Original Research

Reliability and Validity of the Functional Movement Screen™ with a Modified Scoring System for Young Adults with Low Back Pain

Khalid Alkhathami, PT, PhD[1] [a], Yousef Alshehre, PT, PhD[2], Sharon Wang-Price, PT, PhD, OCS, FAAOMPT[3], Kelli Brizzolara, PT, PhD, OCS[3]

[1] Department of Health Rehabilitation, Shaqra University, [2] Physical Therapy Department, University of Tabuk, [3] School of Physical Therapy, Texas Woman's University

International Journal of Sports Physical Therapy

Background

Low back pain (LBP) is one of the most common complaints in individuals who seek medical care and is a leading cause of movement impairments. The Functional Movement Screen (FMS™) was developed to evaluate neuromuscular impairments during movement. However, the reliability and validity of the FMS™ have not yet been established for the LBP population because of a limitation of its original scoring system.

Purpose

The purposes of this study were to determine the reliability and validity of the FMS™ with a modified scoring system in young adults with and without LBP. The FMS™ scores were modified by assigning a zero score only when there was an increase in LBP during the FMS™, not simply for the presence of pain, as in the original FMS™ scoring system.

Study Design

Reliability and validity study.

Methods

Twenty-two participants with LBP (8 males and 14 females, 26.7 ± 4.68 years old) and 22 age- and gender-matched participants without LBP (26.64 ± 4.20 years old) completed the study. Each participant performed the FMS™ once while being scored simultaneously and independently by two investigators. In addition, each participant's FMS™ performance was video-recorded and then was scored by another two investigators separately. The video-recorded performance also was scored twice six weeks apart by the same investigator to determine intra-rater reliability.

Results

The results showed excellent inter-rater and intra-rater reliability of the FMS™ composite score with intraclass correlation coefficients ranging from 0.93 to 0.99 for both groups. In addition, the LBP group scored significantly lower than the group without LBP ($p = 0.008$).

Conclusions

The results indicate that the FMS™ is able to distinguish between individuals with and without LBP, and that it could be a useful test for clinicians to quantify movement quality and to assess movement restrictions in individuals with LBP.

a **Corresponding author**:
Khalid Alkhathami PT, PhD
Assistant Professor
Shaqra University, Department of Health Rehabilitation
Address Department of Health Rehabilitation, College of Applied Medical Sciences, Shaqra University, Shaqra 11961, Kingdom of Saudi Arabia
E-mail: kalkhthami@su

Levels of Evidence

2b.

INTRODUCTION

Low back pain (LBP) is a musculoskeletal disorder that affects more than 80% of people at least once in their lifetime.[1] LBP is considered one of the most common complaints prompting individuals to seek medical care.[2] The total direct and indirect medical spending for LBP is estimated to be between $100 and $200 billion a year.[1] Although a large proportion of individuals who experience an acute episode of LBP experience rapid improvement, the condition is often associated with high recurrence rates.[3,4] In people with LBP, the behavior of fear-avoidance related to LBP and the presence of pain may cause patients to attempt to reduce pain by restricting motions of the spine.[5] In addition to subjective pain complaints, aberrant movement patterns such as painful arc, lateral shifting, or Gower's sign are commonly observed in this patient population.[6] Furthermore, these aberrant movements have been associated with lumbar instability as a result of passive supportive structural lesions and/or lack of muscle control.[6,7] Although aberrant movements can be detected and quantified using imaging, there is no consensus regarding an objective clinical test for quantifying the severity of the abnormal movement patterns in LBP.[8,9]

The Functional Movement Screen (FMS™) was developed to assess functional performance by identifying restrictions and compensations of movement patterns.[10,11] The FMS™ consists of seven component tests, and each test is scored on a scale of 0 to 3, with the total composite score ranging from 0 to 21 points.[10] FMS™ scores less than or equal to 14 have been found to be associated with a higher risk of musculoskeletal injury among firefighters, football players, and female collegiate athletes.[12–14] Additionally, rowers with lower FMS™ scores had a high risk of injury and a higher likelihood of developing LBP.[15] Further, people with chronic pain demonstrated a lower FMS™ composite score as compared to healthy controls.[16]

The reliability of the FMS™ has been established in different healthy populations.[17–19] The inter-rater reliability of FMS™ composite scores in these studies ranges from good (ICC = 0.76) to excellent (ICC = 0.98).[17] In addition, the standard error of measurement (SEM) was found to be 0.92 points, and the minimum detectable change (MDC) was 2.54 points on the 21-point scale.[20] However, these reliability variables were established from physically active populations, such as active-duty service members and athletes.[18] Therefore, these results may not be applied to a general population or a patient population.

The inter-rater reliability and intra-rater reliability of the FMS™ test have been studied in individuals with LBP by Ko et al,[16] using its original scoring system, in which a zero score is assigned when there is presence of pain, regardless of the severity of the pain level. However, Ko et al. did not describe whether or not their patient participants had LBP at the time of testing (baseline). It is likely that some of their participants did not have LBP at bassline testing, because if they did, their FMS™ scores would have been zero. Therefore, it is necessary to modify the original scoring system, as it would underscore severely for those participants who perform movements in proper form but have zero scores simply because they have existing LBP at baseline. In this modified scoring system, a zero score is given only when the participant reports an increase in LBP. Subsequently, when pain intensity does not change during a movement test, it indicates that the movement is performed properly, whereas when there is an increase pain from the baseline during a movement test, it is indicative of an abnormal movement pattern. Therefore, the purposes of this study were to determine inter-rater and intra-rater reliability and construct validity of the FMS™ with a modified scoring system in young adults with and without LBP.

METHODS

STUDY DESIGN

This study is a repeated measure study of the FMS™ with a modified scoring system. In order to determine the inter-rater reliability, two raters scored the same participant simultaneously in real time, and two additional raters scored the same participant independently by reviewing video-recorded sessions. Further, the video-recorded performance was scored twice six weeks apart by the same investigator to determine intra-rater reliability. The construct validity was determined using the known-groups method by comparing the modified FMS™ scores between the participants with and without LBP. Prior to data collection, this study was approved by the Institutional Review Board of Texas Woman's University.

PARTICIPANTS

Using G*Power version 3.1,[21] an *a priori* power analysis was performed to calculate the sample size needed to detect a significant difference between participants with and without LBP. A total of 44 participants, 22 for each group were needed to achieve a power of 0.80 using an α of 0.05 and an effect size of 0.80. The effect size of 0.80 was chosen based on the findings of a previous FMS™ study in an LBP population.[16]

A convenience sample of the participants for the study was recruited from local communities. The eligible participants were young adults between 18 and 40 years of age. In addition, the eligible participants for the LBP group were young adults who experienced repeated episodes of persistent or recurrent LBP in the past year and were not receiving care at the time of the study from a physician or other practitioners, such as a physical therapist or chiropractor.[22] The asymptomatic participants, the group without LBP, were young adults who did not have current LBP and had not had any episode of LBP in the previous year. Participants in both groups were excluded from the study if they had any of the following conditions which might affect the performance of the FMS™: (1) previous surgery to the lumbar spine or abdomen, (2) current pregnancy, (3) any

neurological symptoms in the lower extremities, or (4) any congenital abnormality in the lumbar spine.

RATERS

Four investigators scored the FMS™ test with the modified scoring system for each participant in this study. Prior to data collection, all investigators were required to read the FMS™ manual and to watch FMS™ videos at least once in order to be familiar with each test of the FMS™. Additionally, before data collection began, all investigators practiced the FMS™ for three hours using the modified scoring system, with the principal investigator (KA) who is a certified in FMS[TM].

INSTRUMENTATION

FUNCTIONAL MOVEMENT SCREEN

The Functional Movement Screen™ kit (Functional Movement Systems Inc., Chatham, VA), consisting of a two-inch by six-inch board, one four-foot-long dowel, two short dowels, and an elastic cord, was used to administer the FMS™.[10] The FMS™ includes seven movement tests: deep squat, hurdle step, in-line lunge, shoulder mobility, active straight-leg-raise, trunk stability push-up, and rotary stability.[10] In addition, there are three clearance screens (impingement-clearing test, press-up clearing test, and posterior-rocking clearing test), which are used to determine if the participants have pain associated with internal rotation and flexion of the shoulder, spinal flexion, and spinal extension, repectively.[10,11] Three FMS™ test components are associated with a clearance screen: the shoulder mobility test with the impingement clearance screen, the push-up test with the press-up clearance screen, and the rotator stability test with the posterior rocking clearance screen. However, because this study focused on LBP population, the impingement-clearing test was excluded because it is not related to LBP area and the other two clearing tests (press-up clearing test and posterior-rocking clearing test) were included in this study. In the original scoring system, each FMS™ test is scored on a four-point scale: 3 if the movement task is performed perfectly without compensations, 2 if completion of the task requires compensatory movements, 1 if the participant is unable to perform the movement as required, or 0 if a participant feels pain during the movement task.[10] However, for the purpose of this study, the FMS™ scores were modified so that a zero score was given only if the participants reported an increase in LBP rather than simply for the presence of pain. The total composite score ranging from 0 to 21 was calculated for analysis.[10]

DIGITAL VIDEO CAMERAS

Two digital video cameras (Nikon, Sendai, Japan) were used to record the participant's performance on all of the FMS™ test components. The cameras were positioned to obtain the frontal and sagittal views of each participant from a distance of approximately 10 feet.

PROCEDURES

Participants who met inclusion criteria were informed of the risks and the procedures of the study, and then signed an informed consent form if they agreed to participate. After consent was obtained, the demographic characteristics (i.e., age, sex, height, weight, leg dominance) of each participant were collected, followed by a physical examination to determine the eligibility of the participant. Next, participants with LBP were asked to rate their current pain using the Numeric Pain Rating Scale (NPRS) and to complete the Modified Oswestry Low Back Pain Disability Questionnaire (OSW) to determine their disability level. Both the NPRS and OSW have been shown to be reliable and valid measurements for LBP.[23–26]

Next, each participant performed the seven tests of the FMS™ in the same order as described in Cook et al.'s studies.[10,11] The two cameras were used to record the entire FMS™ testing session, and the two investigators (Rater 1 and Rater 2) were responsible for video recording. Another two investigators (Rater 3 and Rater 4) independently scored the FMS™ test in real time in order to determine inter-rater reliability. One of these two investigators was responsible for the verbal instructions required for performing the FMS™, while the other investigator was responsible for the demonstration of each test of the FMS™. The verbal instructions and demonstration were repeated if necessary. The participants were asked to rate their pain level using the NPRS before each of the FMS™ tests. This pain score was used as baseline to monitor changes of pain level during and after each test. Participants performed three trials for each of the seven FMS™ tests, and the best score from the three trials was recorded. In addition, each participant performed two clearance screens, namely the press-up clearing screen after the push-up test and the posterior-rocking clearing screen after the rotator stability test. The clearance screens were graded as negative or positive. For example, if a participant had no pain or if the pain level was the same as that as the baseline, the clearance screen was considered negative. Conversely, if there was an increase in pain, the clearance screen was considered positive and the associated FMS™ test was scored zero. Five of the seven FMS™ tests (hurdle step, in-line lunge, shoulder mobility, active straight-leg-raise, and rotary stability test) were performed on both sides, first on the participant's right side first and then on the left. For the FMS™ tests that were scored on both limbs, the lower score was used to compute the composite score. The scores from the seven FMS™ tests were added together to get a composite score. Later, the two investigators (Rater 1 and Rater 2) who were responsible for video-recording scored the FMS™ tests independently while viewing the video-recorded sessions. To determine intra-rater reliability, two raters scored each participant separately through watching the video-recording at two separate times, six-weeks apart.[27] Participants were asked whether or not pain increased throughout the FMS™ testing and video-recording, so that the Rater 1 and Rater 2 could score it accordingly when viewing the videos later.

Table 1: Participants' Characteristics for the Low Back Pain and Asymptomatic Groups

Variables	LBP Group (n=22)	Asymptomatic Group (n=22)	*p*-value
Age (yrs)	26.73 ± 4.68	26.64 ± 4.20	0.946
Gender	8 men; 14 women	8 men; 14 women	1.000
Height (m)	1.71 ± 0.06	1.71 ± 0.07	0.826
Weight (kg)	72.08 ± 18.47	66.31 ± 8.73	0.193
BMI (kg/m^2)	24.57 ± 5.63	22.51 ± 1.95	0.113
NPRS (0-10)	2.82 ± 1.01	NA	NA
OSW (%)	12.27 ± 7.38	NA	NA
Modified FMS? (0-21)	14.07 ± 2.80	16.16 ± 2.08	0.008

BMI= body mass index; FMS™= functional movement screen; n= sample size; NPRS= numeric pain rating scale; OSW= modified Oswestry low back pain disability questionnaire.

Table 2: Participants' Characteristics for Intra-rater Reliability Analysis.

Variables	LBP Group (n = 12)	Asymptomatic Group (n = 12)	*p*-value*
Age (yrs)	26.08 ±4.03	25.33 ±2.99	0.610
Gender	3 men; 9 women	4 men; 8 women	0.653
Height (m)	1.73 ±0.05	1.70 ±0.08	0.245
Weight (kg)	70.31 ±16.64	65.85±9.25	0.426
BMI (kg/m2)	23.26±4.67	22.66±2.13	0.692

BMI= body mass index; LBP= low back pain.
*Independent t-tests for ratio data and chi-square tests for categorical data.

STATISTICAL METHODS

The collected data was analyzed using IBM SPSS Version 25 (IBM Corp., Armonk, NY, USA). Descriptive statistics, including mean and standard deviation (SD), were used to describe the participants' demographic characteristics, NPRS scores, OSW scores, and FMS™ composite scores. An independent *t*-test was used to compare the demographic characteristics between the participants with and without LBP. Inter-rater reliabilities were assessed by using model 2 form *k* intraclass correlation coefficients (ICC$_{2, k}$) with a 95% confidence interval (CI) for each group.[28] Intra-rater reliability were assessed by using model 2 form *k* intraclass correlation coefficients (ICC$_{3, k}$) with a 95% CI for each group.[28] In addition, the SEM was calculated for the inter-rater and intra-rater reliability of the FMS™ composite score using the formula: $\text{SEM} = \text{SD}\sqrt{1 - ICC}$ at 95% CI.[28] Then, the MDC was calculated using the following equation: $\text{MDC } 95 = 1.96 \times \text{SEM} \times \sqrt{2}$ to determine the minimal degree of change in the composite score of the FMS™, which reflects the threshold of error measurement at the 95% CI.[28] The values of MDC were rounded to an exact number that reflected the measurement scale of the FMS™. ICC values of 0.90–0.99 indicated excellent reliability, 0.80–0.89 indicated good reliability, 0.70–0.79 indicated fair reliability, and 0–0.69 indicated poor reliability.[29] To determine the construct validity, an independent *t*-test was used for the total modified FMS™ scores and Mann-Whitney *U* test for each FMS™ test component (ordinal data) to compare between participants with and without LBP. The significant levels for all analyses were set at *p* < 0.05.

RESULTS

All of the 44 participants who met the inclusion criteria completed the study. The characteristics of the 44 participants, 22 (8 men, 14 women) in each group, are displayed in Table 1. There were no significant differences in the participant characteristics (i.e., gender, age, height, weight, and body mass index) between individuals with and without LBP.

In addition, 22 of 44 participants were video-recorded. The characteristics of these 22 participants are shown in Table 2. Similarly, there were no significant differences in the participant characteristics between the video-recorded participants with LBP and those without LBP.

Table 3 illustrates the means and standard deviations of the modified FMS™ composite score and each test component scores for each group. Table 4 shows the means, SDs, and range of the FMS™ total composite score for each rater. The inter-rater reliability results and the 95% CI for the FMS™ composite scores are shown in Table 5. Overall, the inter-rater reliability and intra-rater reliability were excellent for the modified FMS™ composite scores collected in real-time, via watching videos, and in real-time vs. via watching videos for each group. In addition, participants with LBP had significantly lower modified FMS™ composite scores as compared to the asymptomatic group (LBP group:

Table 3: Mean and Standard Deviation of the Modified Functional Movement Screen Scores for both Low Back Pain and Asymptomatic Groups. Reported as score ± SD.

Variables	Low Back Pain Group (n=22)	Asymptomatic Group (n=22)	p-value
Deep squat	1.9 ± 1.0	2.5 ± 0.6	0.033*
Hurdle step	2.2 ± 0.7	2.3 ± 0.5	0.813
In-line lunge	2.5 ± 0.9	2.6 ± 0.5	0.774
Shoulder mobility	2.5 ± 0.7	2.5 ± 0.7	0.881
Active straight leg raise	2.0 ± 1.0	2.5 ± 0.5	0.158
Trunk stability push-up	1.3 ± 1.2	2.0 ± 1.0	0.037*
Rotary stability	1.7 ± 0.7	1.8 ± 0.4	0.847
Total score (0-21)	14.1 ± 2.8	16.2 ± 2.1	0.008*

*Statistically significantly different at p < 0.05.

Table 4: The Modified Composite Scores of the Functional Movement Screen for Both Low Back Pain and Asymptomatic Groups. Reported as mean ± SD

Rater	LBP Group	Range		Asymptomatic Group	Range	
	Mean ± SD	(Min -Max)	n	Mean ± SD	(Min-Max)	n
Rater 1 (real-time)	14.1 ± 2.8	9 - 18	22	16.2 ± 2.1	12 - 19	22
Rater 2 (real-time)	14.0 ± 2.8	8 - 18	22	16.1 ± 2.1	12 - 20	22
Rater 3A* (video)	14.3 ± 2.1	11 - 17	12	16.4 ± 2.1	13 - 19	12
Rater 4 (video)	13.9 ± 2.4	10 - 18	12	16.2 ± 2.0	13 - 19	12
Rater 3B* (video)	14.4 ± 2.4	11 - 18	12	16.3 ± 1.8	13 - 19	12

LBP= low back pain; SD= standard deviation.
*3A and 3B were performed 6 weeks apart.

Table 5: Intraclass Correlation Coefficient (ICC), Standard Errors of Measurement (SEM), and Minimal Detectable Change (MDC) Values for Inter-rater and Intra-rater Reliability of the Modified Functional Movement Screen.

Reliability Type	Loa Back Pain Group			Asymptomatic Group		
	ICC [95%CI]	SEM	MDC_{95}	ICC [95%CI]	SEM	MDC_{95}
Inter-rater (real-time)	0.99 [0.98, 0.99]	0.38	1.05	0.96 [0.90, 0.98]	0.38	1.05
Inter-rater (video)	0.94 [0.78, 0.98]	0.58	1.62	0.93 [0.77, 0.98]	0.58	1.62
Inter-rater (real-time vs. video)	0.98 [0.42, 0.99]	0.35	0.98	0.98 [0.94, 0.99]	0.35	0.98
Intra-rater (video vs. video)	0.98 [0.93, 0.99]	0.33	0.90	0.97 [0.91, 0.99]	0.33	0.90

CI= confidence interval; MDC_{95}= minimal detectable change at the 95% level of confidence.

14.07 ± 2.80 points, asymptomatic group: 16.16 ± 2.08 points, $p = 0.008$).

DISCUSSION

The results of this study showed good-to-excellent intra-rater reliability and inter-rater reliability using the modified FMS™ scoring system for both groups when the scores were collected in real time. These results are consistent with what has been reported in previous studies for the asymp-

tomatic group, such as in Onate et al.'s study[29] (ICC = 0.98) and Parenteau-G et al.'s study[27] (ICC = 0.96) in, which the inter-rater reliability was examined on a young active population by using two raters and the real-time method for scoring.[27,29] Excellent inter-rater reliability (ICC = 0.97) was also found in a study in which physically active individuals aged between 18 and 40 years were examined.[30] Similar to this current study, only two investigators performed scoring in real time. On the contrary, Teyhen et al.[20] reported a fair inter-rater reliability (ICC = 0.76) of the FMS™ for an active young population. The conflicting results could be due

to the use of eight raters in the study by Teyhen et al., thus increasing variance among the raters, and therefore, resulting in a lower ICC value.[31]

In addition, when using the video-recording method, the intra-rater and inter-rater reliability of the FMS™ composite score was excellent for both groups. Leeder et al.[32] reported a high inter-rater reliability (ICC = 0.90) of the FMS™ scores among 20 raters in a young adult athletic population who were pain free. In the study by Leeder et al. three cameras were used and were positioned to the front, side, and overhead. Similar to this current study, Gulgin et al.[33] also used two cameras positioned in the sagittal and frontal views to record the FMS™ tests. They also demonstrated good inter-rater reliability (ICC = 0.880) of the FMS™, which was assessed by four raters with various experiences in administering the FMS™.[33] Both the Gulgin et al. study and this current study indicated that the FMS™ does not appear to require extensive experience in order to achieve good reliability. Good reliability achieved by a novice user can be due to the inherent standard criteria of the FMS™ for performance interpretation and scoring.[33] Further, the inter-rater reliability was excellent between the modified FMS™ scores assessed by one investigator watching the video recordings and by the other investigator in real-time. This result suggests that scoring a recorded FMS™ performance was consistent and reliable as scoring the test in real-time.

The ICC values showed good-to-excellent intra-rater reliability of the FMS™ with the modified scoring system for both the asymptomatic and LBP populations, The result of the current study is in agreement with a study which reported ICC = 0.96 for an asymptomatic group using a similar research design.[27] Gribble et al. attributed experiences in use of the FMS™ to their high intra-rater reliability (ICC = 0.94) because their novice group had lower inter-rater ICC values (0.37 - 0.75).[34] However, experience did not appear to be a factor in this current study because the investigator who scored the video-recorded FMS™ performances twice was a novice rater. Consequently, the MDC_{95} values of this study differed from those reported by Teyhen et al.,[20] who found a higher MDC_{95} value (2.54 points) in active duty service members. The variances in MDC_{95} values result from the differences in the ICC values between the two studies as the level of reliability affects the MDC based on the calculation formula.[28]

The results showed a significant difference in the modified FMS™ scores between the asymptomatic group and the LBP group, indicating that the modified FMS™ was capable of distinguish young adults with LBP from those without LBP. The result was in agreement with those in the Ko et al. study,[16] in which the authors also found differences between individuals with and without chronic LBP. Although the FMS™ composite scores were similar in both studies for the asymptomatic participants, the FMS™ composite score for the patient participants of this study (14 points) was higher than that in the Ko et al. study (11 points). Chronicity of LBP in the Ko et al. study could have contributed to difference between the two studies. In addition, the participants in their study were much older (42.2 years) as compared to those in this current study (26.7 years). It has been found that age affects the performance of the FMS™ with younger individuals performing the FMS™ better.[35] Coincidently, an FMS™ score of 14 or lower is associated with a higher risk of musculoskeletal injury among competitive athletes.[12–14] Furthermore, the result of each FMS™ test component revealed that the LBP group appeared to have more deficits in movement performance on the deep squat test and the trunk stability push-up test than the asymptomatic group. These findings are not surprising because both deep squat and push-up demand spinal stability and motor control of the core musculature and participants with LBP likely had deficits in the spinal stabilizers.[36] However, Ko et al. used the original scoring system, which could have under-scored the FMS™ performance for those who had existing LBP, as presence of pain at baseline was given a zero score. Using the modified scoring system in our study, participants who performed movement correctly were awarded a score that was more representation of their quality of movement. It was necessary to make the modification as the FMS™ was designed to assess quality of movement.

This study has some limitations. One limitation was that the participants in this study were young and the intensity of their LBP was low. Therefore, the results might not be generalized to the other ranges of age and to individuals with higher intensities of LBP. Although the two raters who viewed and scored the FMS™ performance four days later after they recorded the real-time performance for inter-rater reliability, and viewed and scored again six weeks later, it is not certain if the two raters could recall the performance.

CONCLUSION

These results of the current study indicate good-to-excellent intra-rated and inter-rater reliability of the FMS™ with a modified scoring system when the scores were collected in both real-time and video-recorded sessions from the young adults with and without LBP. In addition, the modified FMS™ scoring system allows clinicians to quantify the quality of movement in young adults with LBP and identify restrictions and limitations to common body movement patterns requiring minimal time and financial costs. Identification of such factors may allow therapists to address movement impairments in their plan of care. For future study, researchers should assess the reliability and validity of the FMS™ in individuals with older ages and with varying stages of LBP.

REFERENCES

1. Freburger JK, Holmes GM, Agans RP, et al. The rising prevalence of chronic low back pain. *Arch Intern Med*. 2009;169(3):251-258. doi:10.1001/archinternmed.2008.543

2. Waterman BR, Belmont PJ Jr, Schoenfeld AJ. Low back pain in the United States: Incidence and risk factors for presentation in the emergency setting. *Spine J*. 2012;12(1):63-70. doi:10.1016/j.spinee.2011.09.002

3. Lehtola V, Luomajoki H, Leinonen V, Gibbons S, Airaksinen O. Efficacy of movement control exercises versus general exercises on recurrent sub-acute nonspecific low back pain in a sub-group of patients with movement control dysfunction. Protocol of a randomized controlled trial. *BMC Musculoskelet Disord*. 2012;13(1):55. doi:10.1186/1471-2474-13-55

4. Pengel LH, Herbert RD, Maher CG, et al. Acute low back pain: Systematic review of its prognosis. *BMJ*. 2003;327(7410):323. doi:10.1136/bmj.327.7410.323

5. Lamoth CJ, Meijer OG, Daffertshofer A, Wuisman PIJM, Beek PJ. Effects of chronic low back pain on trunk coordination and back muscle activity during walking: Changes in motor control. *Eur Spine J*. 2006;15(1):23-40. doi:10.1007/s00586-004-0825-y

6. Biely SA, Silfies SP, Smith SS, Hicks GE. Clinical observation of standing trunk movements: What do the aberrant movement patterns tell us? *J Orthop Sports Phys Ther*. 2014;44(4):262-272. doi:10.2519/jospt.2014.4988

7. Panjabi MM. The stabilizing system of the spine. part I. function, dysfunction, adaptation, and enhancement. *J Spinal Disord*. 1992;5:383.

8. Fritz JM, Piva SR, Childs JD. Accuracy of the clinical examination to predict radiographic instability of the lumbar spine. *Eur Spine J*. 2005;14(8):743-750. doi:10.1007/s00586-004-0803-4

9. Teyhen DS, Flynn TW, Childs JD, et al. Fluoroscopic video to identify aberrant lumbar motion. *Spine*. 2007;32(7):E229.

10. Cook G, Burton L, Hoogenboom BJ, et al. Functional movement screening: The use of fundamental movements as an assessment of function - part 1. *Int J Sports Phys Ther*. 2014;9(3):396.

11. Cook G, Burton L, Hoogenboom BJ, et al. Functional movement screening: The use of fundamental movements as an assessment of function - part 2. *Int J Sports Phys Ther*. 2014;9(4):549.

12. Butler RJ, Contreras M, Burton LC, et al. Modifiable risk factors predict injuries in firefighters during training academies. *Work*. 2013;46(1):11-17.

13. Chorba RS, Chorba DJ, Bouillon LE, et al. Use of a functional movement screening tool to determine injury risk in female collegiate athletes. *N Am J Sports Phys Ther*. 2010;5(2):47.

14. Kiesel K, Plisky P, Butler R. Functional movement test scores improve following a standardized off-season intervention program in professional football players. *Scand J Med Sci Sports*. 2011;21(2):287-292. doi:10.1111/j.1600-0838.2009.01038.x

15. Clay H, Mansell J, Tierney R. Association between rowing injuries and the functional movement screen[TM] in female collegiate division I rowers. *Int J Sports Phys Ther*. 2016;11(3):345.

16. Ko M-J, Noh K-H, Kang M-H, Oh J-S. Differences in performance on the functional movement screen between chronic low back pain patients and healthy control subjects. *J Phys Ther Sci*. 2016;28(7):2094-2096. doi:10.1589/jpts.28.2094

17. Bonazza NA, Smuin D, Onks CA, Silvis ML, Dhawan A. Reliability, validity, and injury predictive value of the functional movement screen: A systematic review and meta-analysis. *Am J Sports Med*. 2017;45(3):725-732. doi:10.1177/0363546516641937

18. Cuchna JW, Hoch MC, Hoch JM. The interrater and intrarater reliability of the functional movement screen: A systematic review with meta-analysis. *Phys Ther Sport*. 2016;19:57-65. doi:10.1016/j.ptsp.2015.12.002

19. Moran RW, Schneiders AG, Major KM, Sullivan SJ. How reliable are functional movement screening scores? A systematic review of rater reliability. *Br J Sports Med*. 2016;50(9):527-536. doi:10.1136/bjsports-2015-094913

20. Teyhen DS, Shaffer SW, Lorenson CL, et al. The functional movement screen: A reliability study. *J Orthop Sports Phys Ther*. 2012;42(6):530-540. doi:10.2519/jospt.2012.3838

21. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*. 2009;41(4):1149-1160. doi:10.3758/brm.41.4.1149

22. Koumantakis GA, Watson PJ, Oldham JA. Trunk muscle stabilization training plus general exercise versus general exercise only: Randomized controlled trial of patients with recurrent low back pain. *Phys Ther*. 2005;85(3):209-225. doi:10.1093/ptj/85.3.209

23. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. *Spine*. 2005;30(11):1331-1334. doi:10.1097/01.brs.0000164099.92112.29

24. Davidson M, Keating JL. A comparison of five low back disability questionnaires: Reliability and responsiveness. *Phys Ther*. 2002;82(1):8-24. doi:10.1093/ptj/82.1.8

25. Fritz JM, Irrgang JJ. A comparison of a modified Oswestry low back pain disability questionnaire and the Quebec back pain disability scale. *Phys Ther*. 2001;81(2):776-788. doi:10.1093/ptj/81.2.776

26. Vianin M. Psychometric properties and clinical usefulness of the Oswestry disability index. *J Chiropr Med*. 2008;7(4):161-163. doi:10.1016/j.jcm.2008.07.001

27. Parenteau-G E, Gaudreault N, Chambers S, et al. Functional movement screen test: A reliable screening test for young elite ice hockey players. *Phys Ther Sport*. 2014;15(3):169-175. doi:10.1016/j.ptsp.2013.10.001

28. Portney LG. *Foundations of Clinical Research: Applications to Practice 3th*. Third edition.; 2009.

29. Onate JA, Dewey T, Kollock RO, et al. Real-time intersession and interrater reliability of the functional movement screen. *J Strength Cond Res*. 2012;26(2):408-415. doi:10.1519/jsc.0b013e318220e6fa

30. Schneiders AG, Davidsson Å, Hörman E, et al. Functional movement screen^TM normative values in a young, active population. *Int J Sports Phys Ther*. 2011;6(2):75.

31. Saito Y, Sozu T, Hamada C, Yoshimura I. Effective number of subjects and number of raters for inter-rater reliability studies. *Stat Med*. 2006;25(9):1547-1560. doi:10.1002/sim.2294

32. Leeder JE, Horsley IG, Herrington LC. The inter-rater reliability of the functional movement screen within an athletic population using untrained raters. *J Strength Cond Res*. 2016;30(9):2591-2599. doi:10.1519/jsc.0b013e3182a1ff1d

33. Gulgin H, Hoogenboom B. The functional movement screening (FMS)^TM: An inter‐rater reliability study between raters of varied experience. *Int J Sports Phys Ther*. 2014;9(1):14.

34. Gribble PA, Brigle J, Pietrosimone BG, Pfile KR, Webster KA. Intrarater reliability of the functional movement screen. *J Strength Cond Res*. 2013;27(4):978-981. doi:10.1519/jsc.0b013e31825c32a8

35. Teyhen DS, Riebel MA, McArthur DR, et al. Normative data and the influence of age and gender on power, balance, flexibility, and functional movement in healthy service members. *Mil Med*. 2014;179(4):413-420. doi:10.7205/milmed-d-13-00362

36. Jeong U-C, Sim J-H, Kim C-Y, Hwang-Bo G, Nam C-W. The effects of gluteus muscle strengthening exercise and lumbar stabilization exercise on lumbar muscle strength and balance in chronic low back pain patients. *J Phys Ther Sci*. 2015;27(12):3813-3816. doi:10.1589/jpts.27.3813