruited through area orthopaedic and rehabilitation clinics between January 2011 and December 2016. To be included in the FAI syndrome group, individuals had to have been diagnosed with cam, pincer, or mixed FAI syndrome by a physician and had to have their pain reproduced by at least 1 of 3 provocative tests performed during the study visit: (1) the flexion, adduction, internal rotation (FADIR) test; (2) the flexion, abduction, external rotation test; and (3) the resisted straight leg raise. For the FADIR test, the hip was passively flexed to 90° and then adducted and internally rotated.[16] For the flexion, abduction, external rotation test, the hip was passively positioned in flexion, abduction, and external rotation, with the foot of the tested leg on top of the contralateral knee.[38] For the resisted straight leg raise, the leg was passively positioned in 30° of hip flexion with the knee extended.[32] The participant was then asked to keep the leg in that position without assistance and continue to hold the position as resistance was applied at the distal leg. When the test reproduced the individual's pain, the test was considered positive.

While these tests are highly sensitive for intra-articular hip pathology, they have low specificity.[32,34] Therefore, they were used as screening tests to eliminate individuals in the hip pain group when the results of all the tests were negative (suggesting no hip involvement) and to eliminate individuals in the comparison group when a test was positive (suggesting hip involvement, despite the lack of self-reported symptoms). Exclusion criteria included current or recent (within the last 2 months) lower extremity injury, history of lower extremity orthopaedic surgery, history of hip pain, and hip or groin pain or discomfort during any of the provocative tests.

This study was approved by the Institutional Review Boards of Boston University and Boston Children's Hospital, and all individuals provided written informed consent prior to participation. Data from some of the participants included in this study have been published elsewhere.[29]

Instrumentation

As this was part of a larger study for multiple functional tasks, the authors recorded whole-body kinematic data of the trunk, pelvis, and lower extremity using a 10-camera motion-capture system (Oxford Metrics, Yarnton, UK) sampling at 100 Hz. Participants walked on an instrumented split-belt treadmill (Bertec Corporation, Columbus, OH) sampling at 1000 Hz. Retroreflective markers were placed over 30 bony landmarks on the trunk and pelvis and bilaterally on the lower extremities, along with rigid clusters of markers on the thighs and shanks as previously described.[27]

Questionnaires

All participants completed self-report questionnaires, including the University of California at Los Angeles activity score,[3] the modified Harris Hip Score,[6] and the Hip disability and Osteoarthritis Outcome Score.[23] The Western Ontario and McMaster Universities Osteoarthritis Index was scored from the Hip disability and Osteoarthritis Outcome Score.[33] The University of California at Los Angeles activity score is scored from 1 to 10, with 10 being most active. All other questionnaires were scored from 0% to 100%, with 100% corresponding to excellent or no limitations.

Experimental Protocol

For testing, all participants wore a tight-fitting shirt, spandex shorts, and their own exercise shoes. Prior to data collection, the 3 provocative hip tests were performed on each participant. Preferred walking speed was calculated from the average of 5 trials walking a 5-m distance in the lab. The researchers placed reflective markers on each participant and then collected a static calibration trial, with the participant standing in a neutral posture with feet shoulder-width apart and shoulders in approximately 90° of abduction. Joint centers for the hips and knees were created using this trial, but were not normalized in this position. The authors removed the medial knee and ankle markers after this trial.

Participants walked on the treadmill at their preferred speed and at a prescribed speed of 1.25 m/s. After the treadmill achieved the set speed and the participant acclimated, data were recorded for up to 120 seconds of continuous walking. At least 50 strides were used for analysis at each speed, with a median of 90 strides. Strides were excluded from analysis if marker data were missing. The preferred speed was collected first to capture the individual's natural pattern before enforcing the speed constraint. As walking speed affects gait kinematics,[8,25] the prescribed speed was used to obtain kinematics at a standard walking speed. Every 30 seconds, each participant was asked to verbally rate his or her pain on an 11-point (0 is no pain and 10 is extreme pain) numeric rating scale.[14]

Data Analysis

Motion-capture data were processed as previously described.[27] Briefly, marker trajectories were labeled and gaps were filled using Vicon Nexus (Oxford Metrics). Marker and ground reaction force data were filtered using a low-pass, fourth-order Butterworth filter with a cutoff frequency of 6 Hz and 10 Hz, respectively. A participant-specific, 8-segment hybrid model was created in Visual3D (C-Motion, Inc, Germantown, MD) using the CODA pelvis model to define the pelvis and the hip joint centers. Kinematics of the hip, pelvis, and thigh were calculated. Pelvic and thigh segment angles were defined with respect to the laboratory coordinate system. Hip joint angles were defined as the angle between the thigh and pelvis segments. Ground reaction force data were used to determine heel strike.

For each stride, hip, pelvic, and thigh angles were normalized to the gait cycle (heel strike to ipsilateral heel strike). The authors extracted the dependent variables of interest, which included peak hip, pelvic, and thigh angles in the sagittal and frontal planes. The peak angles for each stride were then averaged together for each limb, and the average was used for statistical analysis.

Statistical Analysis

As there are well-documented differences in hip and pelvic kinematics during gait due to speed[8,25] and sex,[28] the researchers performed separate analyses for each speed, as well as for males and females. This study used a linear regression analysis, with group (FAI syndrome versus comparison) as the between-participant factor and limb (more painful versus less painful) as the within-participant factor. For participants with FAI syndrome, the limb with worse self-reported symptoms was the more painful limb. For participants without hip pain, the side designated as more painful was randomly distributed between the left and right sides similar to the distribution of the more painful side in the participants with FAI syndrome. As each limb was included in the analysis and the group sizes were uneven, a generalized estimating equation correction was applied to the linear regression model.[31]

To understand the effects of group and limb, the authors analyzed the main effect of group (FAI syndrome versus comparison) and the interaction of group by limb. A separate generalized estimating equation was performed for each dependent variable. If the group-by-limb interaction was significant, then the researchers performed 2 subsequent analyses. First, they used least-significant-difference pairwise comparisons to analyze the difference between the more painful limb and the less painful limb in individuals with FAI syndrome. Second, they calculated the average of the 2 limbs for each dependent variable for the comparison group, and used pairwise comparisons to analyze the difference between the more painful limb of individuals with FAI syndrome and the average of the 2 limbs for the individuals without hip pain. For each significant pairwise comparison, Cohen's *d* was used to compute the effect size (ES), interpreted as small (0.2), medium (0.5), and large (0.8) effects.[11] Differences less than 1.4° were not interpreted, as this has been reported as the minimal detectable change for hip kinematics in a single testing session.[39] All analyses were run in SPSS Statistics Version 20 (IBM Corporation, Armonk, NY).

RESULTS

The present study included 21 participants (11 males and 10 females) with FAI syndrome and 41 participants (19 males and 22 females) without hip pain (**TABLE 1**). The groups (FAI syndrome and comparison) were not different in terms of height, mass, body mass index, activity score, or preferred walking speed. The majority of the individuals with FAI syndrome had cam morphology. Of the 11 males with FAI syndrome, 7 reported symptoms on 1 limb only; of the 10 females with FAI syndrome, 7 reported symptoms bilaterally. For both males and females, the FADIR test was positive in most individuals (**TABLE 2**). Of the participants with FAI syndrome who reported pain during walking, the average ± SD pain ratings for males and females at the preferred speed were 1.8 ± 0.5 and 2.8 ± 1.3, respectively, and at the prescribed speed were 2.5 ± 1.3 and 1.7 ± 0.8, respectively. The individuals with FAI syndrome generally scored lower on the self-report questionnaires than the individuals without hip pain (**TABLE 3**).

Sex-Specific Analyses: Males

There were significant group differences at the hip and pelvis, but not at the thigh (**TABLES 4** through **6**). There were only 2 significant group-by-limb interactions at the pelvis (**TABLE 5**). No other interactions were significant.

Hip Males with FAI syndrome had 6.0° more peak hip flexion than the comparison group when walking at the preferred speed (95% confidence interval [CI]: 0.4°, 11.5°; P = .035; ES, 0.71) (**TABLE 4**, **FIGURE 1**). Males with FAI syndrome also walked with 8.2° less peak hip extension than males without hip pain at the preferred walking speed (95% CI: 2.8°, 13.5°; P = .003; ES, 1.01), and 6.9° less peak hip extension at the prescribed walking speed

TABLE 1 DEMOGRAPHIC DATA*

	Males		Females	
	FAI Syndrome (n = 11)	Comparison (n = 19)	FAI Syndrome (n = 10)	Comparison (n = 22)
Age, y	25.3 ± 8.0	25.1 ± 6.2	20.7 ± 6.8	22.5 ± 2.8
Height, m	1.80 ± 0.08	1.79 ± 0.07	1.67 ± 0.06	1.64 ± 0.07
Mass, kg	82.0 ± 9.1	77.4 ± 12.2	62.5 ± 7.2	59.9 ± 8.3
BMI, kg/m²	25.4 ± 1.6	24.2 ± 3.1	22.4 ± 2.2	22.4 ± 2.6
UCLA activity score†	9.5 (5-10)	9 (5-10)	8 (4-10)	9 (4-10)
Preferred walking speed, m/s	1.26 ± 0.17	1.27 ± 0.18	1.31 ± 0.14	1.28 ± 0.16

Abbreviations: BMI, body mass index; FAI, femoroacetabular impingement; UCLA, University of California at Los Angeles.
**Values are mean ± SD and analyzed with independent-samples t tests unless otherwise indicated. There were no significant differences between groups for all variables (P>.05).*
†Values are median (range) and analyzed with the Mann-Whitney U test. Data were missing for 1 male with FAI syndrome and 1 female without hip pain.

(95% CI: 1.2°, 12.6°; P = .018; ES, 0.83). Males with FAI syndrome had 1.4° less peak hip abduction than males without hip pain at the prescribed speed (95% CI: 0.02°, 2.7°; P = .047; ES, 0.74).

Pelvis At the pelvis, there were group differences in peak posterior pelvic tilt and peak anterior pelvic tilt in males at both speeds (**TABLE 5**, **FIGURE 2**). Males with FAI syndrome walked with 5.3° less peak posterior pelvic tilt than males without hip pain at the preferred speed (95% CI: 0.9°, 9.7°; P = .018; ES, 0.80), and 5.4° less peak posterior pelvic tilt at the prescribed speed (95% CI: 1.1°, 9.8°; P = .015; ES, 0.83). Males with FAI syndrome also walked with 5.4° more peak anterior pelvic tilt than males without hip pain at the preferred speed (95% CI: 1.0°, 9.9°; P = .017; ES, 0.83), and 5.3° more peak anterior pelvic tilt at the prescribed speed (95% CI: 0.8°, 9.8°; P = .020; ES, 0.81). There were significant group-by-limb interactions for peak pelvic hike (P = .033) and drop (P = .033) at the prescribed speed; no differences were noted in the subsequent pairwise analyses.

Thigh There were no significant group differences or group-by-limb interactions for the thigh in either plane at either speed (**TABLE 6**, **FIGURE 3**).

Sex-Specific Analyses: Females

There were no significant group effects (P≥.069) for any of the variables in females, but there were significant group-by-limb interactions at the hip, pelvis, and thigh (**TABLES 4** through **6**).

Hip At the hip, there were significant group-by-limb interactions for peak hip extension (P = .033 and P = .010 for the preferred and prescribed speeds, respectively), peak hip adduction (P<.001 and P<.001), and peak hip abduction (P = .014 and P = .004) (**TABLE 4**, **FIGURE 1**). In the subsequent pairwise analyses, there were differences between the more painful limb and the less painful limb in females with FAI syndrome. In the FAI syndrome group, individuals walked with 1.8° less peak hip extension on the more painful limb than on the less painful limb at the preferred speed (95% CI: 0.4°, 3.2°; P = .012; ES, 0.75), and 2.1° less peak hip extension on the more painful limb at the prescribed speed (95% CI: 0.7°, 3.4°; P = .004; ES, 0.87). In the

TABLE 2 Number of Individuals With FAI Syndrome Who Had Pain With Provocative Tests and During Gait*

	Males (n = 11)		Females (n = 10)	
	More Painful Limb	Less Painful Limb	More Painful Limb	Less Painful Limb
Provocative test				
FADIR	10 (91)	5 (45)	10 (100)	8 (80)
FABER	5 (45)	1 (9)	4 (40)	1 (10)
SLR	5 (45)	3 (27)	6 (60)	4 (40)
Gait				
Preferred	4 (36)	2 (18)	5 (50)	2 (20)
Prescribed	4 (36)	2 (18)	6 (60)	2 (20)

Abbreviations: FABER, flexion, abduction, external rotation test; FADIR, flexion, adduction, internal rotation test; FAI, femoroacetabular impingement; SLR, straight leg raise resisted at 30°.
**Values are n (percent). Individuals in the comparison group did not have any positive (painful) tests.*

TABLE 3 Data From Self-report Questionnaire Scores*

	FAI Syndrome				Comparison	
	Males (n = 10)†		Females (n = 10)			
Questionnaire	More Painful Limb	Less Painful Limb	More Painful Limb	Less Painful Limb	Males (n = 19)	Females (n = 21)†
mHHS	76.9 ± 12.7	94.1 ± 9.4	71.1 ± 20.4	84.9 ± 16.9	100.0 ± 0.0	99.8 ± 1.2
HOOS subscales‡						
Pain	70.5 ± 18.2	92.8 ± 13.9	68.3 ± 16.9	91.5 ± 8.4	100.0 ± 0.0	100.0 ± 0.0
Symptoms	63.0 ± 13.2	88.5 ± 13.6	68.5 ± 14.2	87.0 ± 12.5	99.0 ± 2.1	97.3 ± 4.4
Functional activities	84.4 ± 12.3	95.9 ± 8.3	84.9 ± 12.3	97.1 ± 5.3	99.9 ± 0.3	100.0 ± 0.0
Recreation/sport activities	68.8 ± 20.8	86.9 ± 15.4	66.3 ± 21.3	90.6 ± 13.9	100.0 ± 0.0	99.7 ± 1.3
Quality of life	49.4 ± 19.4	82.5 ± 19.5	43.8 ± 25.3	81.3 ± 17.7	100.0 ± 0.0	99.7 ± 1.3
WOMAC‡	81.7 ± 13.4	95.6 ± 9.1	81.5 ± 10.9	96.1 ± 5.6	99.9 ± 0.2	100.0 ± 0.0

Abbreviations: FAI, femoroacetabular impingement; HOOS, Hip disability and Osteoarthritis Outcome Score; mHHS, modified Harris Hip Score; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.
**Values are mean ± SD.*
†Data were missing for 1 male with FAI syndrome and 1 female in the control group.
‡Questionnaire scores range from 0% to 100%, with 100% corresponding to excellent or no limitations.

frontal plane, females with FAI syndrome walked with 3.8° more peak hip adduction on the more painful limb than on the less painful limb at the preferred speed (95% CI: 2.2°, 5.3°; $P<.001$; ES, 1.46), and 3.9° more peak hip adduction at the prescribed speed (95% CI: 1.8°, 5.9°; $P<.001$; ES, 1.11). Additionally, females with FAI syndrome walked with 2.8° less peak hip abduction on the more painful limb than on the less painful limb at the preferred speed (95% CI: 0.8°, 4.8°; P = .006; ES, 0.82) and 3.6° less peak hip abduction at the prescribed speed (95% CI: 1.4°, 5.8°; P = .002; ES, 0.95).

Pelvis There were significant group-by-limb interactions for peak posterior pelvic tilt (P = .012 and P = .043 for the preferred and prescribed speeds, respectively) and peak anterior pelvic tilt (P = .047 and $P<.001$) (**TABLE 5**, **FIGURE 2**). In the subsequent pairwise analyses, there were no differences between the FAI syndrome group and the comparison group ($P\geq.867$). While peak posterior pelvic tilt at the preferred speed was different between limbs (P = .044), it was less than the minimal detectable change and was therefore not interpreted.

Thigh There was a significant group-by-limb interaction for peak thigh extension position in the sagittal plane at the prescribed speed (P = .017) (**TABLE 6**, **FIGURE 3**). Within the FAI syndrome group, the thigh of the more painful limb was 1.7° less extended compared to that of the less

TABLE 4 PEAK HIP ANGLES IN THE SAGITTAL AND FRONTAL PLANES OF THE MORE PAINFUL LIMB AND LESS PAINFUL LIMB FOR THE FAI SYNDROME GROUP AND OF THE AVERAGED LEFT AND RIGHT LIMBS FOR THE COMPARISON GROUP, WALKING AT PREFERRED AND PRESCRIBED SPEEDS*

	FAI Syndrome		Comparison Group
Sex/Angle/Speed	More Painful Limb	Less Painful Limb	Average of Limbs
Males			
Hip flexion†			
Preferred‡	32.7 ± 6.5	31.8 ± 7.2	26.3 ± 9.2
Prescribed	30.0 ± 5.8	30.4 ± 6.7	25.9 ± 8.7
Hip extension			
Preferred‡	–6.8 ± 8.1	–7.7 ± 5.5	–15.4 ± 8.8
Prescribed‡	–9.0 ± 8.7	–8.9 ± 6.9	–15.8 ± 8.8
Hip adduction†			
Preferred	2.5 ± 3.6	2.1 ± 3.7	4.1 ± 2.8
Prescribed	3.2 ± 3.3	3.7 ± 3.1	4.2 ± 2.8
Hip abduction			
Preferred	–7.5 ± 2.8	–7.5 ± 3.4	–8.3 ± 1.7
Prescribed‡	–7.4 ± 1.9	–7.0 ± 3.0	–8.5 ± 1.8
Females			
Hip flexion†			
Preferred	32.1 ± 10.0	31.7 ± 8.1	32.6 ± 7.1
Prescribed	32.2 ± 9.2	31.0 ± 8.4	32.3 ± 8.2
Hip extension			
Preferred§‖	–10.4 ± 9.5	–12.2 ± 9.2	–9.9 ± 7.4
Prescribed§‖	–10.6 ± 10.4	–12.7 ± 10.5	–10.2 ± 7.7
Hip adduction†			
Preferred§‖	8.1 ± 3.3	4.4 ± 1.9	7.1 ± 2.4
Prescribed§‖	8.7 ± 4.0	4.8 ± 2.2	7.1 ± 2.3
Hip abduction			
Preferred§‖	–6.4 ± 2.0	–9.2 ± 2.6	–6.8 ± 1.9
Prescribed§‖	–6.0 ± 1.9	–9.6 ± 2.8	–6.9 ± 1.9

Abbreviation: FAI, femoroacetabular impingement.
**Values are mean ± SD degrees.*
†Hip flexion and hip adduction are positive.
‡Significant main effects for group (P<.05).
§Significant interaction effects for group by limb (P<.05).
‖Significant within–FAI syndrome group effects for limb (P<.05).

painful limb (95% CI: 0.5°, 2.9°; P = .004; ES, 0.86) at the prescribed speed. In the frontal plane, the group-by-limb interaction for peak thigh adduction position was significant at the preferred speed (P = .028) (**FIGURE 3**); no differences were noted in the subsequent pairwise analyses.

DISCUSSION

THE RESULTS OF THIS STUDY INDICATE that there are sex-specific differences in the gait alterations observed in individuals with FAI syndrome compared to individuals without hip pain. In males, there were primarily group effects, suggesting person-specific alterations; in females, there were group-by-limb interactions, suggesting limb-specific alterations. These findings indicate that FAI syndrome may contribute to gait alterations differently in males than in females.

Males with FAI syndrome had decreased peak hip extension compared to males without hip pain, a difference that was slightly larger than that noted by Hunt et al.[21] While the authors had expected a unilateral alteration, they found a group difference for males, suggesting a bilateral alteration. This may be due to the increased anterior pelvic tilt in the FAI syndrome group, a group effect for males as well. The anterior pelvic tilt could produce an offset in the hip angle curve, especially as the sagittal plane thigh angle was not different. In a secondary analysis of standing posture, males with FAI syndrome were in more anterior pelvic tilt

TABLE 5 Peak Pelvic Segment Angles in the Sagittal and Frontal Planes of the More Painful Limb and Less Painful Limb for the FAI Syndrome Group and of the Averaged Left and Right Limbs for the Comparison Group, Walking at Preferred and Prescribed Speeds*

	FAI Syndrome		Comparison Group
Sex/Angle/Speed	More Painful Limb	Less Painful Limb	Average of Limbs
Males			
Pelvic posterior tilt†			
Preferred‡	−4.8 ± 5.5	−4.8 ± 5.5	0.5 ± 7.1
Prescribed‡	−4.8 ± 5.5	−4.8 ± 5.5	0.7 ± 7.0
Pelvic anterior tilt			
Preferred‡	−9.4 ± 5.7	−9.4 ± 5.9	−4.0 ± 6.9
Prescribed‡	−9.3 ± 5.8	−9.3 ± 5.9	−4.0 ± 6.9
Pelvic hike†			
Preferred	2.6 ± 1.8	2.8 ± 1.7	3.1 ± 1.2
Prescribed§	3.1 ± 1.9	2.7 ± 1.8	3.2 ± 1.2
Pelvic drop			
Preferred	−2.8 ± 1.7	−2.6 ± 1.7	−3.1 ± 1.2
Prescribed§	−2.7 ± 1.8	−3.2 ± 1.8	−3.2 ± 1.2
Females			
Pelvic posterior tilt†			
Preferred§	−3.4 ± 7.0	−3.5 ± 7.0	−3.2 ± 6.0
Prescribed§	−3.6 ± 6.8	−3.7 ± 6.8	−3.2 ± 6.5
Pelvic anterior tilt			
Preferred§	−7.5 ± 7.1	−7.5 ± 7.1	−7.1 ± 5.6
Prescribed§	−7.6 ± 6.9	−7.6 ± 6.8	−7.2 ± 6.0
Pelvic hike†			
Preferred	3.5 ± 1.3	5.3 ± 2.5	4.1 ± 1.2
Prescribed	3.5 ± 1.6	5.7 ± 2.5	4.3 ± 1.3
Pelvic drop			
Preferred	−5.3 ± 2.5	−3.5 ± 1.3	−4.1 ± 1.2
Prescribed	−5.7 ± 2.5	−3.5 ± 1.6	−4.3 ± 1.3

Abbreviation: FAI, femoroacetabular impingement.
**Values are mean ± SD degrees.*
†Pelvic posterior tilt and pelvic hike of the contralateral side are positive.
‡Significant main effects for group (P<.05).
§Significant interaction effects for group by limb (P<.05).

than were males without hip pain, highlighting the importance of not normalizing data to a position. Increased anterior pelvic tilt (or decreased posterior tilt) has been reported in individuals with FAI syndrome during bilateral squatting[4] and stair climbing.[36] In females, however, this study did not note a difference in pelvic tilt, and the reduction in hip extension was on the more painful limb compared to the less painful limb.

The present study also found sex-specific alterations in the frontal plane. This is in partial agreement with previous studies, which noted decreased abduction of the more painful hip in individuals with FAI compared to healthy participants walking at a preferred speed.[19,22,36] However, the authors found that this was a person-specific alteration for males and a limb-specific alteration for females. Females with FAI syndrome also had increased peak hip adduction on the more painful limb compared to the less painful limb. The increased hip adduction may be due to weakness, which has been noted in the hip abductor muscles in this patient population,[7] or may be an adaptation to reduce compressive forces on the hip due to muscle activation.[30] However, reliance on the hip ligaments for stability may increase hip contact force.[12] Alternatively, the increased hip adduction, which is closer to the impingement position,[16] may contribute to symptoms and explain why females experience symptoms with less severe cam morphology than do males.[18]

A consistent pattern throughout these findings was that there were primarily

TABLE 6 PEAK THIGH SEGMENT ANGLES IN THE SAGITTAL AND FRONTAL PLANES OF THE MORE PAINFUL LIMB AND LESS PAINFUL LIMB FOR THE FAI SYNDROME GROUP AND OF THE AVERAGED LEFT AND RIGHT LIMBS FOR THE COMPARISON GROUP, WALKING AT PREFERRED AND PRESCRIBED SPEEDS*

	FAI Syndrome		Comparison Group
Sex/Angle/Speed	More Painful Limb	Less Painful Limb	Average of Limbs
Males			
Thigh flexion†			
Preferred	25.2 ± 4.9	24.1 ± 5.7	24.5 ± 3.4
Prescribed	23.1 ± 2.3	23.1 ± 3.8	24.3 ± 3.2
Thigh extension			
Preferred	–15.1 ± 4.8	–15.6 ± 4.4	–17.9 ± 3.0
Prescribed	–17.3 ± 4.5	–16.7 ± 3.8	–18.3 ± 2.9
Thigh adduction†			
Preferred	1.4 ± 2.7	1.7 ± 2.3	3.0 ± 2.0
Prescribed	2.0 ± 2.2	2.2 ± 1.8	2.9 ± 2.1
Thigh abduction			
Preferred	–6.4 ± 2.8	–6.1 ± 2.5	–6.6 ± 2.4
Prescribed	–5.5 ± 2.1	–5.7 ± 2.3	–6.8 ± 2.0
Females			
Thigh flexion†			
Preferred	26.4 ± 3.1	25.4 ± 3.5	27.4 ± 2.7
Prescribed	26.3 ± 2.6	24.7 ± 3.1	27.1 ± 3.2
Thigh extension			
Preferred	–16.4 ± 3.3	–17.7 ± 3.1	–15.7 ± 3.0
Prescribed‡§	–16.5 ± 3.9	–18.3 ± 3.7	–16.0 ± 3.1
Thigh adduction†			
Preferred‡	4.4 ± 1.9	3.3 ± 1.3	4.6 ± 1.8
Prescribed	4.5 ± 2.5	3.4 ± 1.5	4.5 ± 1.6
Thigh abduction			
Preferred	–4.1 ± 2.2	–4.8 ± 1.9	–3.3 ± 1.6
Prescribed	–3.8 ± 2.3	–4.9 ± 2.2	–3.4 ± 1.6

Abbreviation: FAI, femoroacetabular impingement.
**Values are mean ± SD degrees.*
†Flexion and adduction of the thigh segment are positive.
‡Significant interaction effects for group by limb (P<.05).
§Significant within–FAI syndrome group effects for limb (P<.05).

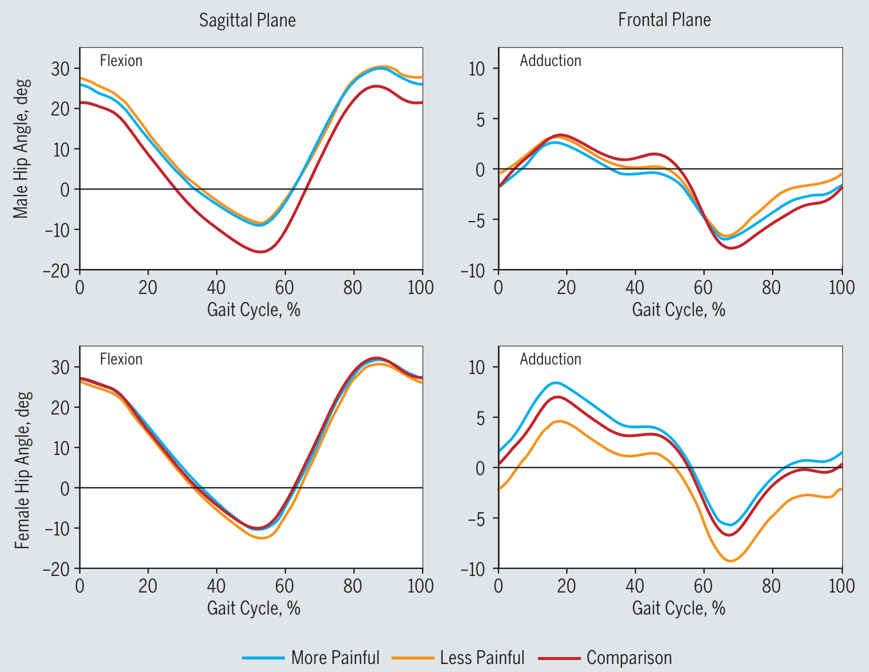


FIGURE 1. Average hip angles in the sagittal and frontal planes of the more and less painful limbs for the femoroacetabular impingement syndrome group and of the averaged left and right limbs for the comparison group, walking at the prescribed speed. Hip flexion and hip adduction are positive. Data are from heel strike to ipsilateral heel strike.

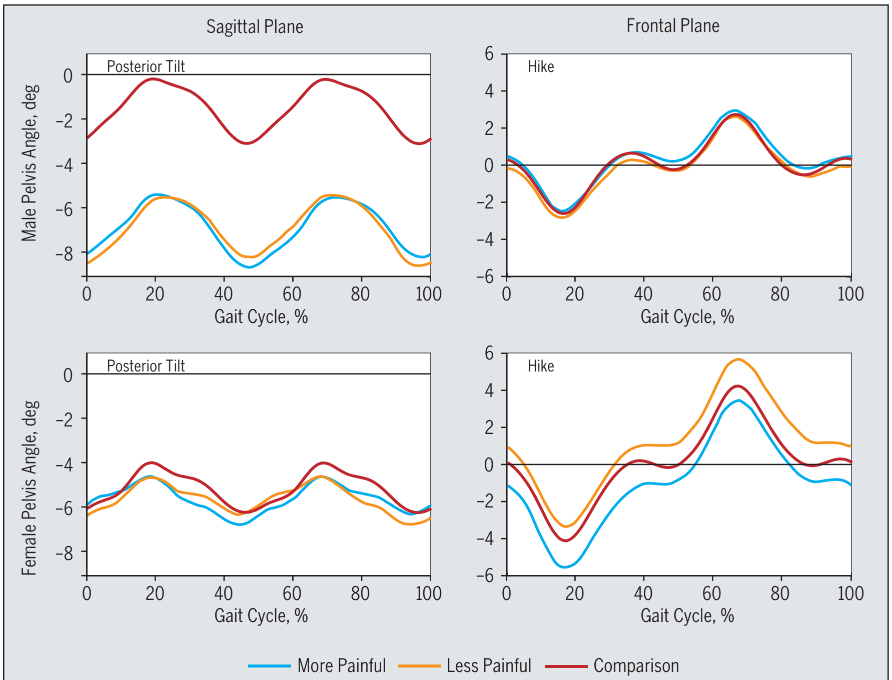


FIGURE 2. Average pelvic segment angles in the sagittal and frontal planes of the more and less painful limbs for the femoroacetabular impingement syndrome group and of the averaged left and right limbs for the comparison group, walking at the prescribed speed. Pelvic posterior tilt and pelvic hike of the contralateral side are positive. Data are from heel strike to ipsilateral heel strike.

group differences for males, but only group-by-limb interactions for females. The group effects in males could imply that males with FAI syndrome displayed the movement alterations bilaterally. However, there was significant variability as to which hip was affected more, raising questions on how to best interpret the alterations. For example, the morphology itself is unlikely to produce a reduction in hip extension. Instead, the reduction could be a result of shortened or overactive hip flexor muscles, or could be an adaptation to reduce anteriorly directed hip joint forces[30] or to reduce tension on anterior hip joint structures (eg, the iliofemoral ligament).[20] As cam morphology, and not pincer morphology,[2,37] has been linked to an increased risk for hip osteoarthritis,[1,37] it could also be an early indicator of osteoarthritis. Based on these arguments, it would follow that hip extension would be limited to a greater extent in the more painful hip than in the less painful one. While this was true in females with FAI syndrome, it was not true in males. Additionally, the analyses of group effects versus group-by-limb interactions suggest that FAI syndrome may manifest bilaterally in males and unilaterally in females. However, in this study, the authors had more females with bilateral symptoms than males with bilateral symptoms. Thus, the researchers expected group effects in females and limb effects in males.

It remains unclear how the gait alterations noted in individuals with FAI syndrome contribute to or result from the morphology or symptoms of FAI syndrome. Although anterior pelvic tilt[35] and hip adduction[16] could cause impingement, the hip does not reach the point of impingement during gait. Nonetheless, individuals with FAI syndrome report pain with prolonged walking.[10] The alterations may be a compensation for pain. The alterations were slightly larger in individuals who reported pain during gait compared to individuals with FAI syndrome who did not. It may also be that individuals with FAI syndrome display these same movement alterations in tasks that are closer to end-range motions, when impingement is more likely.

The present study does have limitations. The groups were small and the researchers did not have the power to detect small differences in movement that might be present (type II error). Also, because of the small numbers, the authors used a statistical approach that might increase the likelihood of detecting a difference when there was not one (type I error). Multicenter studies are necessary to produce larger data sets.

For the individuals with FAI syndrome, the type of bony morphology was reported by the orthopaedic clinic or participant, not measured as part of the study. The authors did not image the comparison group to evaluate hip morphology. The healthy comparison group comprised individuals without hip pain and, therefore, without FAI syndrome,[17] but may have had cam or pincer morphology, which is often present in asymptomatic individuals, especially athletes.[15]

As a cross-sectional study, it was impossible to determine the cause of the altered movement patterns. The alterations might have contributed to the development of FAI syndrome or might have been a compensation. Longitudinal studies are needed to disentangle cause and compensation. Similarly, this study did not test whether a modification of walking patterns might change the symptoms.

CONCLUSION

THE RESULTS OF THIS STUDY SUGGEST that males with FAI syndrome have different gait alterations than females with FAI syndrome when compared to sex-matched individuals without hip pain. In males, the differences were primarily between groups; in females with FAI syndrome, they were between the more painful and less painful limb. These findings suggest that altered movement may be a contributing factor to FAI syndrome and may be modifiable through neuromuscular training. ◉

KEY POINTS

FINDINGS: Gait alterations in individuals with femoroacetabular impingement (FAI) syndrome were sex specific. Males with FAI syndrome displayed a bilateral reduction in peak hip extension and an increase in peak anterior pelvic tilt compared to males without hip pain. Females with FAI syndrome displayed a reduction in peak hip extension and hip abduction and an increase in peak hip adduction on the more painful limb compared to the less painful limb.

IMPLICATIONS: These differences may indicate different etiology and the need for sex-specific movement interventions for individuals with FAI syndrome.

CAUTION: This cross-sectional study does not address the question of cause versus compensation. Future studies are warranted to determine whether these movement alterations are present in more challenging tasks and whether modifying these patterns may affect symptoms.

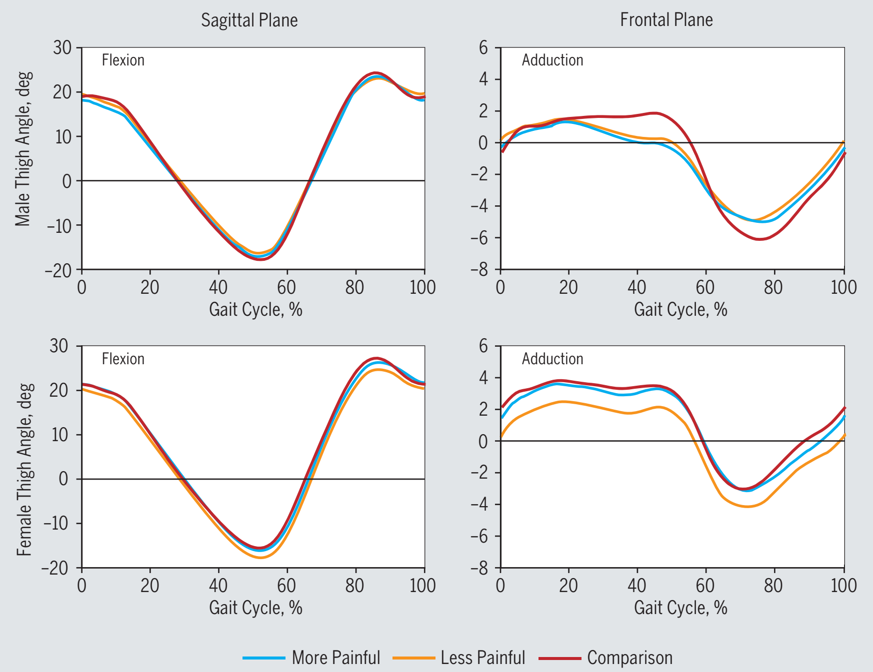


FIGURE 3. Average thigh segment angles in the sagittal and frontal planes of the more painful limb and less painful limb for the femoroacetabular impingement syndrome group and of the averaged left and right limbs for the comparison group, walking at the prescribed speed. Flexion and adduction of the thigh segment are positive. Data are from heel strike to ipsilateral heel strike.

REFERENCES

1. Agricola R, Heijboer MP, Bierma-Zeinstra SM, Verhaar JA, Weinans H, Waarsing JH. Cam impingement causes osteoarthritis of the hip: a nationwide prospective cohort study (CHECK). *Ann Rheum Dis*. 2013;72:918-923. https://doi.org/10.1136/annrheumdis-2012-201643
2. Agricola R, Heijboer MP, Roze RH, et al. Pincer deformity does not lead to osteoarthritis of the hip whereas acetabular dysplasia does: acetabular coverage and development of osteoarthritis in a nationwide prospective cohort study (CHECK). *Osteoarthritis Cartilage*. 2013;21:1514-1521. https://doi.org/10.1016/j.joca.2013.07.004
3. Amstutz HC, Thomas BJ, Jinnah R, Kim W, Grogan T, Yale C. Treatment of primary osteoarthritis of the hip. A comparison of total joint and surface replacement arthroplasty. *J Bone Joint Surg Am*. 1984;66:228-241.
4. Bagwell JJ, Snibbe J, Gerhardt M, Powers CM. Hip kinematics and kinetics in persons with and without cam femoroacetabular impingement during a deep squat task. *Clin Biomech (Bristol, Avon)*. 2016;31:87-92. https://doi.org/10.1016/j.clinbiomech.2015.09.016
5. Brisson N, Lamontagne M, Kennedy MJ, Beaulé PE. The effects of cam femoroacetabular impingement corrective surgery on lower-extremity gait biomechanics. *Gait Posture*. 2013;37:258-263. https://doi.org/10.1016/j.gaitpost.2012.07.016

6. Byrd JW, Jones KS. Prospective analysis of hip arthroscopy with 2-year follow-up. *Arthroscopy*. 2000;16:578-587. https://doi.org/10.1053/jars.2000.7683
7. Casartelli NC, Maffiuletti NA, Item-Glatthorn JF, et al. Hip muscle weakness in patients with symptomatic femoroacetabular impingement. *Osteoarthritis Cartilage*. 2011;19:816-821. https://doi.org/10.1016/j.joca.2011.04.001
8. Chung MJ, Wang MJ. The change of gait parameters during walking at different percentage of preferred walking speed for healthy adults aged 20-60 years. *Gait Posture*. 2010;31:131-135. https://doi.org/10.1016/j.gaitpost.2009.09.013
9. Clohisy JC, Baca G, Beaulé PE, et al. Descriptive epidemiology of femoroacetabular impingement: a North American cohort of patients undergoing surgery. *Am J Sports Med*. 2013;41:1348-1356. https://doi.org/10.1177/0363546513488861
10. Clohisy JC, Knaus ER, Hunt DM, Lesher JM, Harris-Hayes M, Prather H. Clinical presentation of patients with symptomatic anterior hip impingement. *Clin Orthop Relat Res*. 2009;467:638-644. https://doi.org/10.1007/s11999-008-0680-y
11. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
12. Crowninshield RD, Johnston RC, Brand RA, Pedersen DR. Pathologic ligamentous constraint of the hip. *Clin Orthop Relat Res*. 1983:291-297.
13. Diamond LE, Wrigley TV, Bennell KL, Hinman RS, O'Donnell J, Hodges PW. Hip joint biomechanics during gait in people with and without symptomatic femoroacetabular impingement. *Gait Posture*. 2016;43:198-203. https://doi.org/10.1016/j.gaitpost.2015.09.023
14. Downie WW, Leatham PA, Rhind VM, Wright V, Branco JA, Anderson JA. Studies with pain rating scales. *Ann Rheum Dis*. 1978;37:378-381. https://doi.org/10.1136/ard.37.4.378
15. Frank JM, Harris JD, Erickson BJ, et al. Prevalence of femoroacetabular impingement imaging findings in asymptomatic volunteers: a systematic review. *Arthroscopy*. 2015;31:1199-1204. https://doi.org/10.1016/j.arthro.2014.11.042
16. Ganz R, Parvizi J, Beck M, Leunig M, Nötzli H, Siebenrock KA. Femoroacetabular impingement: a cause for osteoarthritis of the hip. *Clin Orthop Relat Res*. 2003:112-120.
17. Griffin DR, Dickenson EJ, O'Donnell J, et al. The Warwick Agreement on femoroacetabular impingement syndrome (FAI syndrome): an international consensus statement. *Br J Sports Med*. 2016;50:1169-1176. https://doi.org/10.1136/bjsports-2016-096743
18. Hetsroni I, Dela Torre K, Duke G, Lyman S, Kelly BT. Sex differences of hip morphology in young adults with hip pain and labral tears. *Arthroscopy*. 2013;29:54-63. https://doi.org/10.1016/j.arthro.2012.07.008
19. Hetsroni I, Funk S, Ben-Sira D, Nyska M, Palmanovich E, Ayalon M. Femoroacetabular impingement syndrome is associated with alterations in hindfoot mechanics: a three-dimensional gait analysis study. *Clin Biomech (Bristol, Avon)*. 2015;30:1189-1193. https://doi.org/10.1016/j.clinbiomech.2015.08.005
20. Hidaka E, Aoki M, Izumi T, Suzuki D, Fujimiya M. Ligament strain on the iliofemoral, pubofemoral, and ischiofemoral ligaments in cadaver specimens: biomechanical measurement and anatomical observation. *Clin Anat*. 2014;27:1068-1075. https://doi.org/10.1002/ca.22425
21. Hunt MA, Guenther JR, Gilbart MK. Kinematic and kinetic differences during walking in patients with and without symptomatic femoroacetabular impingement. *Clin Biomech (Bristol, Avon)*. 2013;28:519-523. https://doi.org/10.1016/j.clinbiomech.2013.05.002
22. Kennedy MJ, Lamontagne M, Beaulé PE. Femoroacetabular impingement alters hip and pelvic biomechanics during gait. Walking biomechanics of FAI. *Gait Posture*. 2009;30:41-44. https://doi.org/10.1016/j.gaitpost.2009.02.008
23. Klässbo M, Larsson E, Mannevik E. Hip disability and osteoarthritis outcome score. An extension of the Western Ontario and McMaster Universities Osteoarthritis Index. *Scand J Rheumatol*. 2003;32:46-51. https://doi.org/10.1080/03009740310000409
24. Kumar D, Dillon A, Nardo L, Link TM, Majumdar S, Souza RB. Differences in the association of hip cartilage lesions and cam-type femoroacetabular impingement with movement patterns: a preliminary study. *PM R*. 2014;6:681-689. https://doi.org/10.1016/j.pmrj.2014.02.002
25. Lelas JL, Merriman GJ, Riley PO, Kerrigan DC. Predicting peak kinematic and kinetic parameters from gait speed. *Gait Posture*. 2003;17:106-112. https://doi.org/10.1016/S0966-6362(02)00060-7
26. Lewis CL, Ferris DP. Walking with increased ankle pushoff decreases hip muscle moments. *J Biomech*. 2008;41:2082-2089. https://doi.org/10.1016/j.jbiomech.2008.05.013
27. Lewis CL, Foch E, Luko MM, Loverro KL, Khuu A. Differences in lower extremity and trunk kinematics between single leg squat and step down tasks. *PLoS One*. 2015;10:e0126258. https://doi.org/10.1371/journal.pone.0126258
28. Lewis CL, Laudicina NM, Khuu A, Loverro KL. The human pelvis: variation in structure and function during gait. *Anat Rec (Hoboken)*. 2017;300:633-642. https://doi.org/10.1002/ar.23552
29. Lewis CL, Loverro KL, Khuu A. Kinematic differences during single-leg step-down between individuals with femoroacetabular impingement syndrome and individuals without hip pain. *J Orthop Sports Phys Ther*. 2018;48:270-279. https://doi.org/10.2519/jospt.2018.7794
30. Lewis CL, Sahrmann SA, Moran DW. Effect of hip angle on anterior hip joint force during gait. *Gait Posture*. 2010;32:603-607. https://doi.org/10.1016/j.gaitpost.2010.09.001
31. Ma Y, Mazumdar M, Memtsoudis SG. Beyond repeated-measures analysis of variance: advanced statistical methods for the analysis of longitudinal data in anesthesia research. *Reg Anesth Pain Med*. 2012;37:99-105. https://doi.org/10.1097/AAP.0b013e31823ebc74
32. Maslowski E, Sullivan W, Forster Harwood J, et al. The diagnostic validity of hip provocation maneuvers to detect intra-articular hip pathology. *PM R*. 2010;2:174-181. https://doi.org/10.1016/j.pmrj.2010.01.014
33. Nilsdotter AK, Lohmander LS, Klässbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS) – validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord*. 2003;4:10. https://doi.org/10.1186/1471-2474-4-10
34. Reiman MP, Goode AP, Cook CE, Hölmich P, Thorborg K. Diagnostic accuracy of clinical tests for the diagnosis of hip femoroacetabular impingement/labral tear: a systematic review with meta-analysis. *Br J Sports Med*. 2015;49:811. https://doi.org/10.1136/bjsports-2014-094302
35. Ross JR, Nepple JJ, Philippon MJ, Kelly BT, Larson CM, Bedi A. Effect of changes in pelvic tilt on range of motion to impingement and radiographic parameters of acetabular morphologic characteristics. *Am J Sports Med*. 2014;42:2402-2409. https://doi.org/10.1177/0363546514541229
36. Rylander J, Shu B, Favre J, Safran M, Andriacchi T. Functional testing provides unique insights into the pathomechanics of femoroacetabular impingement and an objective basis for evaluating treatment outcome. *J Orthop Res*. 2013;31:1461-1468. https://doi.org/10.1002/jor.22375
37. Saberi Hosnijeh F, Zuiderwijk ME, Versteeg M, et al. Cam deformity and acetabular dysplasia as risk factors for hip osteoarthritis. *Arthritis Rheumatol*. 2017;69:86-93. https://doi.org/10.1002/art.39929
38. Troelsen A, Mechlenburg I, Gelineck J, Bolvig L, Jacobsen S, Søballe K. What is the role of clinical tests and ultrasound in acetabular labral tear diagnostics? *Acta Orthop*. 2009;80:314-318. https://doi.org/10.3109/17453670902988402
39. Wilken JM, Rodriguez KM, Brawner M, Darter BJ. Reliability and minimal detectible [sic] change values for gait kinematics and kinetics in healthy adults. *Gait Posture*. 2012;35:301-307. https://doi.org/10.1016/j.gaitpost.2011.09.105

ROBERT MCCUNN, MSc[1,2] • **KAREN AUS DER FÜNTEN**, MD[1] • **MATTHEW WHALAN**, MPhty[3-5]
JOHN A. SAMPSON, PhD[3] • **TIM MEYER**, PhD, MD[1]

Soccer Injury Movement Screen (SIMS) Composite Score Is Not Associated With Injury Among Semiprofessional Soccer Players

Reducing sports injury incidence is a worthwhile endeavor for both applied practitioners and researchers alike. From a competitive point of view, lower injury burden and greater player availability have been linked to superior league ranking in professional soccer, in addition to reduced financial and psychological costs.[11,15,24] However, it should be acknowledged that the financial costs associated with injury are not limited to professional players; for example, health care system and broader economic consequences due to missed days of work may ensue following injury in recreational players. The "sequence of prevention" model proffered by van Mechelen et al[37] posits that the first and second steps to reducing injury incidence are establishing the extent of the problem (ie, incidence) and determining the etiology of injury (ie, risk factors).

BACKGROUND: The association between movement quality and injury is equivocal. No soccer-specific movement assessment has been prospectively investigated in relation to injury risk.

OBJECTIVES: To investigate the association between a soccer-specific movement-quality assessment and injury risk among semiprofessional soccer players.

METHODS: In this prospective cohort study, semiprofessional soccer players (n = 306) from 12 clubs completed the Soccer Injury Movement Screen (SIMS) during the preseason period. Individual training/match exposure and noncontact time-loss injuries were recorded prospectively for the entirety of the 2016 season. Relative risks were calculated, and presented with 90% confidence intervals, for the SIMS composite and individual subtest scores from generalized linear models with Poisson distribution offset for exposure.

RESULTS: When considering noncontact time-loss lower extremity injuries (primary level of analysis), there was a most likely trivial association with the SIMS composite score. Similarly, the SIMS composite score demonstrated most likely to likely trivial associations with all injury categories included in the secondary level of analysis (noncontact time-loss hip/groin, thigh, knee, and ankle injuries). When considering hamstring strains and ankle sprains specifically (tertiary level of analysis), the SIMS composite score demonstrated very likely trivial associations. A total of 262 noncontact time-loss injuries were recorded. The overall (training and match exposure combined) incidence of noncontact time-loss injury was 12/1000 hours.

CONCLUSION: The SIMS composite score demonstrated no association with any of the investigated categories of soccer-related injury. The SIMS composite score should not be used to group players into high- or low-risk groups.

LEVEL OF EVIDENCE: Prognosis, level 4. *J Orthop Sports Phys Ther 2018;48(8):630-636. Epub 8 May 2018. doi:10.2519/jospt.2018.8037*

KEY WORDS: *association football, epidemiology, predict, screening*

Numerous risk factors have been highlighted in relation to soccer-specific injury, including, but not limited to, previous injury, age, running load, and eccentric knee flexor strength.[1,23,34] Movement quality has recently been investigated as a potential injury risk factor within soccer; however, the evidence is equivocal.[4,30,32] While firm consensus on what constitutes movement quality is lacking, one definition offered, at least in the context of movement screening, is that it encapsulates "the maintenance of correct posture and joint alignment in addition to balance while performing the selected movements."[27] An underlying principle behind movement screening as a practice is that poor movement quality increases one's likelihood of injury.[27] Bahr[2] recently challenged the premise of screening to identify injury risk, highlighting that no such test currently displays diagnostic qualities worthy of the tag "predictive." However, while movement screening may not allow applied practitioners to predict

[1]Institute of Sports and Preventive Medicine, Saarland University, Saarbrücken, Germany. [2]Oriam: Scotland's Sports Performance Centre, Heriot-Watt University, Edinburgh, UK. [3]Centre for Human and Applied Physiology, School of Medicine, University of Wollongong, Wollongong, Australia. [4]Wollongong Wolves Football Club, Wollongong, Australia. [5]Player Welfare Department, Football South Coast, Fairy Meadow, Australia. The University of Wollongong's Human Research Ethics Committee (number HE15/340) approved this prospective cohort 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 Robert McCunn, Institute of Sports and Preventive Medicine, Campus, Geb. B8.2, Saarland University, Saarbrücken, Saarland 66123 Germany. E-mail: bob.mccunn@me.com

exactly which players will get injured, highlighting associations between risk factors and injury via prospective studies may help inform general preventive strategies.

Many movement screens exist; however, the majority have been designed for general athletic populations and not soccer players specifically.[27] To date, no soccer-specific movement screen has been prospectively investigated in relation to injury risk, despite the widespread use of movement screens within professional soccer.[25] The Soccer Injury Movement Screen (SIMS) is one such sport-specific tool and has been shown to be a reliable means to assess movement quality.[28] The SIMS comprises 5 movements, chosen to reflect the most common sites (lower extremities) and types (strains and sprains) of soccer-related injury. Hence, these subtests primarily tax the mobility and stability of the hip, knee, and ankle joints, in addition to the strength and flexibility of the surrounding musculature.[28] While other, more general movement screens have found limited associations between movement quality and injury risk,[4,17,30] it remains unclear whether such a trend also applies to sport-specific assessment tools such as the SIMS.

Therefore, the aims of the present study were 2-fold: (1) to investigate the relationship between SIMS composite score and injury risk, and (2) to investigate the relationship between the individual subtests comprising the SIMS and injury risk. The present work represents the first study to prospectively investigate any sport-specific movement screen and injury risk.

METHODS

Participants

The University of Wollongong's Human Research Ethics Committee (number HE15/340) approved this prospective cohort study. The study was conducted in accordance with the Declaration of Helsinki. In total, 306 male soccer players (mean ± SD age, 22 ± 4 years; height, 179 ± 7 cm; body mass, 75 ± 10 kg) from 2 National Premier Leagues New South Wales Division 1 clubs and 10 Illawarra Premier League clubs provided written informed consent to participate. If players were under the age of 18 years, then their legal guardians provided written informed consent and the players provided informed verbal assent. All participants were semiprofessional players who trained 2 to 3 times per week, and each club played at least 1 competitive game per week.

Procedures

Soccer Injury Movement Screen Each participant completed the SIMS exactly as described by McCunn et al[28] during the preseason period (March 2016). The SIMS has previously demonstrated good to excellent intrarater and interrater reliability.[28] The SIMS is primarily a movement-quality assessment comprising 5 subtests: anterior reach, single-leg deadlift (SLDL), in-line lunge, single-leg hop for distance (SLHD), and the tuck jump assessment. Each subtest is scored out of 10 points, resulting in a possible maximum composite score of 50 when the score from each subtest is summed. A higher score indicates poorer performance, with 0 as the best possible score and 50 as the worst. The anterior reach and SLHD scoring criteria are objective in nature and are based on reach and jump distance, respectively. Conversely, the SLDL, in-line lunge, and tuck jump rely on subjective assessment of movement quality from video footage. The exact scoring criteria and guidelines are outlined in the **APPENDIX** (available at www.jospt.org).

The lead researcher was present at every testing session and acted as the test rater, scoring all video footage. Video footage was recorded using iPad 3 devices (Apple Inc, Cupertino, CA). The rater possessed undergraduate and postgraduate sport science qualifications, was an accredited strength and conditioning coach with both the United Kingdom Strength and Conditioning Association and the National Strength and Conditioning Association, and had extensive previous experience conducting/scoring the SIMS (greater than 100 previous tests). In addition to the lead researcher, undergraduate exercise science students assisted in the collection of the SIMS test data. Prior to testing, all student helpers were required to attend 2 training sessions (4 hours in total) with the lead researcher on how to set up the testing equipment and to instruct the participants correctly (see McCunn et al[28]).

All testing was conducted either in a university biomechanics laboratory or at the training ground of the respective club when suitable facilities were available. All testing was conducted on hard, nonslip surfaces. Height, weight, and date of birth were also collected for each participant during testing sessions.

Injury Data Collection Undergraduate exercise science students with additional training (Sports Medicine Australia Level 1 Sports Trainer certification) were recruited to act as injury and exposure data collectors for the present study. In Australia, sports trainers are employed by clubs to deliver onsite first aid and acute injury management; hence, they are also well placed to record injury data.[10] In this study, the sports trainers attended every training session and match for the entirety of the 2016 season for each club.

An electronic version of the injury data recording form presented by Fuller et al[14] was used to record all physical complaints (both time loss and non–time loss). Completed electronic injury forms were sent to the lead researcher every week for review. For each recorded injury, a detailed event description was also requested from the sports trainer. The descriptions included the circumstances that immediately preceded the injury event, weather/pitch conditions, the player's own explanation of how the injury occurred, and any other information that the sports trainer considered relevant.

Each completed injury form was blinded by the lead researcher and then reviewed in conjunction with the injury description by both a chartered physical therapist and an orthopaedic doctor, separately, and assigned a diagnosis based on the Orchard Sports Injury Classification System Version 10.1.[31] If the diagnoses provided by the physical therapist and the orthopaedic doctor differed, then the lead researcher flagged the injury, and all parties reconsidered the case together until consensus on the most likely diagnosis was achieved. This method of retrospective injury diagnosis has recently been advocated for and used in previous research.[16,29]

Only noncontact injuries were included within the analyses, as contact injuries are dependent on interaction with other individuals and were judged by the authors to not be inherently related to movement quality. Sports trainers also recorded training and match exposure time (in minutes) for each individual participant and included these data in their weekly submissions to the lead researcher.

Statistical Analysis

All estimations were made using SPSS Statistics Version 24 (IBM Corporation, Armonk, NY). Data are presented as mean ± SD and absolute or relative frequencies. The effects of the SIMS composite and individual subtest scores on injury risk were analyzed using a generalized linear model with a Poisson distribution, log-linear link function, and offset for minutes of combined training and match exposure. The relative risk (RR) and 90% confidence interval were calculated to express the effect on injury risk per 1-point increase in SIMS composite or individual subtest score.

Several injury categories were analyzed using the generalized linear model. These injury categories were incorporated into 3 levels of analysis (primary, secondary, and tertiary). The primary level included 1 category: all noncontact time-loss lower extremity injuries. The secondary level included 4 separate injury categories: all noncontact time-loss hip/groin, thigh, knee, and ankle injuries—selected because they represent the most frequently injured body locations within soccer.[12] The tertiary level included 2 categories: all noncontact time-loss hamstring muscle strains and ankle sprains, which were selected because these are 2 very commonly investigated specific injury types within soccer.[13,38]

In addition, the observed frequencies of both these injury types exceeded 20 cases. According to Bahr and Holme,[3] 20 to 50 injury cases are required to detect moderate to strong associations between risk factors and injury likelihood. Bonferroni correction was applied to the *P* values for all secondary- and tertiary-level injury categories to counteract the issue of multiple comparisons. Injury rates are reported as the number of injuries per 1000 hours of training, match, and combined (both training and match) exposure.

Inferences regarding the effects of SIMS composite and individual subtest scores were assessed against a predefined smallest worthwhile effect on injury risk, using a spreadsheet to derive a confidence interval and clinical inference from a *P* value.[19] The smallest worthwhile beneficial effect was given by an RR of 0.90 (ie, a 10% lower injury rate), and, conversely, the smallest worthwhile harmful effect was given as an RR of 1.11 (ie, an 11% higher injury rate), as previously established.[18] Effects were classified as clear when there was a greater than 25% likelihood that the true effect was beneficial (reduced injury risk, $RR \leq 0.90$) or harmful (increased injury risk, $RR \geq 1.11$) and an odds ratio greater than 0.66 between benefit and harm; otherwise, the effect was deemed unclear. In instances where the RR indicated neither a beneficial nor harmful effect ($0.90 > RR < 1.11$) and the percentage likelihood of either outcome was less than 25%, the effect was classified as trivial. Effects (risk changes) were qualified against predefined probabilistic terms from the following scale: less than 0.5%, most unlikely; 0.5% to 5%, very unlikely; 5% to 25%, unlikely; 25% to 75%, possibly; 75% to 95%, likely; 95% to 99.5%, very likely; and greater than 99.5%, most likely.[5]

RESULTS

The frequencies and relative distributions of noncontact time-loss injuries, categorized by location and severity, are displayed in **TABLE 1**. A total of 262 noncontact time-loss injuries were recorded. The average ± SD exposure time experienced during training and match play per player was 55 ± 26 and 18 ± 11 hours, respectively. The overall (training and match exposure combined) incidence of noncontact time-loss injury was 12/1000 hours. The incidences of noncontact time-loss injuries sustained during training and matches were 6/1000 hours and 28/1000 hours, respectively. Injuries originating from trauma versus overuse equated to 48% (n = 125) and 52% (n = 137), respectively.

When considering all noncontact time-loss lower extremity injuries (primary level of analysis), there was a most likely trivial association with the SIMS composite score (**TABLE 2**). Similarly, the SIMS composite score demonstrated most likely to likely trivial associations with all injury categories included in the secondary level of analysis (time-loss, noncontact hip/groin, thigh, knee, and ankle injuries) (**TABLE 2**). When considering hamstring strains and ankle sprains specifically (tertiary level of analysis), the SIMS composite score demonstrated very likely trivial associations (**TABLE 2**).

The majority of SIMS individual subtest scores demonstrated trivial to unclear associations with hamstring strain and ankle sprain injuries (**TABLE 3**). However, a greater (worse) SLHD score possibly increased the risk of an ankle sprain. In contrast, a greater (worse) SLDL score possibly decreased the risk of a hamstring strain.

DISCUSSION

The SIMS composite score was not meaningfully related to any of the injury categories investigated (**TABLE 2**). Similarly, the individual subtest scores were not associated with injury, with the exceptions of the SLDL and the SLHD in relation to hamstring strains and ankle sprains, respectively (**TABLE 3**). While a greater (worse) SLHD score was possibly associated with higher risk of ankle sprain injury, it should be noted that the observed association between SLDL score and hamstring strains was counterintuitive, with a theoretically better score equating to increased risk of injury.

SIMS Composite Score

The present study suggests that the SIMS does not display an association (or any predictive relationship) with injury risk. When discussing risk factors, an important distinction should be made between *association* and *prediction*.[26] Bahr[2] explained that while an association can exist between risk factors and injury likelihood, this does not necessarily equate to predictive ability. Outcome statistics related to prediction include, although are not limited to, area under the curve, sensitivity, specificity, and positive/negative predictive value; however, no clear guidelines exist to determine the point at which these values distinguish a test as "predictive."[26] To date, no injury screening test has demonstrated satisfactory predictive ability, yet several have shown an association.[2]

The association with injury for the SIMS composite score was trivial for all categories investigated (**TABLE 2**). Despite a Bonferroni-corrected P value of less than .05 being observed with regard to hamstring strains, the clinical inference was nonetheless trivial, indicating that no meaningful relationship existed between the SIMS composite score and injury likelihood.[21] Similarly, the Functional Movement Screen developed by Cook et al[8,9] is widely used within soccer, yet its association with injury in this population is limited.[4,25,30]

The potential contributors to sports injury are numerous, and while intuitively appealing, movement quality is not strongly associated with injury risk. While it may potentially contribute to injury likelihood in combination with other risk factors, movement quality alone does not appear to be a significant risk factor. The etiology of injury is multifactorial; investigating individual risk factors in isolation, while scientifically sound, may not adequately address the real-world issues of injury prediction and prevention.[6,7]

However, the lack of association with injury risk does not necessarily render movement screening useless.[36] Other benefits of continuing the practice include establishing return-to-play test values, highlighting current musculoskeletal conditions, and establishing trust/rapport between the practitioner and the athlete.[36] Furthermore, movement screening offers a systematic way for applied practitioners to identify fundamental movement patterns relevant to safe strength training and potential performance enhancement. Some evidence suggests that movement quality may be related to physical attributes such as sprinting and jumping; ergo, the application of movement screening may relate more to performance enhancement than to injury prediction.[22,39]

Individual Subtest Scores

The associations with injury for the individual subtests mirrored the results for the composite score for the most part, with both trivial and unclear relationships observed (**TABLE 3**). Two exceptions were the SLHD and SLDL, when considering ankle sprains and hamstring strains, respectively. A higher (worse) SLHD was possibly associated with a greater risk of suffering an ankle sprain. This potential relationship between the SLHD and ankle sprain risk makes intuitive sense, because there is moderate evidence linking ankle instability and poor performance on this test.[17]

However, a higher (worse) SLDL score was possibly associated with a reduced risk of a hamstring strain. The observed relationship between SLDL score and hamstring strain injury is counterintui-

TABLE 1 NONCONTACT TIME-LOSS INJURY PATTERN BY LOCATION AND SEVERITY OF INJURIES*

	Total	1-3 d (Minimal)	4-7 d (Mild)	8-28 d (Moderate)	>28 d (Severe)
Injury location					
Head/face	0	0	0	0	0
Neck/cervical spine	3 (1)	1	0	2 (2)	0
Shoulder/clavicle	0	0	0	0	0
Sternum/ribs/upper back	1	0	1 (2)	0	0
Abdomen	2	0	0	1 (1)	1 (5)
Low back/sacrum/pelvis	11 (4)	7 (6)	2 (4)	2 (2)	0
Upper arm	0	0	0	0	0
Elbow	1	0	0	0	1 (5)
Forearm	0	0	0	0	0
Wrist	0	0	0	0	0
Hand/finger/thumb	0	0	0	0	0
Hip/groin	48 (18)	21 (19)	11 (22)	15 (18)	1 (5)
Thigh	81 (31)	30 (27)	12 (25)	32 (39)	7 (33)
Knee	41 (16)	17 (16)	10 (20)	11 (13)	3 (14)
Lower leg/Achilles tendon	23 (9)	13 (12)	5 (10)	4 (5)	1 (5)
Ankle	48 (18)	19 (17)	8 (16)	14 (17)	7 (33)
Foot/toe	3 (1)	2 (2)	0	1 (1)	0
Total injuries	262	110	49	82	21

**Values are n or n (percent). Percent values below 1% are not shown.*

tive. It is unclear why better performance on this test should potentially result in greater risk. Although not quantified directly by the SLDL test, flexibility, eccentric strength, and neuromuscular control all contribute to successful test performance. These attributes are generally believed to contribute to lower risk of injury; hence, the observed association is surprising.[33-35] However, readers should be aware that while possible associations were observed, in both instances the *P* values were greater than .05 and the 90% confidence intervals encompassed 1, indicating that very tentative conclusions should be drawn.

Methodological Considerations

A number of limitations should be considered when interpreting the results of the present study. Collecting injury data in a nonprofessional environment is fraught with challenges. The injury data-collection method might have influenced the observed injury incidence. McCunn et al[29] highlighted various challenges associated with applying the recommendations presented in the current consensus statement on soccer injury data collection within nonprofessional soccer.[14]

Using time loss to define injury severity has significant limitations when applied within an environment where players are not required to report for training/matches on a daily basis (such as in the present study).[29] In addition, the reality of conducting injury research within nonprofessional soccer dictated that access to advanced medical technology was not always possible. As a result, when deciding on the most appropriate injury diagnosis, objective indicators such as X-ray and magnetic resonance imaging scans were not always available.

Further, the results of the present study are only generalizable to semiprofessional male players, and further research may seek to investigate full professional, female, or youth populations.

A number of methodological strengths should also be acknowledged. The number of injuries observed in the present study allowed for multiple categories to be investigated while still satisfying the suggestion by Bahr and Holme[3] that a minimum of 20 to 50 cases be included for meaningful analysis. In addition, the individuals responsible for collecting the injury data and determining the diagnoses were blinded to the SIMS score of the participants, reducing the likelihood of bias.

The statistical approach used also accounted for multiple injuries to the same player and the exposure time of each individual. This is rare within research that has investigated the association with injury of other movement screening tests. Furthermore, the use of magnitude-based inferences provided an estimation of the strength of relationship between SIMS score and injury risk, rather than simply relying on null hypothesis significance (*P* values) testing.[20]

TABLE 2 ASSOCIATION BETWEEN SOCCER INJURY MOVEMENT SCREEN COMPOSITE SCORE AND INJURY RISK

	Relative Risk*	*P* Value	Bonferroni-Corrected *P* Value	Qualitative Inference
Primary analysis				
Lower extremity injuries (n = 244)	0.98 (0.96, 1.00)	.07	NA	Most likely trivial
Secondary analysis				
Hip/groin injuries (n = 48)	1.01 (0.96, 1.07)	.76	3.04	Most likely trivial
Thigh injuries (n = 81)	0.95 (0.91, 0.99)	.03	.14	Very likely trivial
Knee injuries (n = 41)	0.94 (0.89, 0.99)	.07	.26	Likely trivial
Ankle injuries (n = 48)	1.02 (0.97, 1.07)	.49	1.96	Most likely trivial
Tertiary analysis				
Hamstring muscle strains (n = 64)	0.94 (0.90, 0.98)	.01	.02†	Very likely trivial
Ankle sprains (n = 41)	1.04 (0.99, 1.09)	.21	.42	Very likely trivial

Abbreviation: NA, not applicable.
**Values in parentheses are 90% confidence interval.*
†Bonferroni-corrected P<.05.

TABLE 3 ASSOCIATION BETWEEN SOCCER INJURY MOVEMENT SCREEN SUBTEST SCORES AND HAMSTRING MUSCLE STRAIN/ANKLE SPRAIN INJURY RISK

	Relative Risk*	*P* Value	Qualitative Inference
Hamstring muscle strains (n = 64)			
Anterior reach	0.91 (0.81, 1.02)	.16	Possibly trivial
Single-leg deadlift	0.90 (0.80, 1.02)	.15	Possibly decreased risk
In-line lunge	0.93 (0.78, 1.11)	.49	Possibly trivial
Single-leg hop for distance	0.96 (0.88, 1.05)	.43	Likely trivial
Tuck jump	0.97 (0.85, 1.11)	.71	Likely trivial
Ankle sprains (n = 41)			
Anterior reach	1.06 (0.94, 1.20)	.43	Possibly trivial
Single-leg deadlift	1.10 (0.95, 1.28)	.29	Possibly trivial
In-line lunge	0.90 (0.73, 1.11)	.41	Unclear
Single-leg hop for distance	1.11 (1.00, 1.23)	.10	Possibly increased risk
Tuck jump	0.97 (0.83, 1.14)	.75	Unclear

Abbreviation: NA, not applicable.
**Values in parentheses are 90% confidence interval.*

CONCLUSION

The SIMS composite score was not associated with any of the injury categories investigated. Similarly, the individual subtest scores were not associated with injury. Therefore, the SIMS should not be used to categorize players as "high" or "low" risk. However, the SIMS may be useful in other ways. It may help practitioners identify physical qualities—for example, limb asymmetries related to strength and/or flexibility—that warrant development from a performance-enhancement perspective. ◉

KEY POINTS

FINDINGS: The Soccer Injury Movement Screen composite score was not associated with any of the injury categories investigated. Similarly, the individual subtest scores were not associated with injury.

IMPLICATIONS: The Soccer Injury Movement Screen should not be used to categorize players as "high" or "low" injury risk.

CAUTION: Using time loss to define injury severity has significant limitations when applied within an environment where players are not required to report for training/matches on a daily basis, as in the present study.

REFERENCES

1. Arnason A, Sigurdsson SB, Gudmundsson A, Holme I, Engebretsen L, Bahr R. Risk factors for injuries in football. *Am J Sports Med*. 2004;32:5S-16S. https://doi.org/10.1177/0363546503258912
2. Bahr R. Why screening tests to predict injury do not work—and probably never will...: a critical review. *Br J Sports Med*. 2016;50:776-780. https://doi.org/10.1136/bjsports-2016-096256
3. Bahr R, Holme I. Risk factors for sports injuries—a methodological approach. *Br J Sports Med*. 2003;37:384-392. https://doi.org/10.1136/bjsm.37.5.384
4. Bakken A, Targett S, Bere T, et al. The functional movement test 9+ is a poor screening test for lower extremity injuries in professional male football players: a 2-year prospective cohort study. *Br J Sports Med*. In press. https://doi.org/10.1136/bjsports-2016-097307
5. Batterham AM, Hopkins WG. Making meaningful inferences about magnitudes. *Int J Sports Physiol Perform*. 2006;1:50-57. https://doi.org/10.1123/ijspp.1.1.50
6. Bittencourt NF, Meeuwisse WH, Mendonça LD, Nettel-Aguirre A, Ocarino JM, Fonseca ST. Complex systems approach for sports injuries: moving from risk factor identification to injury pattern recognition—narrative review and new concept. *Br J Sports Med*. 2016;50:1309-1314. https://doi.org/10.1136/bjsports-2015-095850
7. Cook C. Predicting future physical injury in sports: it's a complicated dynamic system. *Br J Sports Med*. 2016;50:1356-1357. https://doi.org/10.1136/bjsports-2016-096445
8. Cook G, Burton L, Hoogenboom BJ, Voight M. Functional movement screening: the use of fundamental movements as an assessment of function - part 1. *Int J Sports Phys Ther*. 2014;9:396-409.
9. Cook G, Burton L, Hoogenboom BJ, Voight M. Functional movement screening: the use of fundamental movements as an assessment of function-part 2. *Int J Sports Phys Ther*. 2014;9:549-563.
10. Ekegren CL, Gabbe BJ, Finch CF. Injury surveillance in community sport: can we obtain valid data from sports trainers? *Scand J Med Sci Sports*. 2015;25:315-322. https://doi.org/10.1111/sms.12216
11. Ekstrand J. Keeping your top players on the pitch: the key to football medicine at a professional level. *Br J Sports Med*. 2013;47:723-724. https://doi.org/10.1136/bjsports-2013-092771
12. Ekstrand J, Hägglund M, Waldén M. Injury incidence and injury patterns in professional football: the UEFA injury study. *Br J Sports Med*. 2011;45:553-558. https://doi.org/10.1136/bjsm.2009.060582
13. Ekstrand J, Waldén M, Hägglund M. Hamstring injuries have increased by 4% annually in men's professional football, since 2001: a 13-year longitudinal analysis of the UEFA Elite Club injury study. *Br J Sports Med*. 2016;50:731-737. https://doi.org/10.1136/bjsports-2015-095359
14. Fuller CW, Ekstrand J, Junge A, et al. Consensus statement on injury definitions and data collection procedures in studies of football (soccer) injuries. *Scand J Med Sci Sports*. 2006;16:83-92. https://doi.org/10.1111/j.1600-0838.2006.00528.x
15. Hägglund M, Waldén M, Magnusson H, Kristenson K, Bengtsson H, Ekstrand J. Injuries affect team performance negatively in professional football: an 11-year follow-up of the UEFA Champions League injury study. *Br J Sports Med*. 2013;47:738-742. https://doi.org/10.1136/bjsports-2013-092215
16. Hammes D, aus der Fünten K, Bizzini M, Meyer T. Injury prediction in veteran football players using the Functional Movement Screen™. *J Sports Sci*. 2016;34:1371-1379. https://doi.org/10.1080/02640414.2016.1152390
17. Hegedus EJ, McDonough SM, Bleakley C, Baxter D, Cook CE. Clinician-friendly lower extremity physical performance tests in athletes: a systematic review of measurement properties and correlation with injury. Part 2—the tests for the hip, thigh, foot and ankle including the star excursion balance test. *Br J Sports Med*. 2015;49:649-656. https://doi.org/10.1136/bjsports-2014-094341
18. Hopkins WG. Linear models and effect magnitudes for research, clinical and practical applications. *Sportscience*. 2010;14:49-58.
19. Hopkins WG. A spreadsheet for deriving a confidence interval, mechanistic inference and clinical inference from a p value. *Sportscience*. 2007;11:16-20.
20. Hopkins WG, Batterham AM. Error rates, decisive outcomes and publication bias with several inferential methods. *Sports Med*. 2016;46:1563-1573. https://doi.org/10.1007/s40279-016-0517-x
21. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc*. 2009;41:3-13. https://doi.org/10.1249/MSS.0b013e31818cb278
22. Lloyd RS, Oliver JL, Radnor JM, Rhodes BC, Faigenbaum AD, Myer GD. Relationships between functional movement screen scores, maturation and physical performance in young soccer players. *J Sports Sci*. 2015;33:11-19. https://doi.org/10.1080/02640414.2014.918642
23. Malone S, Owen A, Mendes B, Hughes B, Collins K, Gabbett TJ. High-speed running and sprinting as an injury risk factor in soccer: can well-developed physical qualities reduce the risk? *J Sci Med Sport*. 2018;21:257-262. https://doi.org/10.1016/j.jsams.2017.05.016
24. McArdle S. Psychological rehabilitation from anterior cruciate ligament–medial collateral ligament reconstructive surgery: a case study. *Sports Health*. 2010;2:73-77. https://doi.org/10.1177/1941738109357173
25. McCall A, Carling C, Nedelec M, et al. Risk factors, testing and preventative strategies for non-contact injuries in professional football: current perceptions and practices of 44 teams from various premier leagues. *Br J Sports Med*. 2014;48:1352-1357. https://doi.org/10.1136/bjsports-2014-093439
26. McCall A, Fanchini M, Coutts AJ. Prediction: the modern-day sport-science and sports-medicine "quest for the holy grail". *Int J Sports Physiol Perform*. 2017;12:704-706. https://doi.org/10.1123/ijspp.2017-0137
27. McCunn R, aus der Fünten K, Fullagar HH, McKeown I, Meyer T. Reliability and association with injury of movement screens: a critical review. *Sports Med*. 2016;46:763-781. https://doi.org/10.1007/s40279-015-0453-1
28. McCunn R, aus der Fünten K, Govus A, Julian R, Schimpchen J, Meyer T. The intra- and inter-rater reliability of the Soccer Injury Movement Screen (SIMS). *Int J Sports Phys Ther*. 2017;12:53-66.
29. McCunn R, Sampson JA, Whalan M, Meyer T. Data collection procedures for football injuries in lower leagues: is there a need for an updated consensus statement? *Sci Med Football*.

2017;1:86-88. https://doi.org/10.1080/24733938.2016.1256554

30. Newton F, McCall A, Ryan D, et al. Functional Movement Screen (FMS™) score does not predict injury in English Premier League youth academy football players. *Sci Med Football*. 2017;1:102-106. https://doi.org/10.1080/24733938.2017.1283436

31. Orchard J, Rae K, Brooks J, et al. Revision, uptake and coding issues related to the open access Orchard Sports Injury Classification System (OSICS) versions 8, 9 and 10.1. *Open Access J Sports Med*. 2010;1:207-214. https://doi.org/10.2147/OAJSM.S7715

32. Padua DA, DiStefano LJ, Beutler AI, de la Motte SJ, DiStefano MJ, Marshall SW. The Landing Error Scoring System as a screening tool for an anterior cruciate ligament injury-prevention program in elite-youth soccer athletes. *J Athl Train*. 2015;50:589-595. https://doi.org/10.4085/1062-6050-50.1.10

33. Thorborg K, Krommes KK, Esteve E, Clausen MB, Bartels EM, Rathleff MS. Effect of specific exercise-based football injury prevention programmes on the overall injury rate in football: a systematic review and meta-analysis of the FIFA 11 and 11+ programmes. *Br J Sports Med*. 2017;51:562-571. https://doi.org/10.1136/bjsports-2016-097066

34. Timmins RG, Bourne MN, Shield AJ, Williams MD, Lorenzen C, Opar DA. Short biceps femoris fascicles and eccentric knee flexor weakness increase the risk of hamstring injury in elite football (soccer): a prospective cohort study. *Br J Sports Med*. 2016;50:1524-1535. https://doi.org/10.1136/bjsports-2015-095362

35. van der Horst N, van de Hoef S, Reurink G, Huisstede B, Backx F. Return to play after hamstring injuries: a qualitative systematic review of definitions and criteria. *Sports Med*. 2016;46:899-912. https://doi.org/10.1007/s40279-015-0468-7

36. van Dyk N, Clarsen B. Prevention forecast: cloudy with a chance of injury. *Br J Sports Med*. 2017;51:1646-1647. https://doi.org/10.1136/bjsports-2017-097861

37. van Mechelen W, Hlobil H, Kemper HC. Incidence, severity, aetiology and prevention of sports injuries. A review of concepts. *Sports Med*. 1992;14:82-99. https://doi.org/10.2165/00007256-199214020-00002

38. Waldén M, Hägglund M, Ekstrand J. Time-trends and circumstances surrounding ankle injuries in men's professional football: an 11-year follow-up of the UEFA Champions League injury study. *Br J Sports Med*. 2013;47:748-753. https://doi.org/10.1136/bjsports-2013-092223

39. Woods CT, McKeown I, Keogh J, Robertson S. The association between fundamental athletic movements and physical fitness in elite junior Australian footballers. *J Sports Sci*. 2018;36:445-450. https://doi.org/10.1080/02640414.2017.1313996

APPENDIX

SCORING CRITERIA

General Rater Instructions

Record each participant's height, weight, and tibial tuberosity height (distance from the floor to the tibial tuberosity). If a participant cannot physically perform any test due to pain, then he or she should be considered injured; this should be reported to the relevant club staff members, and the test should be postponed.

Scoring Guidelines for the Anterior Reach and Single-Leg Hop for Distance (Objective Assessments)

Anterior Reach

Measure the distance (in centimeters) from the start line to the most distal part of the foot of the reaching leg. Round to the nearest centimeter. Three repetitions are performed on each leg, and reach distance should be recorded for each attempt. The maximum reach distances achieved by each leg should be used to calculate the difference between left and right. The maximum theoretical score achievable is 10 and would represent a "poor" score. In contrast, the theoretical minimum score is 0 and would represent a "good" score.

Difference in Reach Distance Between Legs, cm	Test Score
0	0
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
≥10	10

Single-Leg Hop for Distance

Measure the distance (in centimeters) from the start line to the heel of the jumping/landing leg. Round to the nearest centimeter. Three repetitions are performed on each leg, and jump distance should be recorded for each attempt. Both jump distance and limb symmetry are considered when assigning a test score. The maximum jump distance achieved on each leg should be used to calculate the score. Combine the scores for jump distance and jump symmetry to produce the final score out of 10.

Sum of Right and Left Best Jump Distances, cm		
Male	Female	Test Score
<320	<220	5
321-340	221-240	4
341-360	241-260	3
361-380	261-280	2
381-400	281-300	1
>400	>300	0

Difference Between Best Right and Left Jumps, cm	Test Score
>20	5
17-20	4
13-16	3
9-12	2
4-8	1
<4	0

APPENDIX

Scoring Guidelines for the Single-Leg Deadlift, In-Line Lunge, and Tuck Jump (Subjective Assessments)

- If an error occurs once and the rater judges it to be egregious, then it should be scored as an error.
- If an error (but only to a minor extent) is observed once, then it should not be scored.
- If the same error (but only to a minor extent) is observed twice, then it should be scored as an error.

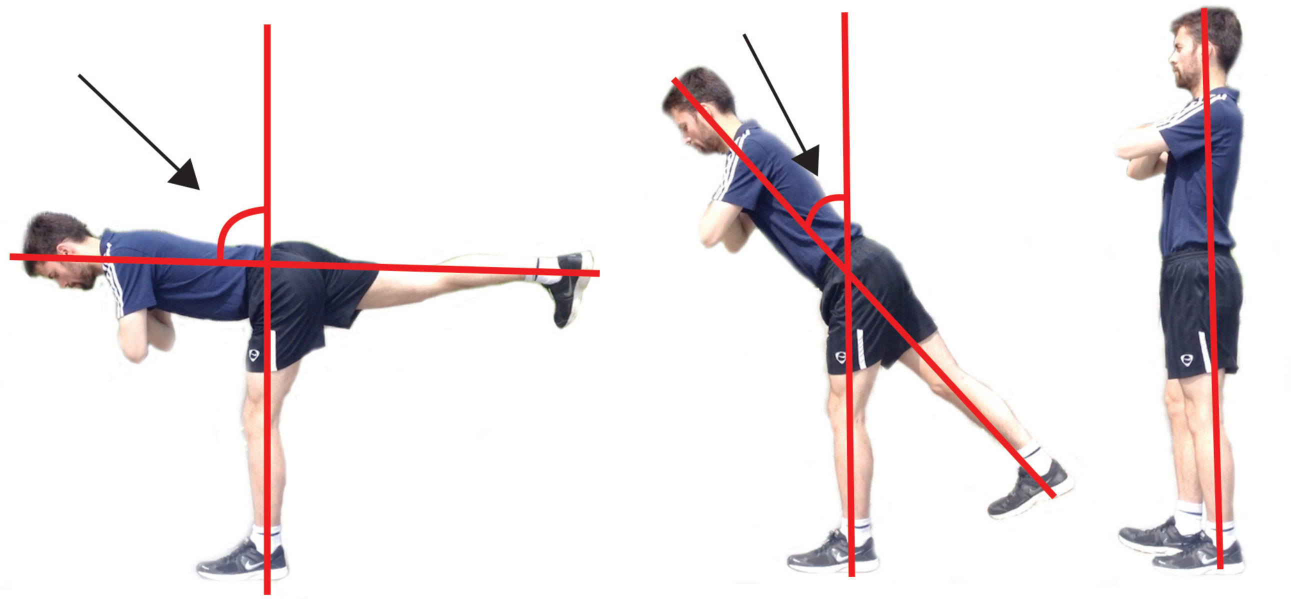
Defining specifically what constitutes "minor extent" or "egregious" is not possible. These judgments are left to the discretion of each individual rater. An important consideration is that raters are consistent in their judgments within themselves.

Single-Leg Deadlift

The score for this test is based on the "movement quality" criteria outlined below. Three repetitions are performed on each leg. The maximum theoretical score achievable is 10 and would indicate "poor" movement quality. In contrast, the theoretical minimum score is 0 and would indicate "good" movement quality. Both legs are scored, and the average of both right and left scores is assigned to the individual.

Item	Score
1. Is external hip rotation (standing leg) visible?	Yes, 1; no, 0
2. Does lumbar spine remain neutral?	Yes, 0; no, 1
3. Does thoracic spine remain neutral?	Yes, 0; no, 1
4. Does knee of raised leg remain extended throughout?	Yes, 0; no, 1
5. Is upper- and lower-body movement synchronized?	Yes, 0; no, 1
6. Is footprint maintained?	Yes, 0; no, 1
7. Is hip abduction (standing leg) present?	Yes, 1; no, 0
8. Does the standing-leg knee remain extended throughout?	Yes, 0; no, 1
9. Is the parallel-to-floor position achieved?*	Parallel (90°), 0; 89°-45°, 1; <45°, 2

The angle being assessed is displayed in the* **FIGURE *below. Scoring is relative to the stance-leg hip flexion angle.*



Sagittal view of the single-leg deadlift: the arrow indicates the angle being observed.

APPENDIX

In-Line Lunge

The score for this test is based on the "movement quality" criteria outlined below. Three repetitions are performed on each side. The maximum theoretical score achievable is 8 and would indicate "poor" movement quality. In contrast, the theoretical minimum score is 0 and would indicate "good" movement quality. Both legs are scored, and the average of both right and left scores is assigned to the individual. To generate a score out of 10, multiply the fractional score out of 8 by 10—for instance, if an individual displays 4 out of 8 possible errors, then the score out of 10 is (4/8) × 10 = 5. The reason for generating a score out of 10 is to maintain the same weighting between the 5 subtests.

Item	Score
1. Does dowel remain vertical in frontal plane throughout?	Yes, 0; no, 1
2. Does torso rotation (transverse plane) occur?	Yes, 1; no, 0
3. Does dowel remain vertical in sagittal plane throughout?	Yes, 0; no, 1
4. Does back knee touch the floor?	Yes, 0; no, 1
5. Does heel of front foot lift off the floor?	Yes, 1; no, 0
6. Is footprint maintained throughout?	Yes, 0; no, 1
7. Are the 3 dowel contact points with the body maintained?	Yes, 0; no, 1
8. Does knee valgus occur during the movement?	Yes, 1; no, 0

Tuck Jump

Mark a cross on the floor using tape (two 60-cm strips that intersect). The score for this test is based on the "movement quality" criteria outlined below. The maximum theoretical score achievable is 10 and would indicate "poor" movement quality. In contrast, the theoretical minimum score is 0 and would indicate "good" movement quality. Myer et al[1] created the tuck jump assessment, and any further clarification on scoring procedures can be sought from their original article (see reference list for full article details).

Item	Score
1. Was there knee valgus at landing?	Yes, 1; no, 0
2. Do thighs reach parallel (peak of jump)?	Yes, 0; no, 1
3. Were thighs equal side to side (during flight)?	Yes, 0; no, 1
4. Was foot placement shoulder-width apart?	Yes, 0; no, 1
5. Was foot placement parallel (front to back)?	Yes, 0; no, 1
6. Was foot contact timing equal?	Yes, 0; no, 1
7. Was there excessive contact landing noise?	Yes, 1; no, 0
8. Was there a pause between jumps?	Yes, 1; no, 0
9. Did technique decline prior to 10 seconds?	Yes, 1; no, 0
10. Were landings in same footprint (within taped cross)?	Yes, 0; no, 1

Reference

1. Myer GD, Ford KR, Hewett TE. Tuck jump assessment for reducing anterior cruciate ligament injury risk. *Athl Ther Today*. 2008;13:39-44.

DANIEL J. MESSER, BAppSc[1,2] • **MATTHEW N. BOURNE,** PhD[3] • **MORGAN D. WILLIAMS,** PhD[4]
AIMAN AL NAJJAR, MSc[5] • **ANTHONY J. SHIELD,** PhD[1,2]

Hamstring Muscle Use in Women During Hip Extension and the Nordic Hamstring Exercise: A Functional Magnetic Resonance Imaging Study

Understanding patterns of hamstring muscle activation in different exercises may have implications for strength training and hamstring and knee injury prevention programs. Hamstring muscle activation has been examined in a range of resistance training exercises using surface electromyography (sEMG)[5,15,26] and functional magnetic resonance imaging (fMRI).[5,13,15] The fMRI technique offers high levels of spatial resolution and potentially provides greater clarity as to the relative contribution of individual muscles than does sEMG, which is prone to cross-talk.[1,7] As far as we are aware, however, fMRI has not previously been employed to assess hamstring activation in women.

BACKGROUND: Understanding hamstring muscle activation patterns in resistance training exercises may have implications for the design of strength training and injury prevention programs. Unfortunately, surface electromyography studies have reported conflicting results regarding hamstring muscle activation patterns in women.

OBJECTIVES: To determine the spatial patterns of hamstring muscle activity during the 45° hip extension and Nordic hamstring exercises in women using functional magnetic resonance imaging (fMRI).

METHODS: This was a cross-sectional study in which 6 recreationally active women with no history of lower-limb injury underwent fMRI on both thighs before and immediately after 5 sets of 6 bilateral eccentric contractions of the 45° hip extension exercise or the Nordic exercise. Using fMRI, the transverse (T2) relaxation times were measured from pre-exercise and postexercise scans, and the percentage increase in T2 was used as an index of muscle activation.

RESULTS: The fMRI revealed a significantly higher biceps femoris long head-to-semitendinosus ratio during the 45° hip extension exercise than in the Nordic exercise (P = .028). The T2 increase after the 45° hip extension exercise was greater for the biceps femoris long head (P<.001), semitendinosus, and semimembranosus ($P \leq$ .001) than that for the biceps femoris short head. During the Nordic exercise, the T2 increase of the semitendinosus was greater than that of the biceps femoris short head (P<.001) and biceps femoris long head (P = .001).

CONCLUSION: While both exercises involve high levels of semitendinosus activation in women, the Nordic exercise preferentially recruits that muscle, while the hip extension exercise more evenly activates all the biarticular hamstrings.

J Orthop Sports Phys Ther 2018;48(8):607-612. Epub 23 Apr 2018. doi:10.2519/jospt.2018.7748

KEY WORDS: *functional magnetic resonance imaging (fMRI), prevention (injury), strength training*

Furthermore, there are currently disparities between sEMG studies that have employed male[5] and female[26] participants. For example, Bourne and colleagues[5] recently utilized a combination of sEMG and fMRI techniques and reported that the Nordic hamstring exercise and leg curl selectively activated[5] the semitendinosus (ST) and biceps femoris short head (BFSH) muscles, while hip extension exercises more uniformly recruited the biarticular hamstrings in men.[5] In contrast, Zebis and colleagues[26] have reported preferential sEMG activation of the biceps femoris long head (BFLH) during the Nordic hamstring exercise and various forms of leg curls in women. These discrepancies could potentially be due to small differences in EMG electrode placement and may be resolved with fMRI measures of hamstring muscle activation in women.

[1]Faculty of Health, School of Exercise and Nutrition Sciences, Queensland University of Technology, Brisbane, Australia. [2]Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Austra[illegible]h Sciences and Menzies Health Institute Queensland, Griffith University, Gold Coast, Australia. [4]Faculty of Life Sciences and Education, School of Hea[illegible]al Practice, University of South Wales, Cardiff, United Kingdom. [5]Centre for Advanced Imaging, The University of Queensland, Brisbane, Australia. The Queensland University of Technology Human Research Ethics Committee and The University of Queensland Medical Research Ethics Committee approved this study (approval number 2014001443). This study was funded by a grant from the Institute of Health and Biomedical Innovation at the Queensland University of Technology. 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 Anthony Shield, School of Exercise and Nutrition Sciences and the Institute of Health and Biomedical Innovation, Queensland University of Technology, Victoria Park Road, Kelvin Grove 4059 Brisbane, QLD, Australia. E-mail: aj.shield@qut.edu.au

Functional magnetic resonance imaging is a noninvasive form of imaging that allows for quantification of muscle activation during exercise.[12] The technique is based on changes in the T2 relaxation time of tissue water,[7,12,20] which can be inferred from signal intensity changes on fMRI. The T2 relaxation times of muscles change in proportion to exercise intensity and mirror the changes in sEMG, while also overcoming the limitations in spatial resolution of sEMG.[1,7]

The purpose of this study was to determine the spatial patterns of hamstring muscle use in women during the 45° hip extension exercise and the Nordic hamstring exercise. Based on previous work in men,[5] the authors hypothesized that the 45° hip extension exercise would display a higher BFLH/ST activation ratio than the Nordic hamstring exercise.

METHODS

Participants

Six recreationally active women (mean ± SD age, 22.5 ± 5.9 years; height, 170.5 ± 7.5 cm; weight, 59 ± 6.9 kg) participated in this study. Participants were free from injuries to the trunk, hips, and lower limbs at the time of testing and had no known history of cardiovascular, metabolic, or neurological disorders. Participants had no history of hamstring strain injury or anterior cruciate ligament (ACL) injury. Prior to testing, participants provided written informed consent to participate in the study, which was approved by the Queensland University of Technology Human Research Ethics Committee and The University of Queensland Medical Research Ethics Committee.

Study Design

A cross-sectional design was used to determine the spatial patterns of hamstring muscle use during the 45° hip extension and Nordic hamstring exercises. These exercises were chosen based on previous work, which reported that out of 10 common exercises, the 45° hip extension exercise most selectively activated the BFLH, while the Nordic hamstring exercise most selectively recruited the ST.[5] At least 7 days (±1 day) before experimental testing, all participants were familiarized with each exercise and had anthropometric measures taken. Experimental testing involved 2 separate sessions separated by at least 14 days (14 ± 4 days). Each session involved fMRI on both thighs before and immediately after one of the exercises. All testing sessions were supervised by the same investigator to ensure consistency of procedures (D.J.M.).

Exercise Protocol

An illustration of the 45° hip extension and Nordic exercises can be found in **FIGURE 1**. Participants performed only the eccentric, or lowering, phase of each exercise. Participants were instructed to perform each eccentric repetition at the slowest possible speed and were loudly encouraged to provide maximum effort during each repetition. During the familiarization session, the Nordic hamstring exercise was performed with body mass only on a device described previously, which enabled forces at the ankle to be assessed.[16,18,19] The ankle braces and load cells attached to the device allowed the forces generated by the knee flexors to be measured through the long axis of the load cells.

To approximate the intensity of the 45° hip extension and the Nordic hamstring exercises, participants performed 3 maximal eccentric repetitions (falling at the slowest possible speed) and concentric repetitions (with the assistance of an elastic band attached across the chest and held by the investigator above and behind the participant) of the Nordic hamstring exercise. The forces measured at each participant's ankles formed an eccentric-to-concentric ratio. During the Nordic hamstring exercise, participants displayed forces that were approximately 20% greater during eccentric repetitions than during concentric repetitions. Subsequently, participants were given an approximate 10-repetition maximum (10-RM) load (the heaviest load that can be lifted 10 times), in the form of weight plates held on the chest for the hip extension exercise (median, 15 kg; range, 10-20 kg). They then performed the exercise with the allocated load, and the weight was gradually increased or reduced until a 10-RM load was found. Using Holten's equation (x kg × [100%/80%], where x is the 10-RM load), the investigators were able to estimate the 1-repetition maximum (1-RM) hip extension load required to match the supramaximal intensity of the Nordic hamstring exercise (120% of the estimated hip extension 1-RM).

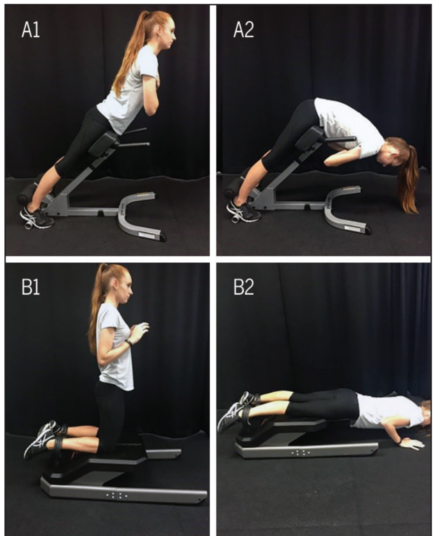


FIGURE 1. (A) 45° hip extension exercise and (B) Nordic hamstring exercise.

In each subsequent exercise session, participants performed 5 sets of 6 repetitions, with 1-minute rest intervals between sets. Participants were vocally encouraged by investigators to foster maximal effort during these tests. During the rest period, participants rested in a seated position (45° hip extension exercise) or lay prone (Nordic hamstring exercise) to minimize activation of the knee flexors. Immediately after the completion of exercise, participants were returned to the scanner for postexercise scans, which commenced within 135.4 ± 20 seconds.

Functional Magnetic Resonance Imaging

All fMRI scans were performed using a 3-T (Trio Tim; Siemens AG, Munich, Germany) imaging system with a spinal coil. Participants lay supine in the magnet bore, with their knees fully extended and hips in a neutral position and straps secured around both limbs to prevent undesired movements. A 180 × 256-mm body image matrix was positioned over the anterior thighs and aligned with the center of the 400 × 281.3-mm field of view, which included both limbs and spanned the distance between the femoral head and the tibial plateau. Consecutive T2-weighted axial images were acquired for both limbs before and immediately following exercise using a Carr-Purcell-Meiboom-Gill spin-echo pulse sequence and the following parameters: transverse relaxation time, 2540 milliseconds; echo times of 8, 16, 24, 32, 40, 48, and 56 milliseconds; number of excitations, 1; slice thickness, 10 mm; interslice gap, 10 mm.

All participants had 2 Carr-Purcell-Meiboom-Gill spin-echo pulse sequences to capture the entire length of the thigh muscles, and the total acquisition time for each sequence was 6 minutes 25 seconds. A localizer adjustment (20 seconds) was applied prior to the first sequence of each scan (pre exercise and post exercise) to standardize the field of view and to align collected images between the pre-exercise and postexercise scans.[2] A postprocessing (B1) filter was applied to minimize any inhomogeneity in magnetic resonance images caused by dielectric resonances at 3 T.[14] Participants were seated for a minimum of 15 minutes prior to pre-exercise scans, and were asked to avoid strength training of the lower limbs for 72 hours prior to data acquisition to ensure that the signal intensity profile of pre-exercise T2-weighted images was not affected by anomalous fluid shifts.[14]

Measurement of T2 Relaxation Times

The T2 relaxation times of each hamstring muscle (BFLH, BFSH, ST, and semimembranosus [SM]) were measured in T2-weighted images acquired before and after exercise sessions to evaluate muscle activation during exercise. All images were transferred to a Windows computer in the Digital Imaging and Communications in Medicine file format. The T2 relaxation time for all hamstring muscles was measured in 5 axial slices, which corresponded to 30%, 40%, 50%, 60%, and 70% of thigh length (defined as the distance between the inferior margin of the ischial tuberosity [0%] and the superior border of the tibial plateau [100%]).[2,14] Image analysis software (Sante DICOM Viewer and DICOM Editor; Santesoft Ltd, Nicosia, Cyprus) was used to measure the signal intensity of each muscle in both limbs in pre-exercise and postexercise scans. The signal intensity was measured in each slice using a 9- to 40-mm^2 circular region of interest (ROI),[12,13] which was placed in a homogeneous area of contractile tissue in the center of each muscle belly (avoiding aponeurosis, fat, tendon, bone, and blood vessels). The signal intensity represented the mean value of all pixels within the ROI and was measured across 7 echo times (8, 16, 24, 32, 40, 48, and 56 milliseconds). For each ROI, T2 relaxation time was calculated using the signal intensity value at each echo time, which was fitted to a monoexponential decay model using a least-squares algorithm: SI = M' exp(echo time/T2),[14] where SI is the signal intensity at a specific echo time and M represents the pre-exercise fMRI signal intensity. To determine the extent to which each ROI was activated during exercise, the mean percentage change in T2 was calculated as (mean postexercise T2/mean pre-exercise T2) × 100.

The percentage change in T2 relaxation time for each hamstring muscle was evaluated using the average value of all ROIs at all 5 thigh levels, which provided a measure of whole-muscle activation. Previous studies have demonstrated excellent intertester reliability of T2 relaxation time measures, with intraclass correlation coefficients ranging from 0.87 to 0.94.[7,14]

Statistical Analysis

The pre-exercise and postexercise T2 values for each exercise session were reported as mean ± SD. A repeated-measures linear mixed model fitted with the restricted maximum-likelihood method was used to compare the spatial patterns of hamstring muscle activation during the 45° hip extension exercise and Nordic hamstring exercise. For each exercise, the log-transformed percentage change in T2 relaxation time was compared between each hamstring muscle. For this analysis, muscle was the fixed factor, and participant identity and the participant identity-by-muscle interaction were the random factors. When a significant main effect was detected for muscle or exercise, post hoc t tests with Bonferroni corrections were used to identify the source and reported as mean differences with 95% confidence intervals (CIs). The adjusted alpha was set at $P<.003$ for these analyses.

The 45° hip extension exercise and the Nordic hamstring exercise differed in terms of movement velocity and hamstring excursion, so it was not appropriate to compare the magnitude of T2 shifts between exercises. To determine differences in the extent of lateral-to-medial hamstring activity between exercises, a repeated-measures linear mixed model fitted with the restricted maximum-likelihood method was used to compare the differences in the ratio of BFLH-to-ST percentage change in T2 relaxation time. For this analysis, exercise was the fixed factor and participant identity the random factor. When a main effect was found for exercise, post hoc t tests were again used to identify the source and reported as mean differences (95% CI); alpha was set at $P<.05$ for this analysis, and Cohen's d was reported as a measure of the effect size.

RESULTS

T2 Relaxation Time Percentage Change Following the 45° Hip Extension Exercise

A significant effect for muscle was noted regarding T2 changes after hip extension exercise ($P<.001$). Post hoc analyses revealed that the exercise-induced T2 increases in the BFSH were significantly less than those in the ST (mean difference in the log-transformed percentage T2 changes, 0.94; 95% CI: 0.5, 1.3; P<.001; Cohen's d = 0.8), the SM (mean difference, 0.67; 95% CI: 0.2, 1.1; P = .001; Cohen's d = 0.6), and the BFLH (mean difference, 0.73; 95% CI: 0.3, 1.1; P<.001; Cohen's d = 0.8) (**FIGURE 2A**). An example of pre-exercise and postexercise T2 images is shown in **FIGURE 3A**. No statistically significant differences were observed between any other muscle pairs (P>.003). The absolute T2 values before and after the 45° hip extension exercise and average percentage T2 increase for each muscle are displayed in the **TABLE**.

T2 Relaxation Time Percentage Change Following the Nordic Hamstring Exercise

A significant effect for muscle was noted regarding T2 changes after the Nordic hamstring exercise (P<.001). Post hoc analyses demonstrated that the exercise-induced T2 increase in the ST was significantly greater than that in the BFSH (mean difference, 0.84; 95% CI: 0.4, 1.2; P<.001; Cohen's d = 0.7) and the BFLH (mean difference, 0.59; 95% CI: 0.2, 0.9; P = .001; Cohen's d = 0.5) (**FIGURE 2B**). **FIGURE 3B** shows an example of pre-exercise and postexercise T2 images for the Nordic hamstring exercise. No statistically significant differences were observed between any other muscle pairs (P>.003). The absolute T2 values before and after the Nordic hamstring exercise and average percentage T2 increase for each muscle are displayed in the **TABLE**.

Comparison of BFLH/ST Ratio Between Exercises

A significant main effect was observed for exercise (P = .028) when comparing the BFLH/ST ratio (**FIGURE 4**). A significantly lower ratio was found during the Nordic exercise compared to the 45° hip extension exercise (mean difference, −0.20; 95% CI: −0.37, −0.03; P = .028).

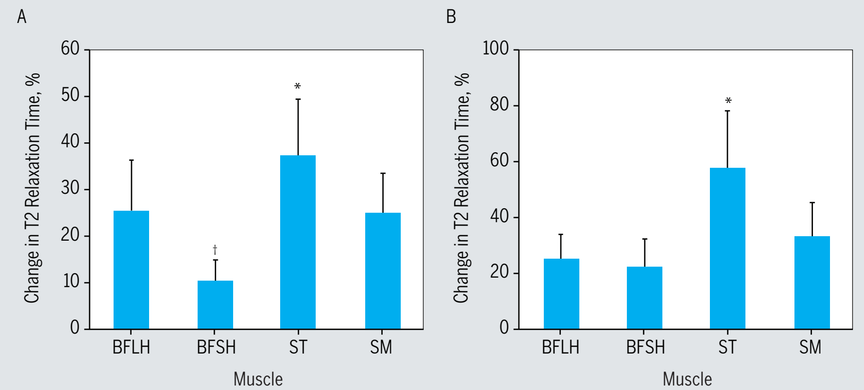


FIGURE 2. Percentage change in functional magnetic resonance imaging T2 relaxation times of each knee flexor muscle following (A) the 45° hip extension exercise and (B) the Nordic hamstring exercise. Values are displayed as the mean percentage change compared to values at rest. In (A): *Significantly different from BFSH (P<.001). †Significantly different from BFLH (P<.001) and SM (P = .001). In (B): *Significantly different from BFSH (P<.001) and BFLH (P = .001). Values are mean ± 95% confidence interval. Abbreviations: BFLH, biceps femoris long head; BFSH, biceps femoris short head; SM, semimembranosus; ST, semitendinosus.

TABLE The Average T2 Relaxation Time Before (Pre) and After (Post) the 45° Hip Extension and Nordic Hamstring Exercises, and the Average T2 Increase of Each Muscle*

	45° Hip Extension			Nordic Hamstring Exercise		
	Pre, ms	Post, ms	Increase, %†	Pre, ms	Post, ms	Increase, %†
BFLH	43.21 ± 4.33	54.99 ± 11.67	25.45 ± 16.94	42.09 ± 2.08	53.81 ± 4.38	25.39 ± 13.69
BFSH	42.35 ± 1.25	46.58 ± 4.55	10.46 ± 6.91	41.60 ± 2.02	51.18 ± 4.90	22.46 ± 15.82
ST	42.21 ± 1.86	63.14 ± 8.06	37.13 ± 19.11	41.50 ± 1.29	68.41 ± 9.77	57.99 ± 32.39
SM	42.22 ± 2.76	56.20 ± 2.88	24.90 ± 13.00	41.36 ± 1.88	56.39 ± 7.83	33.27 ± 19.24

Abbreviations: BFLH, biceps femoris long head; BFSH, biceps femoris short head; SM, semimembranosus; ST, semitendinosus.
**Values are mean ± SD.*
†Average T2 increase relative to T2 relaxation time at rest.

DISCUSSION

This is the first study to have used fMRI to explore the impact of exercise selection on hamstring muscle activation in women. The findings are consistent with a previous study of males[5] in showing high levels of ST activation during both the eccentric 45° hip extension and Nordic hamstring exercises, although the Nordic hamstring exercise preferentially targets the ST, while the 45° hip extension exercise more evenly activates the 3 biarticu-

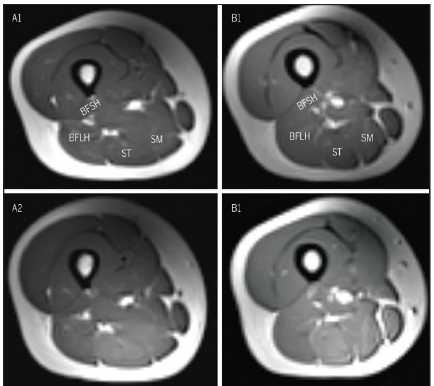


FIGURE 3. (A) 45° hip extension exercise and (B) Nordic hamstring exercise. Region of interest selection in a T2-weighted image (1) before and (2) after exercise. Abbreviations: BFLH, biceps femoris long head; BFSH, biceps femoris short head; SM, semimembranosus; ST, semitendinosus.

lar hamstrings. Given the high spatial resolution of the fMRI technique, these findings provide some clarity regarding hamstring activation patterns in 2 common hamstring exercises, which has not been produced by sometimes conflicting sEMG studies.[5,26] These findings may also have implications for design of strength training and injury prevention programs aimed at reducing hamstring strain and ACL injuries.

Explosive lower-body movements are frequently performed during competitive sport, and these activities impart significant loads on the ACL.[11,26] Given that the hamstrings represent the primary form of muscular support for this ligament,[9] strengthening these muscles is increasingly prioritized in ACL injury prevention programs.[10,24] It has previously been proposed that the ST may play a more significant role than the other hamstrings in unloading the ACL,[25] given that this muscle functions to prevent excessive anterior tibial translation and knee valgus, which are both movements commonly associated with ACL injury.[6,10,26] Accordingly, exercises that selectively activate the ST, like the Nordic hamstring exercise, may be important in ACL injury prevention protocols.

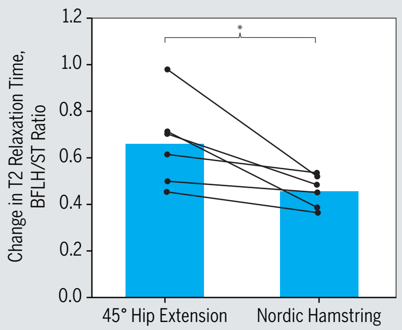


FIGURE 4. The BFLH/ST ratio percentage change in functional magnetic resonance imaging T2 relaxation times following the 45° hip extension exercise and the Nordic hamstring exercise. Column bars depict the average BFLH/ST ratio for exercise, and the lines demonstrate the participants' BFLH/ST ratio response between exercises. *Significant difference between exercises (P = .028). Abbreviations: BFLH, biceps femoris long head; ST, semitendinosus.

Prior BFLH strain injury is associated with persistent deficits in muscle activation,[4,17] BFLH fascicle lengths,[23] and muscle volume.[21] These deficits appear to persist even after a successful return to sport,[8] which suggests that conventional rehabilitation programs are ineffective in restoring optimal structure and function to this most commonly injured muscle. Given that the acute T2 patterns observed after a single exercise bout[5] closely match the hypertrophic adaptations experienced after 10 weeks of training,[3] the results of the current study suggest that the Nordic hamstring exercise is unlikely to be the optimal stimulus for restoring BFLH volume in cases of atrophy. Instead, the 45° hip extension exercise, which elicited a higher BFLH/ST activation ratio, may be a useful alternative for redressing these deficits and should be a focus of future work.

The mechanism for the nonuniformity of muscle activation in different exercises is not fully understood; however, morphological and architectural differences might be at least partly responsible.[5,15] For example, the ST displays a larger moment arm at the knee than at the hip[22] and may, therefore, be preferentially activated during movements involving knee flexion.[22] In contrast, the BFLH moment arm is greater at the hip than at the knee.[22] Moreover, the ST is long, thin, and fusiform and possesses many sarcomeres in series, which may be better suited to contractions at long muscle lengths.[14]

Participants were healthy, recreationally active women, so it cannot be assumed that similar results would occur in highly trained female athletes or those with a history of hamstring or knee pathology. Furthermore, the T2 response following exercise is influenced by a range of factors, such as the metabolic capacity of the active tissue, which is likely to differ between individuals.[1,7] The investigators attempted to minimize any variability by recruiting only female participants with a similar age and training status. Despite these attempts to approximately standardize the exercise intensity, the 45° hip extension and Nordic hamstring exercises differed in terms of movement velocity and hamstring excursion, so it was not appropriate to compare the absolute magnitudes of T2 shifts between exercises. However, these findings can offer insights into the relative metabolic activity and reliance on different hamstring muscles during the eccentric contraction of both exercises.

CONCLUSION

Female participants display different spatial patterns of hamstring muscle activation during hip- and knee-based strength exercises. The ST muscle displays high levels of muscle activation during both the eccentric 45° hip extension and Nordic hamstring exercises. Hip extension exercise more evenly activates the biarticular hamstrings, while the Nordic hamstring exercise preferentially targets the ST. Consequently, the 45° hip extension exercise displayed a BFLH/ST activation ratio that was approximately 20% higher than that of the Nordic hamstring exercise. These findings may have implications for the design of hamstring and ACL injury prevention programs. ◉

KEY POINTS

FINDINGS: While both exercises strongly activate the semitendinosus, the Nordic hamstring exercise preferentially recruits the semitendinosus muscle, while the 45° hip extension exercise activates the biarticular hamstring muscles more evenly.

IMPLICATIONS: These findings may have implications for design of strength training and injury prevention programs aimed at reducing hamstring strain and anterior cruciate ligament injuries.

CAUTION: Participants were healthy, recreationally active women, so it cannot be assumed that similar results would occur in highly trained female athletes or those with a history of hamstring or knee pathology.

ACKNOWLEDGMENTS: *The authors acknowledge the facilities and the scientific and technical assistance of the National Imaging Facility at the Centre for Advanced Imaging at The University of Queensland.*

REFERENCES

1. Adams GR, Duvoisin MR, Dudley GA. Magnetic resonance imaging and electromyography as indexes of muscle function. *J Appl Physiol (1985)*. 1992;73:1578-1583. https://doi.org/10.1152/jappl.1992.73.4.1578
2. Bourne M, Opar D, Shield A. Hamstring muscle activation during high-speed overground running: impact of previous strain injury [abstract]. *Br J Sports Med*. 2014;48:571-572. https://doi.org/10.1136/bjsports-2014-093494.32
3. Bourne MN, Duhig SJ, Timmins RG, et al. Impact of the Nordic hamstring and hip extension exercises on hamstring architecture and morphology: implications for injury prevention. *Br J Sports Med*. 2017;51:469-477. https://doi.org/10.1136/bjsports-2016-096130
4. Bourne MN, Opar DA, Williams MD, Al Najjar A, Shield AJ. Muscle activation patterns in the Nordic hamstring exercise: impact of prior strain injury. *Scand J Med Sci Sports*. 2016;26:666-674. https://doi.org/10.1111/sms.12494
5. Bourne MN, Williams MD, Opar DA, Al Najjar A, Kerr GK, Shield AJ. Impact of exercise selection on hamstring muscle activation. *Br J Sports Med*. 2017;51:1021-1028. https://doi.org/10.1136/bjsports-2015-095739
6. Buchanan TS, Kim AW, Lloyd DG. Selective muscle activation following rapid varus/valgus perturbations at the knee. *Med Sci Sports Exerc*. 1996;28:870-876.
7. Cagnie B, Elliott JM, O'Leary S, D'Hooge R, Dickx N, Danneels LA. Muscle functional MRI as an imaging tool to evaluate muscle activity. *J Orthop Sports Phys Ther*. 2011;41:896-903. https://doi.org/10.2519/jospt.2011.3586
8. Fyfe JJ, Opar DA, Williams MD, Shield AJ. The role of neuromuscular inhibition in hamstring strain injury recurrence. *J Electromyogr Kinesiol*. 2013;23:523-530. https://doi.org/10.1016/j.jelekin.2012.12.006
9. Hewett TE, Lindenfeld TN, Riccobene JV, Noyes FR. The effect of neuromuscular training on the incidence of knee injury in female athletes. *Am J Sports Med*. 1999;27:699-706. https://doi.org/10.1177/03635465990270060301
10. Hewett TE, Stroupe AL, Nance TA, Noyes FR. Plyometric training in female athletes. Decreased impact forces and increased hamstring torques. *Am J Sports Med*. 1996;24:765-773. https://doi.org/10.1177/036354659602400611
11. Johansson H, Sjölander P, Sojka P. Receptors in the knee joint ligaments and their role in the biomechanics of the joint. *Crit Rev Biomed Eng*. 1991;18:341-368.
12. Mendiguchia J, Arcos AL, Garrues MA, Myer GD, Yanci J, Idoate F. The use of MRI to evaluate posterior thigh muscle activity and damage during Nordic hamstring exercise. *J Strength Cond Res*. 2013;27:3426-3435. https://doi.org/10.1519/JSC.0b013e31828fd3e7
13. Mendiguchia J, Garrues MA, Cronin JB, et al. Nonuniform changes in MRI measurements of the thigh muscles after two hamstring strengthening exercises. *J Strength Cond Res*. 2013;27:574-581. https://doi.org/10.1519/JSC.0b013e31825c2f38
14. Ono T, Higashihara A, Fukubayashi T. Hamstring functions during hip-extension exercise assessed with electromyography and magnetic resonance imaging. *Res Sports Med*. 2011;19:42-52. https://doi.org/10.1080/15438627.2011.535769
15. Ono T, Okuwaki T, Fukubayashi T. Differences in activation patterns of knee flexor muscles during concentric and eccentric exercises. *Res Sports Med*. 2010;18:188-198. https://doi.org/10.1080/15438627.2010.490185
16. Opar DA, Piatkowski T, Williams MD, Shield AJ. A novel device using the Nordic hamstring exercise to assess eccentric knee flexor strength: a reliability and retrospective injury study. *J Orthop Sports Phys Ther*. 2013;43:636-640. https://doi.org/10.2519/jospt.2013.4837
17. Opar DA, Williams MD, Timmins RG, Dear NM, Shield AJ. Knee flexor strength and bicep femoris electromyographical activity is lower in previously strained hamstrings. *J Electromyogr Kinesiol*. 2013;23:696-703. https://doi.org/10.1016/j.jelekin.2012.11.004
18. Opar DA, Williams MD, Timmins RG, Hickey J, Duhig SJ, Shield AJ. Eccentric hamstring strength and hamstring injury risk in Australian footballers. *Med Sci Sports Exerc*. 2015;47:857-865. https://doi.org/10.1249/MSS.0000000000000465
19. Opar DA, Williams MD, Timmins RG, Hickey J, Duhig SJ, Shield AJ. The effect of previous hamstring strain injuries on the change in eccentric hamstring strength during preseason training in elite Australian footballers. *Am J Sports Med*. 2015;43:377-384. https://doi.org/10.1177/0363546514556638
20. Schuermans J, Van Tiggelen D, Danneels L, Witvrouw E. Biceps femoris and semitendinosus—teammates or competitors? New insights into hamstring injury mechanisms in male football players: a muscle functional MRI study. *Br J Sports Med*. 2014;48:1599-1606. https://doi.org/10.1136/bjsports-2014-094017
21. Silder A, Heiderscheit BC, Thelen DG, Enright T, Tuite MJ. MR observations of long-term musculotendon remodeling following a hamstring strain injury. *Skeletal Radiol*. 2008;37:1101-1109. https://doi.org/10.1007/s00256-008-0546-0
22. Thelen DG, Chumanov ES, Hoerth DM, et al. Hamstring muscle kinematics during treadmill sprinting. *Med Sci Sports Exerc*. 2005;37:108-114.
23. Timmins RG, Bourne MN, Shield AJ, Williams MD, Lorenzen C, Opar DA. Short biceps femoris fascicles and eccentric knee flexor weakness increase the risk of hamstring injury in elite football (soccer): a prospective cohort study. *Br J Sports Med*. 2016;50:1524-1535. https://doi.org/10.1136/bjsports-2015-095362
24. Wilkerson GB, Colston MA, Short NI, Neal KL, Hoewischer PE, Pixley JJ. Neuromuscular changes in female collegiate athletes resulting from a plyometric jump-training program. *J Athl Train*. 2004;39:17-23.
25. Zebis MK, Andersen LL, Bencke J, Kjaer M, Aagaard P. Identification of athletes at future risk of anterior cruciate ligament ruptures by neuromuscular screening. *Am J Sports Med*. 2009;37:1967-1973. https://doi.org/10.1177/0363546509335000
26. Zebis MK, Skotte J, Andersen CH, et al. Kettlebell swing targets semitendinosus and supine leg curl targets biceps femoris: an EMG study with rehabilitation implications. *Br J Sports Med*. 2013;47:1192-1198. https://doi.org/10.1136/bjsports-2011-090281

[VIEWPOINT]

GARY O'DONOVAN, PhD[1,2] • OLGA L. SARMIENTO, PhD[1] • MARK HAMER, PhD[2,3]

The Rise of the "Weekend Warrior"

J Orthop Sports Phys Ther 2018;48(8):604-606. doi:10.2519/jospt.2018.0611

It is recommended that individuals aged 18 to 64 years perform at least 150 minutes per week of moderate-intensity aerobic activity, at least 75 minutes per week of vigorous-intensity aerobic activity, or equivalent combinations.[29] The "weekend warrior" performs the recommended amount of aerobic activity using 1 or 2 sessions per week. The health benefits of the weekend warrior physical activity pattern were first described in 2004: Lee and colleagues[14] reported that all-cause mortality risk was 15% lower in weekend warriors than in inactive men in their study of 8421 men in the Harvard Alumni Health Study (hazard ratio = 0.85; 95% confidence interval: 0.65, 1.11). O'Donovan and colleagues[18] confirmed these benefits in one of the most talked-about studies of 2017[2]: they reported that all-cause mortality risk was 30% lower in weekend warriors than in inactive adults in their study of 63 591 men and women in the Health Survey for England and the Scottish Health Survey (hazard ratio = 0.70; 95% confidence interval: 0.60, 0.82). In this Viewpoint, we celebrate sport, exercise, and the weekend warrior. We show that the weekend warrior is thriving in the United Kingdom, the United States, and Latin America. We argue that vigorous activity and the pursuit of cardiorespiratory fitness are important to the health of the weekend warrior. Finally, we suggest that the weekend warrior physical activity pattern should be accommodated in future physical activity guidelines and interventions.

Sport, Exercise, and the Weekend Warrior

Sport is a form of physical activity that includes rules and is usually competitive. The true sense of sport is broad: "Sport means all forms of physical activity, which, through casual or organised participation, aim at expressing or improving physical fitness and mental well-being, forming social relationships or obtaining results in competition at all levels."[5] Exercise is a form of leisure-time physical activity that is planned, structured, and repetitive. Exercise training is purposeful and is performed with specific external goals, including the improvement or maintenance of physical fitness, physical performance, or health.[4] Some 1300 men and 1000 women in O'Donovan and colleagues' study[18] were weekend warriors. Ninety-four percent of the weekend warriors reported taking part in sport and exercise.[18] More than 40% of the weekend warriors were in desk-bound occupations,[18] and O'Donovan and colleagues[17] have suggested that participation in sport and exercise once or twice per week is enough to increase cardiorespiratory fitness and to reduce the all-cause mortality risk associated with today's sedentary lifestyles. The benefits of strength training were not assessed in the weekend warriors in O'Donovan and colleagues' study.[18] Stamatakis and colleagues[26] investigated the benefits of strength training in the same cohorts, and they reported that all-cause mortality risk was lower in those who reported meeting the strength training guideline[29] of at least 2 sessions per week.

Millions of adults in England enjoy running, cycling, and sports participation at least once per week.[24,25] The available evidence suggests that the weekend warrior physical activity pattern is also popular among adults with higher incomes in the United States.[23] Sport, exercise, and the weekend warrior physical activity pattern are encouraged in Latin America through Ciclovías Recreativas and other community programs. Ciclovías Recreativas are innovative, multisectorial community programs in which main roads are temporarily closed to motor vehicle traffic and opened exclusively for people to enjoy a safe, free space to exercise in a city.[22] Every Sunday morning and on public holidays, each event takes over between 1 and 114 km of road and

[1]Department of Public Health, School of Medicine, Universidad de los Andes, Bogotá, Colombia. [2]School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, UK. [3]Research Department of Epidemiology and Public Health, University College London, London, UK. 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 Gary O'Donovan, School of Medicine, Universidad de los Andes, Carrera 1, 18A-12, Bogotá, Colombia 111711. E-mail: g.odonovan@uniandes.edu.co

attracts between 40 and 1500000 participants from all walks of life.[22] Around half of the participants are on foot and half on wheels.[15] The main reasons given for taking part include health or exercise (53%), having fun (31%), sharing time with family and friends (18%), and other reasons, such as protecting the environment (5%).[22] Some 42% of adult participants report at least 3 hours of moderate to vigorous activities during a Ciclovía Recreativa, and most of them may be regarded as weekend warriors; indeed, most of them say they would not exercise if it were not for the Ciclovía Recreativa.[15]

Lessons learned from Ciclovías Recreativas in Colombia emphasize the importance of a rich policy framework, public funding, and community appropriation.[20] The National Constitution of 1991 established sport and recreation as a right for all Colombians. Funding and promoting sport and exercise were also defined as responsibilities of the state. An intersectorial government commission for physical activity was created in 2008, and in 2009 the Colombian congress passed a national obesity law that included strategies for improving environments, policies, and programs for physical activity. The national sports institute (Coldeportes) launched a national physical activity program in 2003 and expanded it to healthy lifestyles in 2011, reaching all governmental departments of Colombia. It is important to note that the sustainability and growth of the Ciclovías Recreativas are ongoing processes, and that the features relevant to their success are dynamic. Both government support and community support are necessary; neither is sufficient on its own. Community support becomes increasingly important when government priorities change and budgets fall.[6]

Activity, Fitness, and Health

It is important to reiterate that physical activity increases cardiorespiratory fitness because fitness may be a stronger predictor of mortality than smoking, blood pressure, and other established risk factors.[21] It is also important to reiterate that a relatively low amount of vigorous-intensity physical activity increases cardiorespiratory fitness more than a relatively high amount of moderate-intensity physical activity,[21] because a lack of time is regarded as a barrier to participation in physical activity.[27] Forty-five percent of the weekend warriors in O'Donovan and colleagues' study[18] reported taking part in 1 session, and 55% reported taking part in 2 sessions, of physical activity per week. In a classic series of experiments, Hickson and colleagues[9,10] and Hickson and Rosenkoetter[11] showed that cardiorespiratory fitness could be maintained with 2 sessions of vigorous-intensity exercise per week. The average body mass index was 27 kg/m^2 in the weekend warriors in O'Donovan and colleagues' study,[18] and it is noteworthy that moderate to high levels of fitness attenuate, if not negate, the association between overweight and cardiovascular disease.[19] This is the fat-but-fit paradigm that is present in one fifth of obese individuals.[19] Musculoskeletal injury risk was not assessed in the weekend warriors in O'Donovan and colleagues' study,[18] but the available evidence suggests that physical fitness is inversely associated with musculoskeletal injury risk.[12,13] Indeed, it has been suggested that athletes train hard in order to develop the physical capacities required to reduce the risk of injury.[8] The dose-response relationship to exercise training varies between individuals, and exercise should be prescribed on an individual basis.[8]

A low level of cardiorespiratory fitness of less than 5 metabolic equivalents (METs) is associated with increased all-cause mortality risk.[21] A moderate level of fitness of 5 to 7 METs is associated with a substantial reduction, and a high level of greater than 8 to 10 METs is associated with a further reduction in all-cause mortality risk.[21] We would suggest that exercise be prescribed to achieve moderate and high levels of fitness. Many exercise scientists[7] and laymen[3] struggle to understand the terms *moderate intensity* and *vigorous intensity*, and we recommend that the terms *good* and *better* be used instead. There is a compendium of physical activities[1]: "good" exercises would be 5 to 7 METs and "better" exercises would be greater than 8 to 10 METs. In this way, an exercise scientist might consult the compendium and say to someone in his or her care: "It is good that you walk at a very brisk pace." And, "It would be better if you were to run." Walking is an ideal exercise for beginners, and it is important to set achievable goals that provide success, build confidence, and increase motivation.[16] We would suggest that middle-aged and older adults take part in at least 12 weeks of walking or another "good" exercise before gradually adding running or another "better" exercise. Anyone who has experienced chest pain, dizziness, or fainting should see their physician before becoming more active.

Future Physical Activity Guidelines and Interventions

The recommended frequency is not specified in prevailing physical activity guidelines.[29] Future guidelines should be amended, because we now understand that the weekend warrior physical activity pattern is the healthy choice[18] of millions of adults around the world.[22-25] Ciclovías Recreativas inspired by the Latin American model are implemented in at least 496 cities in 27 countries on all continents.[22] The United Nations has estimated that two thirds of the world's population will be living in urban areas by 2050,[28] and reclaiming the streets on Sunday mornings and public holidays may be an ideal strategy to continue the rise of the weekend warrior.

Key Points

- The weekend warrior physical activity pattern is associated with a 30% reduction in all-cause mortality risk.
- The pursuit of cardiorespiratory fitness is likely important to the health of the weekend warrior.
- "Good" exercises of 5 to 7 METs increase cardiorespiratory fitness and decrease all-cause mortality risk.

- "Better" exercises of greater than 8 to 10 METs further increase cardiorespiratory fitness and further decrease all-cause mortality risk.
- Reclaiming the streets on Sunday mornings and public holidays may be an ideal strategy to continue the rise of the weekend warriors who run, ride, and have fun.

REFERENCES

1. Ainsworth BE, Haskell WL, Herrmann SD, et al. The Compendium of Physical Activities Tracking Guide. Available at: https://sites.google.com/site/compendiumofphysicalactivities/. Accessed December 31, 2017.
2. Altmetric. The Altmetric top 100: what academic research caught the public imagination in 2017? Available at: https://www.altmetric.com/top100/2017/#. Accessed January 2, 2018.
3. Canning KL, Brown RE, Jamnik VK, Salmon A, Ardern CI, Kuk JL. Individuals underestimate moderate and vigorous intensity physical activity. *PLoS One*. 2014;9:e97927. https://doi.org/10.1371/journal.pone.0097927
4. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep*. 1985;100:126-131.
5. Council of Europe Committee of Ministers. European Sports Charter. Strasbourg Cedex, France: Council of Europe; 1992.
6. Díaz del Castillo A, Sarmiento OL, Reis RS, Brownson RC. Translating evidence to policy: urban interventions and physical activity promotion in Bogotá, Colombia and Curitiba, Brazil. *Transl Behav Med*. 2011;1:350-360. https://doi.org/10.1007/s13142-011-0038-y
7. Ekkekakis P, Petruzzello SJ. Acute aerobic exercise and affect: current status, problems and prospects regarding dose-response. *Sports Med*. 1999;28:337-374. https://doi.org/10.2165/00007256-199928050-00005
8. Gabbett TJ. The training—injury prevention paradox: should athletes be training smarter *and* harder? *Br J Sports Med*. 2016;50:273-280. https://doi.org/10.1136/bjsports-2015-095788
9. Hickson RC, Bomze HA, Holloszy JO. Linear increase in aerobic power induced by a strenuous program of endurance exercise. *J Appl Physiol Respir Environ Exerc Physiol*. 1977;42:372-376. https://doi.org/10.1152/jappl.1977.42.3.372
10. Hickson RC, Kanakis C, Jr., Davis JR, Moore AM, Rich S. Reduced training duration effects on aerobic power, endurance, and cardiac growth. *J Appl Physiol Respir Environ Exerc Physiol*. 1982;53:225-229. https://doi.org/10.1152/jappl.1982.53.1.225
11. Hickson RC, Rosenkoetter MA. Reduced training frequencies and maintenance of increased aerobic power. *Med Sci Sports Exerc*. 1981;13:13-16.
12. Hootman JM, Macera CA, Ainsworth BE, Martin M, Addy CL, Blair SN. Association among physical activity level, cardiorespiratory fitness, and risk of musculoskeletal injury. *Am J Epidemiol*. 2001;154:251-258. https://doi.org/10.1093/aje/154.3.251
13. Knapik JJ, Spiess A, Swedler D, et al. Retrospective examination of injuries and physical fitness during Federal Bureau of Investigation new agent training. *J Occup Med Toxicol*. 2011;6:26. https://doi.org/10.1186/1745-6673-6-26
14. Lee IM, Sesso HD, Oguma Y, Paffenbarger RS, Jr. The "weekend warrior" and risk of mortality. *Am J Epidemiol*. 2004;160:636-641. https://doi.org/10.1093/aje/kwh274
15. Montes F, Sarmiento OL, Zarama R, et al. Do health benefits outweigh the costs of mass recreational programs? An economic analysis of four Ciclovía programs. *J Urban Health*. 2012;89:153-170. https://doi.org/10.1007/s11524-011-9628-8
16. O'Donovan G, Blazevich AJ, Boreham C, et al. The ABC of Physical Activity for Health: a consensus statement from the British Association of Sport and Exercise Sciences. *J Sports Sci*. 2010;28:573-591. https://doi.org/10.1080/02640411003671212
17. O'Donovan G, Hamer M, Stamatakis E. Is weekend-only physical activity enough to compensate for a sedentary lifestyle?—reply [letter]. *JAMA Intern Med*. 2017;177:1224-1225. https://doi.org/10.1001/jamainternmed.2017.2386
18. O'Donovan G, Lee IM, Hamer M, Stamatakis E. Association of "weekend warrior" and other leisure time physical activity patterns with risks for all-cause, cardiovascular disease, and cancer mortality. *JAMA Intern Med*. 2017;177:335-342. https://doi.org/10.1001/jamainternmed.2016.8014
19. Ortega FB, Lavie CJ, Blair SN. Obesity and cardiovascular disease. *Circ Res*. 2016;118:1752-1770. https://doi.org/10.1161/CIRCRESAHA.115.306883
20. Pratt M, Perez LG, Goenka S, et al. Can population levels of physical activity be increased? Global evidence and experience. *Prog Cardiovasc Dis*. 2015;57:356-367. https://doi.org/10.1016/j.pcad.2014.09.002
21. Ross R, Blair SN, Arena R, et al. Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association. *Circulation*. 2016;134:e653-e699. https://doi.org/10.1161/CIR.0000000000000461
22. Sarmiento OL, Díaz del Castillo A, Triana CA, Acevedo MJ, Gonzalez SA, Pratt M. Reclaiming the streets for people: insights from Ciclovías Recreativas in Latin America. *Prev Med*. 2017;103S:S34-S40. https://doi.org/10.1016/j.ypmed.2016.07.028
23. Shuval K, Li Q, Gabriel KP, Tchernis R. Income, physical activity, sedentary behavior, and the 'weekend warrior' among U.S. adults. *Prev Med*. 2017;103:91-97. https://doi.org/10.1016/j.ypmed.2017.07.033
24. Sport England. Active people survey. Available at: https://www.sportengland.org/research/about-our-research/active-people-survey/. Accessed December 31, 2017.
25. Sport England. Sport satisfaction survey. Available at: https://www.sportengland.org/our-work/partnering-local-government/tools-directory/sport-satisfaction-survey/. Accessed December 31, 2017.
26. Stamatakis E, Lee IM, Bennie J, et al. Does strength-promoting exercise confer unique health benefits? A pooled analysis of data on 11 population cohorts with all-cause, cancer, and cardiovascular mortality endpoints. *Am J Epidemiol*. 2018;187:1102-1112. https://doi.org/10.1093/aje/kwx345
27. Trost SG, Owen N, Bauman AE, Sallis JF, Brown W. Correlates of adults' participation in physical activity: review and update. *Med Sci Sports Exerc*. 2002;34:1996-2001.
28. United Nations. World's population increasingly urban with more than half living in urban areas. Available at: http://www.un.org/en/development/desa/news/population/world-urbanization-prospects-2014.html. Accessed January 2, 2018.
29. World Health Organization. Global Recommendations on Physical Activity for Health. Geneva, Switzerland: World Health Organization; 2010.

@ MORE INFORMATION
WWW.JOSPT.ORG

SAURAB SHARMA, PT, MPT[1,2] • **JOSHNA PALANCHOKE,** BPT[3] • **J. HAXBY ABBOTT,** DPT, PhD, FNZCP[1]

Cross-cultural Adaptation and Validation of the Nepali Translation of the Patient-Specific Functional Scale

The Patient-Specific Functional Scale (PSFS) is a patient-reported outcome measure (PROM) in which patients identify the activities that are most important to them and rate their ability to perform these activities on a numerical scale from 0 to 10, where higher scores indicate better physical function.[20] The advantages of the PSFS over other measures of physical function or disability are that (1) it is brief, easy to understand, and comprehensive, and therefore can be completed in less time; (2) it is patient generated and considers activities important at an individual level; (3) it can be administered verbally and so does not require patients to be literate; and (4) it can be applied across a variety of conditions and body regions, thus eliminating the need for multiple measures and enabling comparison of functional outcomes across conditions and between studies.[1,2,4,11,20]

BACKGROUND: The Patient-Specific Functional Scale (PSFS) is among the most commonly used measures to assess physical function.

OBJECTIVES: We aimed to translate and cross-culturally validate the PSFS to Nepali and further assess its psychometric properties.

METHODS: This longitudinal, single-arm cohort study translated and cross-culturally adapted the PSFS to Nepali (PSFS-NP) following recommended guidelines. A sample of 104 Nepalese with musculoskeletal pain was recruited to evaluate the psychometric properties of the PSFS-NP. We assessed the internal consistency (Cronbach alpha), 2-week test-retest reliability (intraclass correlation coefficient [$ICC_{3,2}$]), the smallest detectable change at the 90% confidence interval (CI), and construct validity. Concurrent validity was assessed against the Nepali versions of the Oswestry Disability Index, global rating of change, and numeric pain-rating scale. Receiver operating characteristic curves were plotted to measure responsiveness and area under the curve, and the minimum important change (MIC) was estimated.

RESULTS: The PSFS-NP showed good reliability, with a Cronbach alpha of .75, an ICC of 0.89 (95% CI: 0.78, 0.94), and a smallest detectable change at the 90% CI of 1.46. It demonstrated significant correlations with the Nepali versions of the Oswestry Disability Index (r = –0.47, P = .001), global rating of change (r = 0.71, P<.001), and numeric pain-rating scale (r = –0.32 and –0.55, P<.001). Areas under the curve ranged from 0.72 to 0.99. The MIC was 2.00 in the main analysis. Secondary analyses revealed MICs of 0.50, 0.66, and 2.00 for small, medium, and large improvement, respectively.

CONCLUSION: The PSFS-NP is a reliable, valid, and responsive measure. It can be used in clinical practice and research in Nepalese with musculoskeletal pain. *J Orthop Sports Phys Ther 2018;48(8):659-664. Epub 6 Apr 2018. doi:10.2519/jospt.2018.7925*

KEY WORDS: *clinimetrics, musculoskeletal pain, outcome measures, psychometric, responsiveness*

The validity of patient-specific scales for comparing across and between groups has been questioned; however, recent studies have shown that the PSFS is valid for use in group-level research and clinical data.[2] Additionally, a systematic review published in 2012 reported that the psychometric properties of the PSFS were adequate in various musculoskeletal conditions.[13] This scale is also more responsive than other longer measures of disability.[9,20]

Assessment of physical function is the primary focus of physical therapy interventions; however, this evaluation is hampered in Nepal because of limited availability of PROMs to assess physical function. Although the Oswestry Disability Index (ODI) is validated in Nepali,[3] administering this measure verbally can be challenging, given its length and inclusion of sensitive questions (eg, sex life). Making the PSFS available in Nepali would greatly facilitate assessment of physical function across a variety of musculoskeletal conditions in both clinical practice and research in Nepal. Accordingly, we aimed to translate and cross-culturally adapt the PSFS into Nepali (PSFS-NP), and to assess its clinimetric properties in Nepalese for the assessment of musculoskeletal pain.

[1]Centre for Musculoskeletal Outcomes Research, Dunedin School of Medicine, University of Otago, Dunedin, New Zealand. [2]Kathmandu University School of Medical Sciences, Dhulikhel, Nepal. [3]Scheer Memorial Hospital, Banepa, Nepal. This study was approved by the Institutional Review Committee of Kathmandu University School of Medical Sciences, Dhulikhel, Nepal (ethical approval number 74/15). 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 Saurab Sharma, Centre for Musculoskeletal Outcomes Research, Dunedin School of Medicine, University of Otago, Dunedin, New Zealand. E-mail: saurabsharma1@gmail.com

METHODS

A longitudinal, single-arm cohort design was used, according to a methodology described in greater detail elsewhere.[19] The study was conducted in 2 phases. Phase 1 involved the translation and cross-cultural adaptation of the PSFS-NP using recommended guidelines.[5] Refer to **APPENDIX A** (available at www.jospt.org) for the steps of the translation.

Phase 2 involved measurement of clinimetrics of the PSFS-NP. For this phase, adults experiencing musculoskeletal pain and who could count numbers from 0 to 10 and could understand and speak Nepali fluently were recruited from Dhulikhel Hospital and the community (rural and semi-urban). Participants were excluded if they had undergone any surgeries or had a recent history of trauma, a diagnosed psychiatric illness, or red flags suggestive of a tumor or infection. Ethical approval was obtained from the Institutional Review Committee of Kathmandu University School of Medical Sciences. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) recommendations guided the methodology of the study.[16]

Sociodemographic characteristics, pain history, the PSFS-NP, and Nepali versions of the numeric pain-rating scale (NPRS-NP)[19] and Oswestry Disability Index (ODI-NP)[3] were assessed at baseline. The PSFS-NP and NPRS-NP were readministered at a 2-week follow-up, along with a Nepali 7-item global rating of change (GROC-NP)[19] as an external anchor for computation of measurement error and responsiveness.[10] A GROC score of 4 was categorized as "stable," and scores between 5 and 7 were categorized as "improved" (5, slight improvement; 6, medium improvement; 7, large improvement). All measures were administered verbally to allow inclusion of participants with poor or no literacy. The details of the measures used are presented in **TABLE 1**. Data were analyzed using SPSS Version 24 (IBM Corporation, Armonk, NY). The level of significance was set at $P<.05$.

Reliability

Internal consistency was reported using the Cronbach alpha, with a score of .90 or greater indicating excellent internal consistency.[7] Two-week test-retest reliability was computed for the stable group using a 2-way mixed-effects model (with absolute agreement) and intraclass correlation coefficient model 3,2 ($ICC_{3,2}$). An ICC value higher than 0.75 indicates excellent test-retest reliability.[7] We used a Bland-Altman plot to report limits of agreement.[6] Standard error of measurement (SEM) was calculated as SD change × $(1 - ICC)^{1/2}$, where SD change equals SD (baseline – final).[10] We computed individual-level smallest detectable change (SDC) at the 90% confidence interval (CI) as $z \times \sqrt{2} \times$ SEM ($z = 1.64$ at the 90% CI). We hypothesized that the PSFS-NP would demonstrate excellent internal consistency and test-retest reliability, and have an SDC_{90} between 1 and 2.5, as previously reported.[13]

Validity

Construct validity of the PSFS-NP was examined by testing the hypotheses that (1) PSFS change score (PSFS baseline score – PSFS final score) would change significantly within the improved group using a 1-sample *t* test, and (2) PSFS change scores would differ significantly between the stable and improved groups using an independent-samples *t* test.[12]

Concurrent validity was evaluated by comparing PSFS baseline scores with the ODI baseline scores for the subgroup with low back pain (LBP), and with the NPRS baseline scores for the total sample. We hypothesized a moderate significant

TABLE 1 Nepali Versions of PROMs Used in the Study

PROM	Items	Scale	Construct Assessed	Scoring	Psychometrics
PSFS	3	0-10, ordinal	Physical function	Mean of item scores (range, 0-10). Lower scores indicate greater disability	...
NPRS[19]	3	0-10, ordinal	Pain intensity	Mean of 3 item scores (current, best, and worst in past 24 h) (range, 0-10). Higher scores indicate greater pain intensity 0 is no pain and 10 is maximum pain	ICC = 0.81; SDC_{90}, 1.13; MIC, 1.17; concurrent validity (with GROC), $r = 0.45$
ODI[3]	10	0-5, ordinal	Physical function and pain	Sum of item scores/number of items rated × 100 (range, 0-100). Higher scores indicate greater disability	Cronbach α = .72; ICC = 0.87
GROC[19]	1	1-7, ordinal	Change in global status of the patient's musculoskeletal condition (transitional scale)	Single-item score (4 is no change). Scores higher than 4 mean greater improvement and scores lower than 4 mean greater worsening in health status 7 is a lot of improvement, 6 is medium improvement, 5 is slight improvement, 4 is no change, 3 is slightly worse, 2 is moderately worse, 1 is a lot worse	MIC, 1-point change[14,19]

Abbreviations: GROC, global rating of change; ICC, intraclass correlation coefficient; MIC, minimum important change; NPRS, numeric pain-rating scale; ODI, Oswestry Disability Index; PROM, patient-reported outcome measure; PSFS, Patient-Specific Functional Scale; SDC, smallest detectable change.

negative correlation. We also correlated PSFS change scores with GROC-NP and NPRS change (NPRS baseline – NPRS final) scores for the total sample. We hypothesized that PSFS change would correlate strongly (significantly and positively) with the GROC-NP, but moderately (significantly and negatively) with the NPRS change scores. We considered Pearson correlation coefficients of 0.30 to 0.70 as a moderate correlation, and greater than 0.70 as a strong correlation.

Responsiveness

Receiver operating characteristic (ROC) curves were plotted to assess the responsiveness of the PSFS-NP, using the GROC-NP as an external anchor.[10] The ROC curves were plotted for the PSFS change scores for the stable group compared with the improved group. Secondary analyses assessed (1) stable group versus small improvement group, (2) stable group versus medium improvement group, and (3) stable group versus large improvement group. Area under the curve (AUC) was calculated to indicate the capacity of the PSFS-NP to differentiate between the stable and improved groups. Values of AUC closer to 1 indicate better agreement with the GROC.[10] Minimum important change (MIC) values were also calculated.[10] We hypothesized that MIC values would range from 1 to 4, as typically reported in a previous systematic review.[13]

RESULTS

In phase 1, the translation of the PSFS to Nepali was successfully completed. The summary of the translation history is reported in **APPENDIX A**. The PSFS-NP can be found in **APPENDIX B** (available at www.jospt.org).

In phase 2, 104 adults with musculoskeletal pain (75 hospital, 29 community) consented to participate in the study. All participants completed both the baseline and final assessments at a mean ± SD interval of 11.5 ± 3.5 days (range, 6-18 days). The participants' characteristics are described in **TABLE 2**. Thirty-six participants (35%) with complete follow-up data were classified as stable, 64 (62%) as improved, and 4 (4%) as "worsened" based on the GROC scores per a priori definition. Forty-five of 48 participants (94%) in the LBP subgroup completed the ODI-NP.

The PSFS-NP demonstrated acceptable internal consistency of .75 and excellent test-retest reliability of 0.89 (95% CI: 0.78, 0.94). The SEM and individual-level SDC_{90} were 0.63 and 1.46, respectively. The Bland-Altman plot is shown in the **FIGURE**.

The PSFS-NP demonstrated construct validity by t tests: t_{63} = 8.65 (P<.001) within the improved group and t_{98} = 5.21 (P<.001) between the stable and improved groups. Concurrent validity was supported by moderate correlations of PSFS baseline score with ODI baseline score (r = −0.47, P = .001) and NPRS baseline score

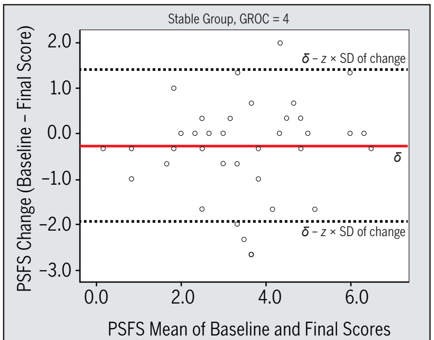


FIGURE. Bland-Altman plot for the PSFS-NP. The *y*-axis is the change in PSFS-NP scores between baseline and follow-up measurements, and the *x*-axis is the mean of PSFS-NP scores at baseline and at final measurement. The solid line is the mean change in score (*δ*), and dotted lines are *δ* ± *z* × SD change, where *z* = 1.64 at the 90% confidence interval. Abbreviation: PSFS-NP, Nepali version of the Patient-Specific Functional Scale; GROC, Global Rating of Change.

TABLE 2 DESCRIPTION OF THE PARTICIPANTS

Variable	Value
Age, y*	41.2 ± 13.5
Sex, n (%)	
Male	32 (31)
Female	72 (69)
Total	104 (100)
Ethnicity, n (%)	
Newar	34 (33)
Brahmin	23 (22)
Chettri	16 (15)
Other	31 (30)
Education, n (%)	
No school	41 (39)
Primary (grades 1-5)	11 (11)
Secondary (grades 6-10)	17 (16)
Higher secondary (grades 11-12)	16 (15)
Bachelor and above	19 (18)
Occupation, n (%)	
Agriculture and housework	28 (27)
Household work only	22 (21)
Agriculture only	8 (8)
Sitting job (office/business)	8 (8)
No work	6 (6)
Other	32 (31)

Table continues on page 662.

TABLE 2 Description of the Participants (continued)

Variable	Value
Site of pain, n (%)	
Low back	48 (46)
Knee	21 (20)
Shoulder	13 (13)
Neck	9 (9)
Elbow	5 (5)
Other	8 (8)
Total duration of pain, mo*	21.70 ± 34.00
Time between evaluations, d*	11.50 ± 3.50
GROC at follow-up, n (%)	
Worsened group (<4)	4 (4)
No improvement (4)	36 (35)
Improved group (5-7)	64 (62)
Small improvement (5)	30 (29)
Medium improvement (6)	23 (22)
Large improvement (7)	11 (11)
Average PSFS score (0-10)*	
Baseline	3.70 ± 1.73
Final	5.03 ± 2.27
Change (baseline – final)	–1.32 ± 1.89
Average NPRS score (0-10)*	
Baseline	4.27 ± 1.63
Follow-up	3.36 ± 1.56
Change (baseline – follow-up)	0.90 ± 1.49

Abbreviations: GROC, global rating of change; NPRS, numeric pain-rating scale; PSFS, Patient-Specific Functional Scale.
**Values are mean ± SD.*

TABLE 3 Responsiveness of the Nepali Patient-Specific Functional Scale

	AUC*	MIC
Primary analysis (GROC 4 versus GROC 5-7)	0.83 (0.74, 0.91)	2.00
Small improvement (GROC 4 versus GROC 5)	0.72 (0.59, 0.84)	0.50
Medium improvement (GROC 4 versus GROC 6)	0.89 (0.80, 0.98)	0.66
Large improvement (GROC 4 versus GROC 7)	0.99 (0.97, 1.00)	2.00

Abbreviations: AUC, area under the curve; GROC, global rating of change; MIC, minimum important change.
**Values in parentheses are 95% confidence interval.*

(r = −0.32, P = .001), a moderate correlation of PSFS change score with NPRS change score (r = −0.55, P<.001), and a significant, strong positive correlation of PSFS change score with the GROC-NP (r = 0.71, P<.001). Four ROC curves for the PSFS change scores were plotted (see **APPENDIX C**, available at www.jospt.org) for the 4 groups based on GROC scores, as described in the Methods. The AUCs with their CIs and the respective MICs are reported in **TABLE 3**.

DISCUSSION

The PSFS-NP, after translation in accordance with recommended guidelines, demonstrated acceptable clinimetric properties, as hypothesized. Although the PSFS has been validated in many languages in a variety of clinical conditions, this study supports its validation in individuals with low literacy (50% of this study's participants had only primary education or less) when administered verbally.

Test-retest reliability of the PSFS-NP was excellent, in line with our a priori hypothesis. The 95% CI of the ICC of the PSFS-NP (0.78, 0.94) is consistent with 6 of 8 studies included in a previous systematic review reporting clinimetric properties of the PSFS in musculoskeletal conditions (ranging between 0.76 and 0.97).[13] Only 1 study reported a lower ICC (0.76, for chronic lateral epicondylalgia), and 1 reported a higher ICC (0.97, for LBP). Similarly, the Japanese PSFS reported almost perfect 1-week reliability (ICC = 0.98).[17] Such high reliability could be because participants were informed of the baseline scores, which may have increased the reliability. Likewise, the SDC_{90} of the PSFS-NP (1.46) was also within the hypothesized range (1.0-2.5)[13] and equal to that reported for chronic LBP.[15]

Similarly, the PSFS-NP also demonstrated validity as hypothesized. The construct validity was established by a statistically significant mean difference within the improved group, and between the stable and improved groups, as in a previous study.[12]

Concurrent validity was also confirmed, based on the a priori hypotheses of moderate to strong correlations with the criterion variables. First, the PSFS baseline scores demonstrated moderate correlation (r = −0.47) with the ODI baseline scores in this study, which is a lower correlation than those previously reported (r = 0.51-0.74)[20] with the Roland-Morris Disability questionnaire, a measure of back-related disability similar to the ODI. The strength of correlation of the PSFS and ODI was only moderate, which may be because of the verbal administrations of the ODI, which likely affected responses to the item related to sex life. Culturally, Nepalese patients prefer to say that "sex life is absent" rather than "sex life is normal" when interviewed, which is evident by the lowest scores for this item.

Second, correlations of PSFS change and GROC scores were strong, as hypothesized, because both assessed change (ie, physical function and overall change, respectively). As physical function is a prime concern of patients, their overall reporting of change (assessed by the GROC) could be highly influenced by change in physical function (assessed by the PSFS).[8,17]

Finally, as hypothesized, the correlation of the PSFS-NP with the NPRS-NP was moderate. It is worth noting that neither the GROC nor NPRS directly assesses the construct of physical function; the findings relating to validity would have benefited from use of scales that assess the construct of physical function specifically. However, due to few available valid measures in Nepali, we were limited in the present study to investigating this only in people with LBP, using the ODI-NP, which supported concurrent validity. Nevertheless, we can confirm the construct validity of the PSFS-NP, because, as proposed by Terwee and colleagues,[21] more than 75% of the a priori hypotheses were achieved.

The MIC value (2.00) of the PSFS-NP in the current study lies within the range reported previously, as hypothesized.[13] The MIC values obtained in this study are consistent with those reported previously for chronic LBP[8,15] using the same method of assessment, by ROC curve. The stepwise increase of MICs for small, medium, and large change for the PSFS-NP (0.50, 0.66, and 2.00, respectively) supports its construct validity. This method of estimating the MIC for small, medium, and large change separately provides a conservative estimate of MIC; that is, calculation of the MIC using cut points for medium (or lesser) change or large (or lesser) change would result in lower estimates for MIC than this discrete-groups method.

Although the current study is robust in terms of its methodology and complies with COSMIN recommendations,[16] the results should be interpreted with consideration of its limitations. First, the findings related to reliability of the PSFS-NP are based on a relatively small number of individuals in the stable group (n = 36). A larger sample size may provide greater certainty for reliability coefficients.

Second, because the findings on responsiveness are based on a relatively short duration of follow-up (6-18 days, which is shorter than many studies), the magnitude of change may be smaller than that observed in other studies. This disadvantage of a shorter follow-up period is offset by the current study's shorter recall time for the GROC, which likely reduced recall bias.[18]

Finally, the findings of this study are limited to individuals with musculoskeletal pain and so may not be generalized to other health conditions, such as cardiopulmonary or neurological conditions. Future research may consider the usefulness of the PSFS in other health conditions.[13]

The findings of this research have important clinical and research implications. As the assessment of physical function is recommended in core outcome sets,[7,10] the availability of a validated PSFS-NP will facilitate its use in the assessment of physical function in musculoskeletal conditions in Nepal, in both research and clinical practice.

CONCLUSION

The Nepali version of the PSFS is a reliable, valid, and responsive measure for assessment of physical function in individuals with musculoskeletal pain. Clinicians should consider a change of score lower than 1.5 on the 0-to-10 PSFS-NP as measurement error, and a score change of 2 points as a meaningful change in function for people with musculoskeletal pain.

KEY POINTS

FINDINGS: The Nepali translation of the Patient-Specific Functional Scale (PSFS) is a reliable, responsive, and valid measure for assessing physical function of Nepalese adults with musculoskeletal pain.

IMPLICATIONS: Clinicians and researchers should consider a change of 2 (out of 10) of the average score of 3 items as a clinically meaningful change for patients with musculoskeletal pain. A score of 1.46 (out of 10) is the smallest detectable change, and any change score less than this should be considered a measurement error.

CAUTION: The validity of the PSFS was established in adult Nepalese with musculoskeletal pain with sufficient numeracy to understand a numerical scale. The measure should not be considered valid or reliable in individuals with a lower level of numerical skill, or in other patient populations, such as patients with neurological, cardiopulmonary, or pediatric conditions.

ACKNOWLEDGMENTS: *The authors would like to thank the translators who volunteered to contribute to the translation of the PSFS to Nepali; the developer of the PSFS, Dr Paul W. Stratford, for reviewing the back translation of the final Nepali version of this outcome measure; and the participants who volunteered for this study.*

REFERENCES

1. Abbott JH, Schmitt J. Minimum important differences for the Patient-Specific Functional Scale, 4 region-specific outcome measures, and the numeric pain rating scale. *J Orthop Sports Phys Ther*. 2014;44:560-564. https://doi.org/10.2519/jospt.2014.5248
2. Abbott JH, Schmitt JS. The Patient-Specific Functional Scale was valid for group-level change comparisons and between-group discrimination. *J Clin Epidemiol*. 2014;67:681-688. https://doi.org/10.1016/j.jclinepi.2013.11.002
3. Acharya RS, Al-Oraibi S, Adhikari SP, Parajuli N, Limbu H, Al Enezi F. Validation in the cross-

cultural adaptation of the Nepali version of the Oswestry Disability Index. *Indian J Physiother Occup Ther*. 2014;8:158-163. https://doi.org/10.5958/j.0973-5674.8.2.079
4. Barten JA, Pisters MF, Huisman PA, Takken T, Veenhof C. Measurement properties of patient-specific instruments measuring physical function. *J Clin Epidemiol*. 2012;65:590-601. https://doi.org/10.1016/j.jclinepi.2011.12.005
5. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)*. 2000;25:3186-3191.
6. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;327:307-310. https://doi.org/10.1016/S0140-6736(86)90837-8
7. Cicchetti D. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess*. 1994;6:284-290.
8. Cleland JA, Fritz JM, Whitman JM, Palmer JA. The reliability and construct validity of the Neck Disability Index and Patient Specific Functional Scale in patients with cervical radiculopathy. *Spine (Phila Pa 1976)*. 2006;31:598-602. https://doi.org/10.1097/01.brs.0000201241.90914.22
9. 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
10. de Vet HC, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide*. Cambridge, UK: Cambridge University Press; 2011.
11. Fairbairn K, May K, Yang Y, Balasundar S, Hefford C, Abbott JH. Mapping Patient-Specific Functional Scale (PSFS) items to the *International Classification of Functioning, Disability and Health* (ICF). *Phys Ther*. 2012;92:310-317. https://doi.org/10.2522/ptj.20090382
12. Hefford C, Abbott JH, Arnold R, Baxter GD. The Patient-Specific Functional Scale: validity, reliability, and responsiveness in patients with upper extremity musculoskeletal problems. *J Orthop Sports Phys Ther*. 2012;42:56-65. https://doi.org/10.2519/jospt.2012.3953
13. Horn KK, Jennings S, Richardson G, van Vliet D, Hefford C, Abbott JH. The Patient-Specific Functional Scale: psychometrics, clinimetrics, and application as a clinical outcome measure. *J Orthop Sports Phys Ther*. 2012;42:30-42. https://doi.org/10.2519/jospt.2012.3727
14. 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
15. Maughan EF, Lewis JS. Outcome measures in chronic low back pain. *Eur Spine J*. 2010;19:1484-1494. https://doi.org/10.1007/s00586-010-1353-6
16. 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
17. Nakamaru K, Aizawa J, Koyama T, Nitta O. Reliability, validity, and responsiveness of the Japanese version of the Patient-Specific Functional Scale in patients with neck pain. *Eur Spine J*. 2015;24:2816-2820. https://doi.org/10.1007/s00586-015-4236-z
18. 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
19. Sharma S, Palanchoke J, Reed D, Abbott JH. Translation, cross-cultural adaptation and psychometric properties of the Nepali versions of numerical pain rating scale and global rating of change. *Health Qual Life Outcomes*. 2017;15:236. https://doi.org/10.1186/s12955-017-0812-8
20. Stratford P, Gill C, Westaway M, Binkley J. Assessing disability and change on individual patients: a report of a patient specific measure. *Physiother Can*. 1995;47:258-263. https://doi.org/10.3138/ptc.47.4.258
21. 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

APPENDIX A

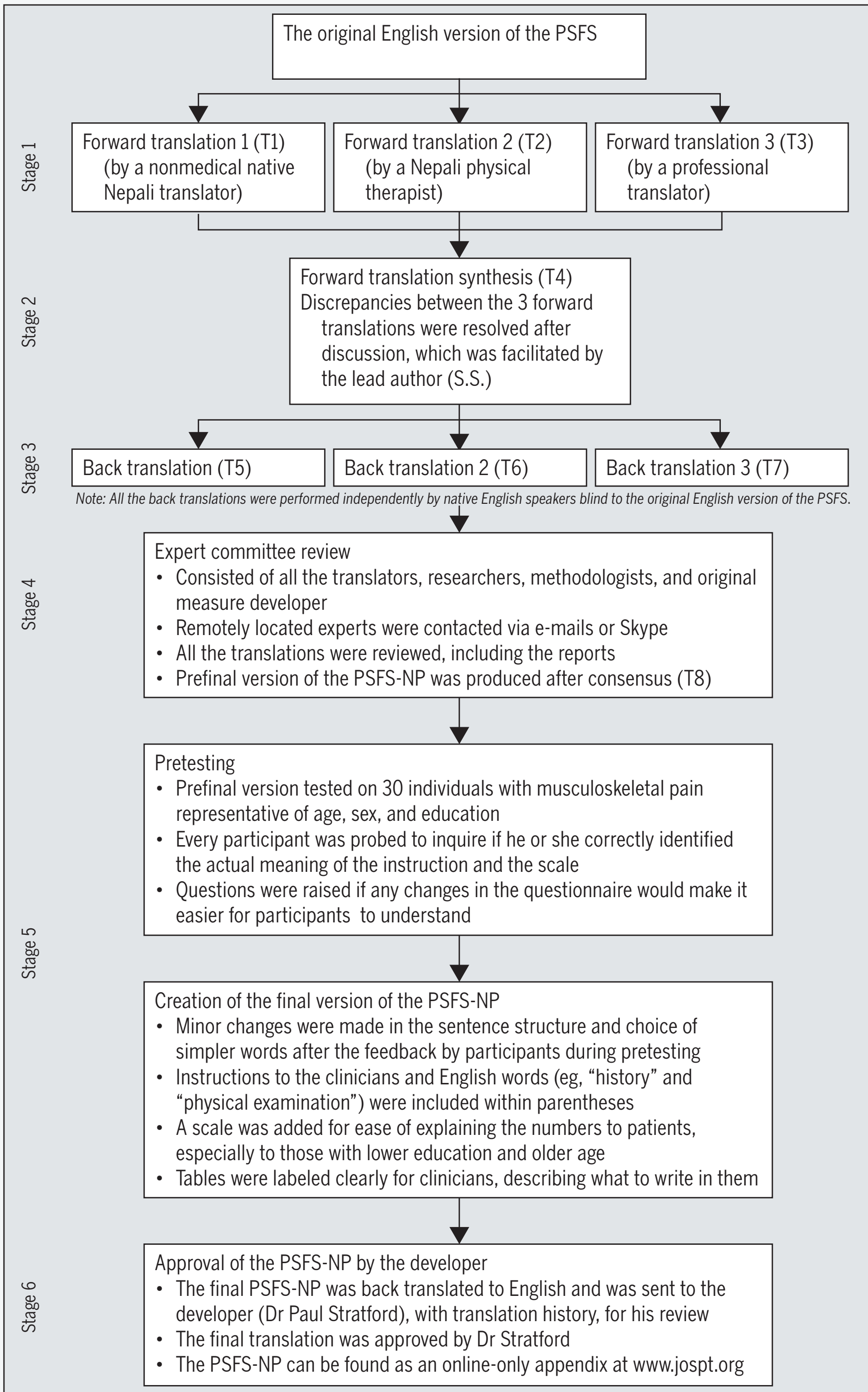


FIGURE. Translation history of the Nepali version of the PSFS. Abbreviations: PSFS, Patient-Specific Functional Scale; PSFS-NP, Nepali version of the Patient-Specific Functional Scale.

APPENDIX B

Nepali version of the Patient-Specific Functional Scale (PSFS-NP)

बिरामी-विशेष कार्य प्रश्नावली (स्केल)

चिकित्सकले तल लेखिएका पङ्क्तिहरु बिरामीलाई पढेर सुनाएपछि बिरामीले उत्तर दिएको क्रियाकलापहरु तल तालिकामा भर्नु पर्ने हुन्छ ।

नोटः बिरामीको समस्याहरु *(history)* सोधे पछि र शारीरिक जाँच *(physical examination)* भन्दा अगाडि भर्नुहोस् ।

प्रारम्भिक जाँचमा पढ्नुहोस् :

म तपाईंलाई तीन देखि पाँचवटा त्यस्ता कार्यहरु पहिचान गर्न लगाउँछु, जुन तपाईंको ____________ समस्याको कारणले गर्न सक्नु हुन्न वा गर्न गाह्रो हुन्छ । के आज ती क्रियाकलापहरु छन् जुन तपाईंले गर्न सक्नुहुन्न वा गर्न गाह्रो हुन्छ ? *(चिकित्सकः बिरामीलाई तलको स्केल देखाउनुहोस्)*

थपः अरु केहि कार्य छन् जुन गर्न थोरै मात्र गाह्रो हुन्छ ? जस्तै कि कार्य जसलाई तपाईं “६” वा बढी अङ्क दिनुहुन्छ ? थप दुईवटासम्म कार्य भन्नुहोस् । *(थप १ र थप २ मा उल्लेख गर्नुहोस्)*

पूनः जाँचमा पढ्नुहोस् :

जब मैले तपाईंलाई पहिले _______ *(मिति उल्लेख गर्नुहोस्)* मा जाँच्दा *(सूचिमा लेखिएका प्रत्येक कार्यहरु पढेर बिरामीलाई सुनाउनुहोस्)* ___________ कार्यहरुमा तपाईंले समस्या छ भन्नुभएको थियो ।

के आज पनि तपाईंलाई कार्य १ गर्न गाह्रो छ *(बिरामीलाई स्कोर / मूल्याङ्कन गर्न लगाउनुहोस्)* ? कार्य २ गर्न गाह्रो छ *(बिरामीलाई स्कोर / मूल्याङ्कन गर्न लगाउनुहोस्)* ? कार्य ३ गर्न गाह्रो छ *(बिरामीलाई स्कोर / मूल्याङ्कन गर्न लगाउनुहोस्)* ? इत्यादि ।

बिरामी विशेष - कार्य प्रश्नावली - एउटा कुनै अङ्क छान्नुहोस् ।

० १ २ ३ ४ ५ ६ ७ ८ ९ १०

कार्य गर्नै सकिँदैन

घाइते वा समस्या हुनुभन्दा अगाडि जस्तैगरि कार्य गर्न सकिन्छ

मिति र अङ्क

	प्रारम्भिक जाँच	पूनः जाँच १	पूनः जाँच २	पूनः जाँच ३
तपाईंलाई कुन कार्य गर्न गाह्रो हुन्छ, तल उल्लेख गर्नुहोस् ।	मितिः	मितिः	मितिः	मितिः
१.				
२.				
३.				
थप १.				
थप २.				

माथिका बाकसहरुमा बिरामीलाई “कार्य गर्न“ कत्तिको गाह्रो हुन्छ, बिरामीले मुल्याङ्कन गरेको ० - १० अंक लेख्नुहोस् ।

Sharma et al.,JOSPT, 2018

APPENDIX C

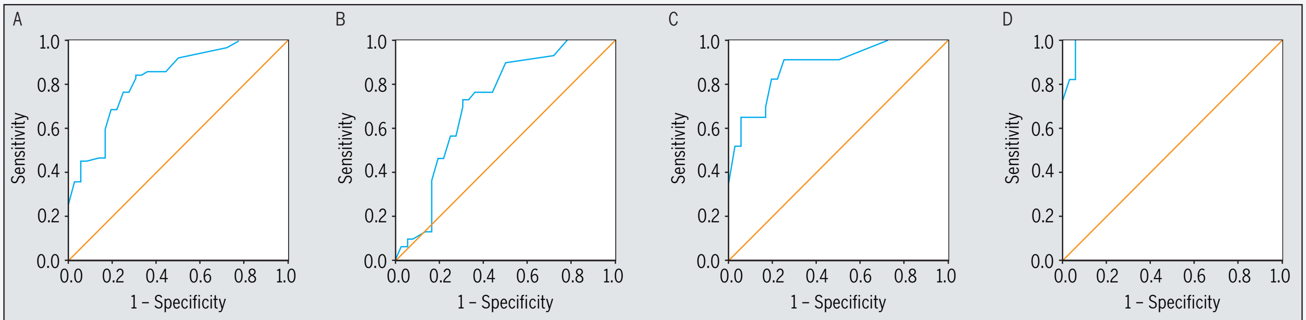


FIGURE. Receiver operating characteristic curves. (A) Stable group (GROC, 4) versus improved group (GROC, 5-7), (B) stable group (GROC, 4) versus small improvement group (GROC, 5), (C) stable group (GROC, 4) versus medium improvement group (GROC, 6), and (D) stable group (GROC, 4) versus large improvement group (GROC, 7). Abbreviation: GROC, global rating of change.

TISHYA A.L. WREN, PhD[1,2] • NICOLE M. MUESKE, MS[1] • CHRISTOPHER H. BROPHY, MD[2] • J. LEE PACE, MD[3]
MIA J. KATZEL, DPT[1] • BIANCA R. EDISON, MD[1,2] • CURTIS D. VANDENBERG, MD[1,2] • TRACY L. ZASLOW, MD[1,2]

Hop Distance Symmetry Does Not Indicate Normal Landing Biomechanics in Adolescent Athletes With Recent Anterior Cruciate Ligament Reconstruction

Anterior cruciate ligament (ACL) injuries in adolescent athletes have been increasing over the past 2 decades, due to a greater number of adolescents participating in high-demand, organized sports.[12] Though surgical anterior cruciate ligament reconstruction (ACLR) is the preferred treatment, there remains a significant risk of graft retear or a new injury in the contralateral healthy knee within the first 7 months after young athletes return to sport.[15,25] Despite considerable debate over when adolescents with ACLR should be allowed to return to sport, clear and objective clinical guidelines have yet to gain consensus.[1]

BACKGROUND: Return-to-sport protocols after anterior cruciate ligament reconstruction (ACLR) often include assessment of hop distance symmetry. However, it is unclear whether movement deficits are present, regardless of hop symmetry.

OBJECTIVES: To assess biomechanics and symmetry of adolescent athletes following ACLR during a single-leg hop for distance.

METHODS: Forty-six patients with ACLR (5-12 months post surgery; 27 female; mean ± SD age, 15.6 ± 1.7 years) were classified as asymmetric (operative-limb hop distance less than 90% that of nonoperative limb [n = 17]) or symmetric (n = 29) in this retrospective cohort. Lower extremity biomechanics were compared among operative and contralateral limbs and 24 symmetric controls (12 female; mean ± SD age, 14.7 ± 1.5 years) using analysis of variance.

RESULTS: Compared to controls, asymmetric patients hopped a shorter distance on their operative limb ($P<.001$), while symmetric patients hopped an intermediate distance on both sides ($P\geq.12$). During landing, the operative limb, regardless of hop distance, exhibited lower knee flexion moments compared to controls and the contralateral side ($P\leq.04$), with lower knee energy absorption than the contralateral side ($P\leq.006$). During takeoff, both symmetric and asymmetric patients had less hip extension and smaller ankle range of motion on the operative side compared with controls ($P\leq.05$). Asymmetric patients also had lower hip range of motion on the operative, compared with the contralateral, side ($P = .001$).

CONCLUSION: Both symmetric and asymmetric patients offloaded the operative knee; symmetric patients achieved symmetry, in part, by hopping a shorter distance on the contralateral side. Therefore, hop distance symmetry may not be an adequate test of single-limb function and return-to-sport readiness. *J Orthop Sports Phys Ther 2018;48(8):622-629. Epub 30 Mar 2018. doi:10.2519/jospt.2018.7817*

KEY WORDS: *biomechanics, motion analysis, pediatrics, single-leg hop*

Commonly utilized clinical return-to-sport criteria include time since surgery, visual movement analysis, and symmetry of strength and hop distance.[16] Although symmetry is often considered a prerequisite for return to sport, recent research has indicated that comparison with the contralateral limb may not be ideal because of deficits on the contralateral side after surgery.[7,24] Gokeler et al[7] found that young adult athletes had bilateral deficits on 4 different hop tests 7 months post ACLR compared with controls. Wellsandt et al[24] concluded that patients who met a 90% symmetry criterion for strength and hop tests 6 months post ACLR would not pass if compared against performance on the contralateral limb before, rather than after, surgery. Thus, symmetry of strength and hop distance may not indicate adequate recovery and readiness for return to sport.

In addition, assessment of hop distance symmetry does not offer information about movement quality, which has been associated with ACL injury risk for functional tasks such as cutting and landing.[9,14,21] Xergia et al[28] found no relationship between hop distance symmetry and kinematic or kinetic performance, only a correlation with isokinetic strength. Orishimo et al[20] showed that,

[1]Children's Orthopaedic Center, Children's Hospital Los Angeles, Los Angeles, CA. [2]Keck School of Medicine, University of Southern California, Los Angeles, CA. [3]Connecticut Children's Medical Center, Hartford, CT. All study procedures were approved by the Children's Hospital Los Angeles Institutional Review Board. 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 Tishya A.L. Wren, Children's Hospital Los Angeles, 4650 Sunset Boulevard, Number 69, Los Angeles, CA 90027. E-mail: twren@chla.usc.edu

post ACLR, patients meeting an 85% hop distance symmetry criterion still exhibited kinematic and kinetic asymmetries during both takeoff and landing, including reduced range of motion at all lower extremity joints and lower moments and power absorption at the knee, with compensation at the hip and ankle. Previous studies of limb asymmetry have also focused on adults, although ACL injury and reinjury have a much higher incidence in adolescents.[12,14] Little is known about movement quality and its relationship to hop distance symmetry in adolescent patients following ACLR.

The purpose of this study was to compare single-leg hop biomechanics in adolescent patients with recent ACLR among operative limbs and contralateral nonoperative limbs, and in uninjured controls, to determine whether 90% hop distance symmetry may indicate normalization of biomechanics and, therefore, return-to-sport readiness. The authors hypothesized that both symmetric and asymmetric adolescents would demonstrate reduced flexion and loading of the reconstructed knee, with compensation at the hip and ankle, during the rehabilitation period.

METHODS

This retrospective study examined data from a consecutive series of patients aged 12 to 18 years who were seen in the Children's Orthopaedic Center, Children's Hospital Los Angeles Motion and Sports Analysis Laboratory between February 2013 and February 2017 for assessment of rehabilitation progress following unilateral ACLR. Patients were excluded if they had a history of other serious lower extremity injury or surgery within the previous 5 years, had a previous ACL injury, could not complete the tasks, or had missing motion-analysis data during landing. Patients were not yet cleared for return to full activity at the time of testing. The authors also examined retrospective data from controls in the same age range who had been tested to provide normative data for the laboratory between July 2013 and August 2016. Each control participant played organized sports at least 3 times per week and had no history of lower extremity injury or surgery. Informed consent and assent were obtained from parents and participants in accordance with protocols approved by the Children's Hospital Los Angeles Institutional Review Board. A waiver of consent approved by the Institutional Review Board was used to access some patient data retrospectively.

Data collection was performed by 2 experienced pediatric physical therapists with specialized training in sports biomechanical assessment and motion analysis. Anthropometric measurements were obtained using standard clinical procedures, and a single-leg hop for maximal distance was performed as part of more extensive biomechanical testing. For the single-leg hop, participants were instructed to stand on 1 leg and jump as far as possible, landing on the same leg on a target force plate. For a trial to be successful, participants were required to stick the landing for a minimum of 2 seconds. Participants warmed up for approximately 5 minutes prior to testing and practiced the task 2 to 3 times until they felt comfortable. Three successful trials were performed on each limb, and the trial with the longest hop distance was used for analysis.

During the single-leg hop trials, 3-D motion-analysis data were recorded using an 8- to 10-camera motion-capture system (Vicon 612 and Nexus 2; Oxford Metrics, Yarnton, UK) and a triaxial analog force plate (OR6-5; Advanced Mechanical Technology, Inc, Watertown, MA). A modified plug-in gait[4] marker set was used; the modifications included using patella markers instead of thigh wands[27] and placing the tibia markers directly over the proximal tibial crest.[17,22] During a static calibration trial, the knee axis was defined using knee-alignment devices, which create virtual markers to define the knee flexion axis based on visual alignment of physical axes by the assessor[11]; the ankle axis was defined using markers on the medial and lateral malleoli. Motion data were collected at 120 Hz and force-plate data at 2400 Hz. Motion and force data were filtered using a Woltring filter with a mean-square error of 10 mm. Hop distance was measured as the horizontal displacement of the toe marker at the beginning and end of the jump. Lower extremity kinematics and kinetics were calculated using standard commercial software (Vicon Workstation or Nexus 2; Oxford Metrics).

Kinematic and kinetic measures reflecting shock absorption (sagittal angles, moments, and energy absorption) and dynamic-limb valgus (pelvic drop, hip adduction and internal rotation, knee abduction and knee abduction moments) were evaluated at initial contact and between initial foot contact and maximum knee flexion of the weight-bearing limb during landing. This deceleration phase was studied because it represents the period when the majority of noncontact ACL injuries occur, particularly in the first 40 milliseconds after initial contact.[2,26] Shock absorption is important for dissipating landing forces,[20] and dynamic-limb valgus has been identified as a risk factor for ACL injury.[9] Positive values indicate anterior pelvic tilt and ipsilateral elevation; hip flexion, adduction, and internal rotation; knee flexion and adduction; and ankle dorsiflexion. External moments are reported, with positive values indicating hip and knee flexion moments, ankle dorsiflexion moments, and knee adduction moments. Energy absorption was calculated as net joint power integrated over time, from initial contact to maximum knee flexion, in regions with negative internal power. The sagittal kinematic variables were also examined during takeoff from maximum knee flexion to foot-off to investigate possible biomechanical contributors to hop distance.

Kinetics were not available during takeoff, because participants did not jump from a force plate and takeoff data were missing for some participants who started their jump outside the motion-capture volume.

Moments, ground reaction forces, and power integrals were normalized using nondimensional normalization to account for differences in size among participants.[10] Nondimensional normalization was used instead of traditional mass normalization because it reduces size effects more effectively.[23] Nondimensional normalization accounts for all dimensions of mass, length, and time, resulting in a unitless measure, whereas traditional normalization accounts for only some of these effects (eg, moments in Newton meters per kilogram have remaining components of length and time, in square meters per square seconds).

For analysis, patients were grouped based on their limb symmetry index (LSI), which was defined as the hop distance of the operative limb divided by the hop distance of the contralateral limb, expressed as a percentage. Following typical clinical criteria, patients with an LSI of less than 90% were classified as asymmetric, while patients with an LSI of 90% or greater were classified as symmetric. Controls were considered asymmetric if they hopped less than 90% of the contralateral distance on either limb. Because symmetry of hop distance is considered to be the ideal outcome, the main analyses compared reconstructed limbs of patients to contralateral limbs of patients and to limbs of symmetric controls.

Demographic and anthropometric characteristics were compared between symmetric patients, asymmetric patients, and symmetric controls using chi-square tests for categorical variables and analysis of variance with Bonferroni-adjusted post hoc tests for continuous variables. For the main analysis, analysis of variance was used to assess differences among groups and limbs, including a group-by-limb interaction. The between-subject factor was group (asymmetric, symmetric, or control), and the within-subject factor was limb (operative or nonoperative; all control limbs were considered nonoperative). When the overall analysis of variance indicated significant effects (group, limb, or interaction), pairwise post hoc tests were performed comparing group-limb subgroups (control, symmetric operative, symmetric nonoperative, asymmetric operative, asymmetric nonoperative). These post hoc tests used paired t tests for comparison of operative versus contralateral limbs within patients and unpaired t tests for all other comparisons. Bonferroni adjustment of P values was performed to adjust for the multiple post hoc comparisons. All analyses were performed using Stata Version 14.0 (StataCorp LLC, College Station, TX), with a significance level of .05.

RESULTS

Forty-six patients post ACLR met the eligibility criteria and were included in the study. Their mean ± SD time following surgery was 7.2 ± 1.3 months (range, 5.1-11.7 months). Of 39 controls meeting the eligibility criteria, 15 (38%) had hop distance asymmetry greater than the 10% threshold. Of these, 6 hopped farther on the dominant limb and 9 hopped farther on the nondominant limb. There was no difference in hop distance between the dominant and nondominant limbs of controls (dominant, 1.56 ± 0.36 leg lengths [LL]; 95% confidence interval: 1.44, 1.67 LL; nondominant, 1.56 ± 0.37 LL; 95% confidence interval: 1.44, 1.69 LL; P = .93). Dominant and nondominant control limbs were, therefore, combined in the analysis. Because symmetric hop distance is considered ideal, only the 24 symmetric controls were used for comparison with the patients. Seventeen of 46 patients (37%) were classified as asymmetric.

No significant differences were identified between asymmetric patients, symmetric patients, and symmetric controls in terms of sex, age, height, or time since surgery (**TABLE 1**). Patients had higher weight and body mass index than controls, but the difference was statistically significant for symmetric patients only.

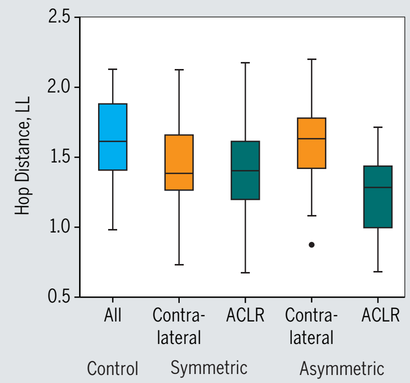


FIGURE 1. Comparison of hop distance among patient reconstructed, patient contralateral, and control limbs for symmetric and asymmetric patients. Abbreviations: ACLR, anterior cruciate ligament reconstruction; LL, leg lengths.

TABLE 1 Comparison of Participant Demographic and Clinical Characteristics Among Groups*

Characteristic	Controls (n = 24)	Symmetric Patients (n = 29)	Asymmetric Patients (n = 17)	P Value
Sex (female), n (%)	12 (50)	16 (55)	11 (65)	.65
Age, y	14.7 ± 1.5	15.6 ± 1.7	15.6 ± 1.7	.08
Height, cm	166.0 ± 9.9	166.9 ± 11.2	164.5 ± 10.9	.78
Weight, kg	54.4 ± 8.8	64.3 ± 11.7†	60.2 ± 16.7	.02
Body mass index, kg/m^2	19.6 ± 1.8	23.0 ± 2.8†	22.2 ± 5.9	.004
Time since surgery, mo	...	7.2 ± 1.4	7.1 ± 1.2	.77
Limb symmetry index, %	100.0 ± 4.8	99.9 ± 8.7	76.6 ± 9.8‡	<.001

Values are mean ± SD unless otherwise indicated. Statistical significance was evaluated using an analysis of variance, with Bonferroni-adjusted post hoc tests for continuous variables and chi-square tests for proportions.
†*P≤.02 versus controls.*
‡*P<.001 versus controls and symmetric patients.*

Compared with controls, asymmetric patients hopped a similar distance on the uninjured limb but a significantly shorter distance on the reconstructed limb (**TABLE 2**, **FIGURE 1**). The average LSI for asymmetric patients was 77% ± 10%, ranging from 53% to 89%. In contrast, despite having an LSI greater than 90%, symmetric patients hopped an intermediate distance on both the operative and nonoperative limbs.

During landing, asymmetric patients had lower knee flexion moments and energy absorption at the knee on the operative side compared with both their contralateral side and with uninjured limbs (controls and nonoperative limbs of symmetric and asymmetric patients) (**FIGURE 2**, **TABLE 2**). They also had lower peak hip and knee flexion angles on the operative side compared with their contralateral side and both sides of symmetric patients, greater plantar flexion at initial contact compared with their contralateral side and controls, and lower peak ankle dorsiflexion compared with both limbs of symmetric patients.

Symmetric patients also had lower knee flexion moments on the operative side compared with controls and the contralateral side, as well as lower energy absorption at the knee compared with their contralateral side. Symmetric patients also had greater hip flexion angles and moments on both sides compared with controls, as well as greater energy absorption at the hip on the operative side.

TABLE 2 Comparison of Kinematic and Kinetic Variables Between Symmetric and Asymmetric Patients and Symmetric Controls During Single-Leg Hop Landing*

	Controls	Symmetric Patients		Asymmetric Patients		*P* Values		
	Both Sides (n = 48 sides)	Nonoperative (n = 29)	Operative (n = 29)	Nonoperative (n = 17)	Operative (n = 17)	Group	Limb	Group by Limb
Distance jumped, LL	1.63 ± 0.32	1.43 ± 0.34	1.43 ± 0.36	1.60 ± 0.33	1.22 ± 0.29[†‡]	.38	<.001	<.001
At initial contact, deg								
Pelvic tilt	11.1 ± 8.6	13.1 ± 13.3	13.3 ± 12.3	13.8 ± 8.3	13.8 ± 10.2	.82	.92	.95
Hip flexion	39.6 ± 8.9	43.2 ± 11.5	45.3 ± 9.1	43.8 ± 7.6	39.4 ± 9.5	.19	.37	.009
Knee flexion	10.8 ± 5.8	11.1 ± 5.9	10.7 ± 6.3	10.6 ± 5.4	8.7 ± 7.3	.70	.28	.48
Ankle dorsiflexion	–1.9 ± 12.2	–4.1 ± 15.4	–6.2 ± 15.0	1.7 ± 8.9	–13.3 ± 15.4[†‡]	.98	<.001	.005
Pelvic obliquity	–9.9 ± 4.2	–10.7 ± 4.0	–10.6 ± 4.1	–11.5 ± 3.7	–12.5 ± 3.7	.12	.64	.54
Hip adduction	–9.7 ± 5.7	–10.4 ± 5.8	–8.8 ± 6.7	–10.8 ± 7.1	–9.9 ± 4.7	.77	.23	.74
Knee adduction	2.8 ± 2.7	1.1 ± 2.7	0.5 ± 3.1	0.9 ± 4.4	0.2 ± 4.1	.13	.14	.88
Hip rotation	3.4 ± 7.3	0.3 ± 6.5	2.5 ± 6.8	4.1 ± 7.0	1.2 ± 8.1	.48	.76	.06
Initial contact to peak knee flexion								
Maximum pelvic tilt, deg	–0.7 ± 4.6	–2.3 ± 4.7	0.02 ±5.4	–3.0 ± 4.4	–1.8 ± 6.4	.72	.66	.46
Maximum hip flexion, deg	55.1 ± 13.1	68.5 ± 13.9[†]	71.1 ± 14.4[†]	65.8 ± 17.0	55.9 ± 16.1[‡§‖]	.002	.02	<.001
Maximum knee flexion, deg	61.1 ± 14.1	68.9 ± 10.4	65.9 ± 10.7	66.3 ± 15.5	53.6 ± 16.6[‡§‖]	.03	<.001	.005
Maximum ankle dorsiflexion, deg	14.8 ± 8.3	17.1 ± 7.1	16.8 ± 7.2	13.2 ± 7.6	9.3 ± 9.1[§‖]	.03	.07	.12
Minimum pelvic obliquity, deg	–10.2 ± 3.8	–10.8 ± 4.0	–10.6 ± 4.1	–11.8 ± 3.8	–12.6 ± 4.0	.10	.76	.58
Maximum hip adduction, deg	4.1 ± 5.9	1.9 ± 5.5	5.2 ± 6.7	1.4 ± 7.6	4.9 ± 7.3	.23	.01	.95
Minimum knee adduction, deg	1.5 ± 3.4	0.4 ± 2.9	0.2 ± 2.8	-0.6 ± 4.7	–0.9 ± 4.7	.15	.66	.94
Maximum hip internal rotation, deg	7.1 ± 6.7	5.4 ± 5.6	7.6 ± 5.8	7.8 ± 7.7	7.7 ± 7.3	.63	.36	.33
Peak ground reaction force, BW	3.1 ± 0.6	2.8 ± 0.5	2.8 ± 0.4	3.1 ± 0.6	3.0 ± 0.6	.13	.12	.61
Average hip flexion moment, ND	0.121 ± 0.060	0.181 ± 0.057[†]	0.189 ± 0.051[†]	0.159 ± 0.070	0.150 ± 0.076	.004	.93	.17
Average knee flexion moment, ND	0.152 ± 0.046	0.133 ± 0.043	0.106 ± 0.032[†§]	0.155 ± 0.059[‖]	0.086 ± 0.061[†‡§]	.68	<.001	.01
Average ankle dorsiflexion moment, ND	0.073 ± 0.050	0.077 ± 0.029	0.077 ± 0.039	0.048 ± 0.052	0.091 ± 0.046[‡]	.66	.009	.009
Energy absorption at hip, ND	0.060 ± 0.044	0.084 ± 0.038	0.096 ± 0.050[†]	0.084 ± 0.056	0.056 ± 0.036[‖]	.04	.19	<.001
Energy absorption at knee, ND	0.122 ± 0.056	0.144 ± 0.052	0.107 ± 0.046[§]	0.150 ± 0.064	0.078 ± 0.048[†‡§]	.17	<.001	.03
Energy absorption at ankle, ND	0.038 ± 0.027	0.044 ± 0.025	0.045 ± 0.024	0.037 ± 0.028	0.048 ± 0.032	.82	.18	.24
Average knee adduction moment, ND	0.098 ± 0.047	0.071 ± 0.032[†]	0.056 ± 0.022[†]	0.075 ± 0.036	0.060 ± 0.024[†]	.02	.01	.98

Abbreviations: BW, body weight; LL, leg lengths; ND, nondimensional.
**Values are mean ± SD unless otherwise indicated.*
[†]P<.05 versus control.
[‡]P<.05 versus asymmetric nonoperative.
[§]P<.05 versus symmetric nonoperative.
[‖]P<.05 versus symmetric operative.

Both symmetric and asymmetric patients had lower average knee adduction moments compared with controls, with the difference being statistically significant for both limbs of symmetric patients and the operative limb of asymmetric patients (**FIGURE 3**, **TABLE 2**).

Takeoff kinematics were available for 25 of 29 symmetric patients, 15 of 17 asymmetric patients, and 19 of 24 symmetric controls. Asymmetric patients demonstrated lower peak knee flexion and hip flexion range of motion on the operative side compared with the contralateral side (**TABLE 3**). Ankle dorsiflexion range of motion was lower on the operative side of both symmetric and asymmetric patients compared with controls. Hip extension (ie, minimum hip flexion) was reduced compared to controls on the operative side in asymmetric patients and bilaterally in symmetric patients.

DISCUSSION

A SIMILAR PERCENTAGE OF ACLR PAtients (37%) and controls (38%) were asymmetric based on a typical 90% hop distance threshold. This suggests that symmetry of hop distance may not indicate ideal biomechanics or return-to-sport readiness. Regardless of hop distance symmetry, adolescent patients with recent ACLR exhibited biomechanical asymmetries and differences from symmetric controls. Both symmetric and asymmetric patients offloaded the reconstructed knee, reducing knee flexion moments and energy absorption. Symmetric patients appeared to offload the knee to the hip, while asymmetric patients offloaded the knee to the ankle. Only minor differences were observed in the frontal or transverse plane. Moreover, a high percentage of controls were also asymmetric in hop distance, suggesting that asymmetry is not solely indicative of injury and healing but may also reflect

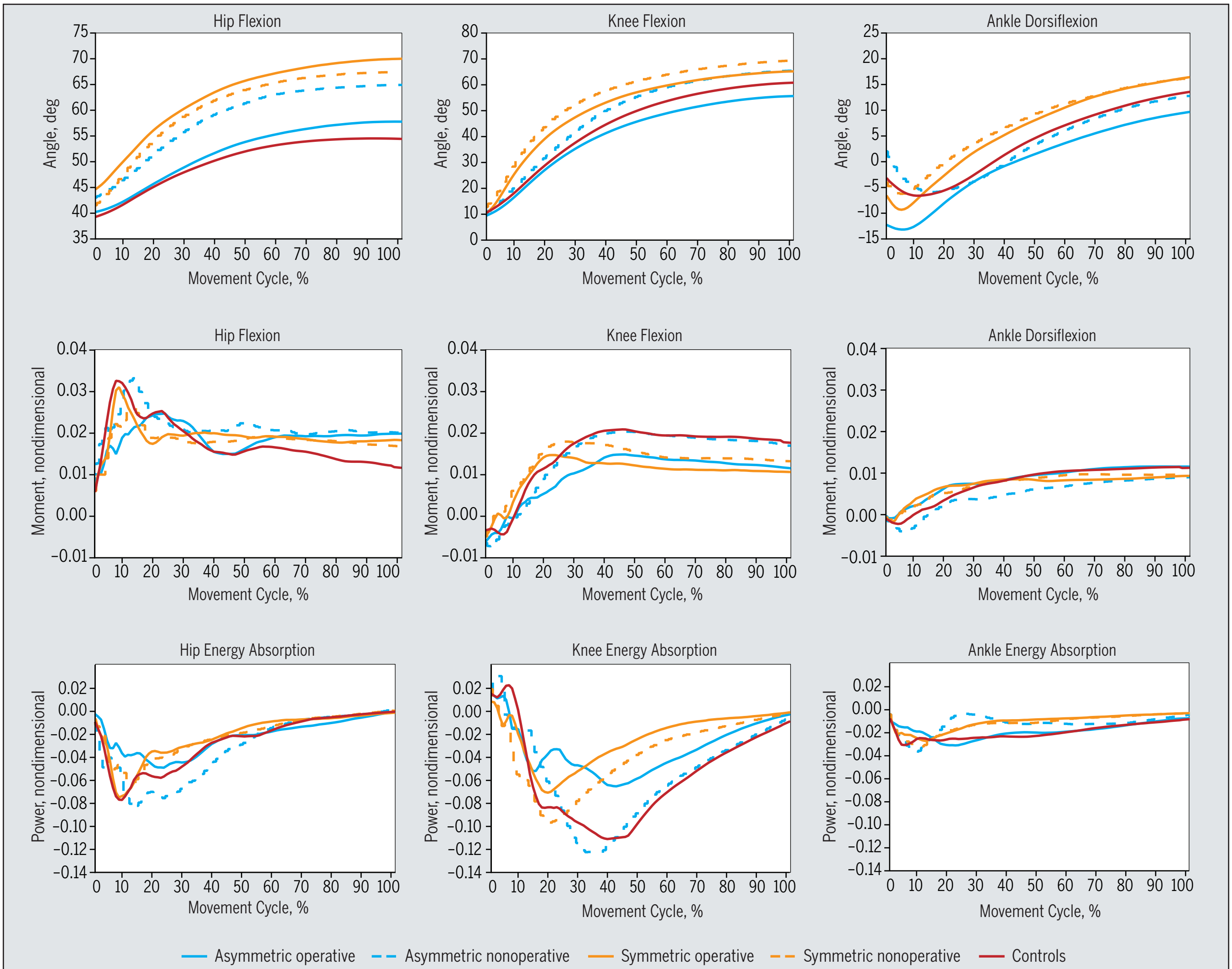


FIGURE 2. Sagittal plane hip, knee, and ankle angles, moments, and powers during landing from initial contact to peak knee flexion.

suboptimal biomechanics in many uninjured adolescent athletes.

Asymmetric patients, by definition, hopped a shorter distance on the operative side, with LSIs ranging from 53% to 89%. However, symmetric patients tended to hop a shorter distance than controls on both sides, suggesting that symmetry may be achieved, at least in part, by decreasing task achievement on the nonoperative side. This is consistent with other recent reports of decreased contralateral-limb performance in ACLR patients who meet limb symmetry criteria[7,24] and may be due to deconditioning, fear, or lack of motivation. In a recent study, Wellsandt et al[24] found that 8 of 11 patients who went on to a second ACL injury had passed 90% LSI criteria for strength and 4 different hop tests, but 6 of the 8 would not have passed if the reconstructed limb had been compared against contralateral-limb function prior to surgery. They therefore recommended that the benchmark for operative-limb function should be based on the performance of the contralateral limb before, rather than after, surgery, due to decreased performance of the contralateral limb after surgery. It is unclear whether and to what extent the smaller differences observed in the current study's symmetric patient group were due to shorter hop distance on the nonoperative side.

With regard to movement quality, all operative limbs, regardless of symmetry, displayed altered movement strategies resulting in decreased loading of the surgical knee. Asymmetric patients exhibited a stiffer landing pattern, with decreased hip flexion and increased ankle plantar flexion. Symmetric patients shifted loading to the hip, increasing hip flexion angles, moments, and energy absorption. These patterns are consistent with previous studies in adults following ACLR.[18,20] Oberländer et al[18] observed offloading of the reconstructed knee to adjacent joints in adult ACLR patients 6 to 12 months after surgery. Orishimo et al[20] showed decreased knee range of motion and power absorption during single-leg hop takeoff and landing in adults 4 to 12 months post ACLR despite 85% hop distance symmetry, with increased moments and power at the hip during takeoff and increased moments and/or power at the ankle during both takeoff and landing. The authors of the current study observed reduced ankle range of motion during

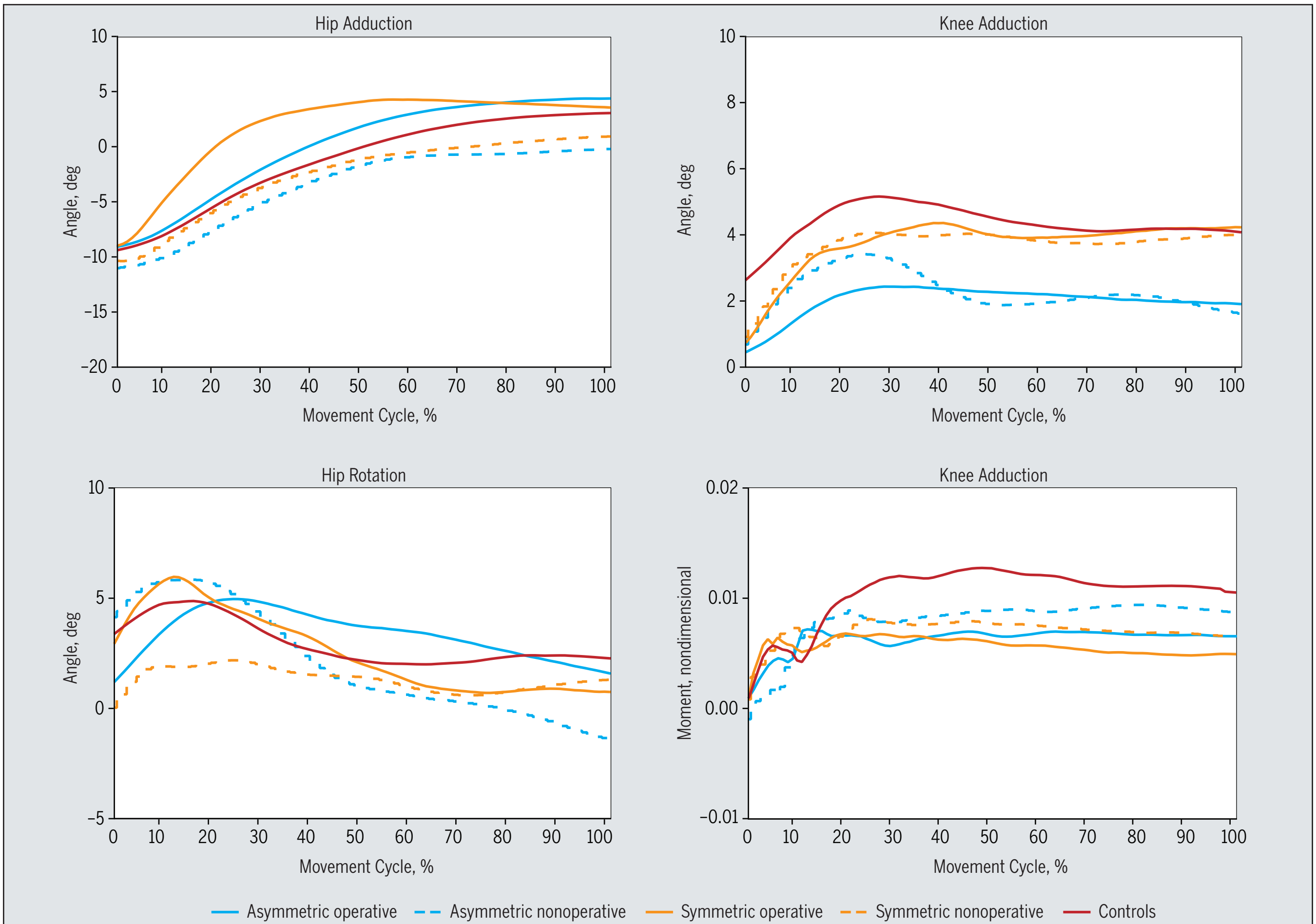


FIGURE 3. Frontal and transverse plane variables during landing from initial contact to peak knee flexion.

TABLE 3 Comparison of Kinematic and Kinetic Variables Between Symmetric and Asymmetric Patients and Symmetric Controls During Single-Leg Hop Takeoff*

	Controls	Symmetric Patients		Asymmetric Patients		*P* Values		
	Both Sides (n = 38 sides)	Nonoperative (n = 25)	Operative (n = 25)	Nonoperative (n = 15)	Operative (n = 15)	Group	Limb	Group by Limb
Takeoff, deg								
Maximum hip flexion	53.9 ± 11.4	58.9 ± 14.6	59.9 ± 11.3	63.0 ± 10.2	55.4 ± 15.8	.23	.04	.007
Maximum knee flexion	56.2 ± 9.8	56.9 ± 7.5	55.7 ± 7.6	61.2 ± 9.2	51.5 ± 10.7[†]	.67	<.001	.002
Maximum ankle dorsiflexion	36.0 ± 28.8	28.5 ± 6.1	29.4 ± 6.0	30.8 ± 6.0	26.1 ± 5.0	.27	.61	.47
Minimum hip flexion	–2.8 ± 10.7	8.6 ± 10.3[‡]	9.7 ± 9.9[‡]	2.8 ± 9.8	9.0 ± 10.2[†‡]	.01	.006	.049
Minimum knee flexion	8.3 ± 7.8	8.6 ± 12.1	10.6 ± 7.6	11.2 ± 8.5	10.4 ± 9.9	.72	.79	.50
Minimum ankle dorsiflexion	–20.8 ± 26.1	–20.8 ± 10.2	–18.3 ± 12.2	–22.6 ± 7.0	–21.3 ± 8.1	.78	.67	.89
Hip flexion range	56.6 ± 15.1	50.2 ± 13.3	50.2 ± 13.4	60.2 ± 11.3	46.4 ± 12.0[†]	.61	<.001	<.001
Knee flexion range	47.9 ± 11.5	48.3 ± 14.0	45.1 ± 10.5	50.0 ± 8.1	41.1 ± 13.5	.85	.03	.30
Ankle dorsiflexion range	56.8 ± 11.8	49.3 ± 10.0	47.7 ± 11.7[‡]	53.3 ± 7.6	47.4 ± 10.1[‡]	.13	.02	.18

**Values are mean ± SD unless otherwise indicated.*
[†]$P<.05$ versus asymmetric nonoperative.
[‡]$P<.05$ versus control.

operative-limb takeoff, regardless of hop distance symmetry. The alternative landing strategies that offload the operative knee in both asymmetric and symmetric patients may reflect an avoidance strategy, but paradoxically may also increase their long-term risk for osteoarthritis in the operative knee by curbing stimulation of normal cartilage production.[3,6,19] Inadequate shock absorption at the knee has been shown to be one risk factor associated with ACL injury during landing tasks.[13] When determining return-to-sport readiness using the single-leg hop for distance test, it appears necessary also to evaluate knee motion in addition to distance symmetry to obtain a more comprehensive and informative measure.

Limitations of the present study include the retrospective sample of patients receiving postsurgical evaluation. These patients were between 5 and 12 months post surgery and were not yet cleared for return to sport, which may limit generalizability to other populations. Due to the relatively small sample size, male and female participants were not analyzed separately, despite there being some evidence for biomechanical variations between sexes.[8] As this study's focus was on "ideal" biomechanics, the primary analysis involved comparison to symmetric controls and excluded asymmetric controls, who were considered to have less than ideal biomechanics. Takeoff data were not available for all participants, and kinetics were not measured during takeoff. Finally, the Vicon plug-in gait model has been shown to have high intersubject variance of frontal plane variables compared to cluster-based models, which may have contributed to the lack of differences in frontal plane variables.[5]

CONCLUSION

Adolescent patients following ACLR surgery landed with decreased loading of the reconstructed knee, regardless of whether their hop distance was symmetric. In fact, patients achieved symmetry, in part, by jumping a shorter distance on the contralateral limb. Therefore, symmetry of hop distance alone appears to be an inadequate benchmark for single-limb function and return-to-sport readiness; quality of motion during the single-leg hop should be considered in conjunction with hop distance symmetry. Because premature return to sport may predispose athletes to future injury,[15] further research is necessary to establish more comprehensive, objective return-to-sport criteria for adolescents post ACLR that can sufficiently reveal biomechanical deficits. ◉

KEY POINTS

FINDINGS: Regardless of hop distance symmetry, patients offload the reconstructed knee 5 to 12 months after surgery.
IMPLICATIONS: During return-to-sport assessment, movement biomechanics should be considered in addition to hop distance symmetry.
CAUTION: This study did not assess strength or motivation, which also should be considered during functional return-to-sport testing.

ACKNOWLEDGMENTS: *The authors would like to thank the sports team in the Motion and Sports Analysis Laboratory: Bitte Healy, Kyle Chadwick, and Henry Lopez.*

REFERENCES

1. Ardern CL, Webster KE, Taylor NF, Feller JA. Return to sport following anterior cruciate ligament reconstruction surgery: a systematic review and meta-analysis of the state of play. *Br J Sports Med*. 2011;45:596-606. https://doi.org/10.1136/bjsm.2010.076364

2. Boden BP, Dean GS, Feagin JA, Jr., Garrett WE, Jr. Mechanisms of anterior cruciate ligament injury. *Orthopedics*. 2000;23:573-578.
3. Chmielewski TL. Asymmetrical lower extremity loading after ACL reconstruction: more than meets the eye. *J Orthop Sports Phys Ther*. 2011;41:374-376. https://doi.org/10.2519/jospt.2011.0104
4. Davis RB, 3rd, Õunpuu S, Tyburski D, Gage JR. A gait analysis data collection and reduction technique. *Hum Mov Sci*. 1991;10:575-587. https://doi.org/10.1016/0167-9457(91)90046-Z
5. Duffell LD, Hope N, McGregor AH. Comparison of kinematic and kinetic parameters calculated using a cluster-based model and Vicon's plug-in gait. *Proc Inst Mech Eng H*. 2014;228:206-210. https://doi.org/10.1177/0954411913518747
6. Englund M. The role of biomechanics in the initiation and progression of OA of the knee. *Best Pract Res Clin Rheumatol*. 2010;24:39-46. https://doi.org/10.1016/j.berh.2009.08.008
7. Gokeler A, Welling W, Benjaminse A, Lemmink K, Seil R, Zaffagnini S. A critical analysis of limb symmetry indices of hop tests in athletes after anterior cruciate ligament reconstruction: a case control study. *Orthop Traumatol Surg Res*. 2017;103:947-951. https://doi.org/10.1016/j.otsr.2017.02.015
8. Hewett TE, Myer GD, Ford KR. Anterior cruciate ligament injuries in female athletes: part 1, mechanisms and risk factors. *Am J Sports Med*. 2006;34:299-311. https://doi.org/10.1177/0363546505284183
9. Hewett TE, Myer GD, Ford KR, et al. Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: a prospective study. *Am J Sports Med*. 2005;33:492-501. https://doi.org/10.1177/0363546504269591
10. Hof AL. Scaling gait data to body size. *Gait Posture*. 1996;4:222-223. https://doi.org/10.1016/0966-6362(95)01057-2
11. Kadaba MP, Ramakrishnan HK, Wootten ME. Measurement of lower extremity kinematics during level walking. *J Orthop Res*. 1990;8:383-392. https://doi.org/10.1002/jor.1100080310
12. LaBella CR, Hennrikus W, Hewett TE, Council on Sports Medicine and Fitness, Section on Orthopaedics. Anterior cruciate ligament injuries: diagnosis, treatment, and prevention. *Pediatrics*. 2014;133:e1437-e1450. https://doi.org/10.1542/peds.2014-0623
13. Leppänen M, Pasanen K, Kujala UM, et al. Stiff landings are associated with increased ACL injury risk in young female basketball and floorball players. *Am J Sports Med*. 2017;45:386-393. https://doi.org/10.1177/0363546516665810
14. Maletis GB, Chen J, Inacio MC, Funahashi TT. Age-related risk factors for revision anterior cruciate ligament reconstruction: a cohort study of 21,304 patients from the Kaiser Permanente Anterior Cruciate Ligament Registry. *Am J Sports Med*. 2016;44:331-336. https://doi.org/10.1177/0363546515614813
15. Morgan MD, Salmon LJ, Waller A, Roe JP, Pinczewski LA. Fifteen-year survival of endoscopic anterior cruciate ligament reconstruction in patients aged 18 years and younger. *Am J Sports Med*. 2016;44:384-392. https://doi.org/10.1177/0363546515623032
16. Myer GD, Paterno MV, Ford KR, Quatman CE, Hewett TE. Rehabilitation after anterior cruciate ligament reconstruction: criteria-based progression through the return-to-sport phase. *J Orthop Sports Phys Ther*. 2006;36:385-402. https://doi.org/10.2519/jospt.2006.2222
17. Nazareth A, Mueske NM, Wren TA. Effect of tibia marker placement on kinematics in pathological gait. *J Appl Biomech*. 2016;32:603-607. https://doi.org/10.1123/jab.2015-0219
18. Oberländer KD, Brüggemann GP, Höher J, Karamanidis K. Altered landing mechanics in ACL-reconstructed patients. *Med Sci Sports Exerc*. 2013;45:506-513. https://doi.org/10.1249/MSS.0b013e3182752ae3
19. Øiestad BE, Holm I, Aune AK, et al. Knee function and prevalence of knee osteoarthritis after anterior cruciate ligament reconstruction: a prospective study with 10 to 15 years of follow-up. *Am J Sports Med*. 2010;38:2201-2210. https://doi.org/10.1177/0363546510373876
20. Orishimo KF, Kremenic IJ, Mullaney MJ, McHugh MP, Nicholas SJ. Adaptations in single-leg hop biomechanics following anterior cruciate ligament reconstruction. *Knee Surg Sports Traumatol Arthrosc*. 2010;18:1587-1593. https://doi.org/10.1007/s00167-010-1185-2
21. Paterno MV, Schmitt LC, Ford KR, et al. Biomechanical measures during landing and postural stability predict second anterior cruciate ligament injury after anterior cruciate ligament reconstruction and return to sport. *Am J Sports Med*. 2010;38:1968-1978. https://doi.org/10.1177/0363546510376053
22. Peters A, Sangeux M, Morris ME, Baker R. Determination of the optimal locations of surface-mounted markers on the tibial segment. *Gait Posture*. 2009;29:42-48. https://doi.org/10.1016/j.gaitpost.2008.06.007
23. Pinzone O, Schwartz MH, Baker R. Comprehensive non-dimensional normalization of gait data. *Gait Posture*. 2016;44:68-73. https://doi.org/10.1016/j.gaitpost.2015.11.013
24. Wellsandt E, Failla MJ, Snyder-Mackler L. Limb symmetry indexes can overestimate knee function after anterior cruciate ligament injury. *J Orthop Sports Phys Ther*. 2017;47:334-338. https://doi.org/10.2519/jospt.2017.7285
25. Wiggins AJ, Grandhi RK, Schneider DK, Stanfield D, Webster KE, Myer GD. Risk of secondary injury in younger athletes after anterior cruciate ligament reconstruction: a systematic review and meta-analysis. *Am J Sports Med*. 2016;44:1861-1876. https://doi.org/10.1177/0363546515621554
26. Withrow TJ, Huston LJ, Wojtys EM, Ashton-Miller JA. The effect of an impulsive knee valgus moment on *in vitro* relative ACL strain during a simulated jump landing. *Clin Biomech (Bristol, Avon)*. 2006;21:977-983. https://doi.org/10.1016/j.clinbiomech.2006.05.001
27. Wren TA, Do KP, Hara R, Rethlefsen SA. Use of a patella marker to improve tracking of dynamic hip rotation range of motion. *Gait Posture*. 2008;27:530-534. https://doi.org/10.1016/j.gaitpost.2007.07.006
28. Xergia SA, Pappas E, Georgoulis AD. Association of the single-limb hop test with isokinetic, kinematic, and kinetic asymmetries in patients after anterior cruciate ligament reconstruction. *Sports Health*. 2015;7:217-223. https://doi.org/10.1177/1941738114529532