The Functional Movement Screen: A Review

Chris Beardsley, MA Hons[1] and Bret Contreras, BSc, CSCS[2]
[1]Strength and Conditioning Research Limited, Loughborough, Leicestershire, United Kingdom; [2]School of Sport and Recreation, Auckland University of Technology, Auckland, New Zealand

ABSTRACT

THE FUNCTIONAL MOVEMENT SCREEN (FMS) IS A PRE-PARTICIPATION SCREENING TOOL COMPRISING 7 INDIVIDUAL TESTS FOR WHICH BOTH INDIVIDUAL SCORES AND AN OVERALL SCORE ARE GIVEN. THE FMS DISPLAYS BOTH INTERRATER AND INTRARATER RELIABILITY BUT HAS BEEN CHALLENGED ON THE BASIS OF A LACK OF VALIDITY IN SEVERAL RESPECTS. THE FMS SEEMS TO HAVE SOME DEGREE OF PREDICTIVE ABILITY FOR IDENTIFYING ATHLETES WHO ARE AT AN INCREASED RISK OF INJURY. HOWEVER, A POOR SCORE ON THE FMS DOES NOT PRECLUDE ATHLETES FROM COMPETING AT THE HIGHEST LEVEL NOR DOES IT DIFFERENTIATE BETWEEN ATHLETES OF DIFFERING ABILITIES.

The functional movement screen (FMS) is a pre-participation screening tool comprising 7 individual tests for which both individual scores and an overall score are given (11). The 7 tests are rated from 0 to 3 by an examiner and include the deep squat, hurdle step, in-line lunge, shoulder mobility, active straight leg raise, trunk stability push-up, and rotary stability (11,12). The score of 0 is given if pain occurs during a test, the score of 1 is given if the subject is not able to perform the movement, the score of 2 is given if the subject is able to complete the movement but compensates in some way, and the score of 3 is given if the subject performs the movement correctly (11).

It has been suggested that a less-than-perfect score on a single individual test of the FMS reveals a "compensatory movement pattern." Such compensatory movement patterns have been proposed to lead to athletes "sacrificing efficient movements for inefficient ones" (11), which implies the replacement of either a more economical or more effective pattern with a less economical or less effective one. It has also been proposed that such compensatory movement patterns predispose an athlete to injury and reduced performance and may be corrected by performing specific exercises. As a designer of the FMS states: "an athlete who is unable to perform a movement correctly … has uncovered a significant piece of information that may be the key to reducing the risk of chronic injuries, improving overall sport performance, and developing a training or rehabilitation program …" (9). This seems to imply that the FMS is put forward as a valid test for identifying certain movement patterns that lead to greater injury risk and reduced athletic performance. In the course of our review, we did not identify a formal definition of the concept "compensatory movement pattern." We suggest that it can be defined as a kinematic feature or sequence of features observed during the performance of a movement that deviate from a template that is thought to represent the least injurious way of performing the movement.

In the FMS, the individual scores for each movement are combined into a final score out of 21 total possible points. It has been suggested that lower overall scores predict individuals who are at a greater risk of injury than those with higher scores (11). In practice, researchers have generally identified 14 points as the ideal cut-off point for those at greater or less risk of injury (5,8,36,38,41,43). The cut-off value of 14 points was in certain studies identified by means of a statistical method known as a receiver-operator characteristic (ROC) curve (5,38,43). This technique allows researchers to identify the numerical score that maximizes the correct prediction of injury classification (66). However, in other cases (8,36), the researchers simply adopted the cut-off value of 14 points based on the findings of previous studies. Although this may not maximize the predictability of the cut-off point in those individual studies that elected not to use a ROC curve, it does have the advantage of enhancing comparability between trials. Studies investigating the norms for FMS overall scores have identified that the normal FMS score in healthy but untrained populations ranges from 14.14 ± 2.85 points (51) to 15.7 ± 1.9 points (53). This suggests that most untrained people are slightly above the cut-off score of ≤14 points, which is thought to be indicative of prevalent compensation patterns and which is also believed to be predictive of increased risk of injury and reduced performance.

KEY WORDS:
functional movement screen; pre-participation screen; biomechanics; injury prevention; injury risk

 Copyright © National Strength and Conditioning Association

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

The FMS was designed in 1997 by Cook et al. (10). It seems that most if not all of the individual tests within the FMS were selected based mainly on the clinical experiences of the designers, which included orthopedic assessments of athletes with injuries (9). The FMS was first presented commercially as a manual for screening athletes and the product line was later expanded to include a range of equipment for performing the screen, certifications for those screening athletes, seminars, books, and videos. The commercial marketing of the screen, along with a lack of research to support the tests that are included within it, has led to some controversy regarding its effectiveness. In recent years, many studies have been published exploring various aspects of the FMS, including its validity, reliability, its ability to predict the likelihood of injury and suboptimal performance, and the factors that are correlated with overall FMS score. It is the purpose of this review to provide a summary of the current literature. The review has been based on a literature search conducted on December 5, 2013, in the following databases: PubMed, Google Scholar, and Scopus, and using the key terms "functional movement screen" and "FMS." The references were screened by both reviewers based on the titles and abstracts, and further studies were retrieved from the studies cited within the full articles accessed in the initial search. Any peer-reviewed study involving the FMS was incorporated into the review.

RELIABILITY OF THE FUNCTIONAL MOVEMENT SCREEN

For a test to be usable, it needs to be reliable. Reliability describes whether a test can be repeated either by the same person at a slightly different time (intrarater) or by different people at the same time (interrater) and produce the same result (34). At least 11 peer-reviewed studies have assessed either the interrater or intrarater reliability of the sum FMS score (6,17,21,28,42,45,48,53,55,57,59) as shown in Table 1. All of these 11 studies reported on the interrater reliability of the sum FMS score while only 7 reported on intrarater reliability of the sum FMS score (21,28,45,48,55,57,59). Of the 11 studies investigating interrater reliability of the sum FMS score, only 1 study reported that reliability was less than moderate (55). The 11 studies investigated reliability using 3 different statistical methods. Most commonly, the intraclass correlation coefficient was used (63), although Kappa (60) was used twice and the Krippendorff alpha (K alpha) (31) was used once. It is noteworthy that the one study reporting less than moderate reliability (55) used a different statistical method of analyzing correlation (29) and also used raters with quite diverse backgrounds. Different statistical methods might be expected to produce different outcomes (either better or worse) while the use of raters with diverse backgrounds might reasonably be expected to lead to substantially worse reliability. Of the 7 studies reporting on intrarater reliability, 6 studies reported at least moderate reliability (21,45,48,55,57,59) while 1 study found that student raters (but not more experienced raters) demonstrated poor reliability (28). The FMS seems to display an acceptable degree of reliability for a field test in most populations and with most types of raters.

VALIDITY OF THE FUNCTIONAL MOVEMENT SCREEN

In addition to reliability, for a test to be useful, it also needs to be valid. Validity describes whether a test actually measures what it is intended to measure (35). Validity is often discussed under various headings, including construct validity, face validity, content validity, and criterion-referenced validity (30). Construct validity is the more exact term for overall validity or the ability of the test to reflect the underlying theoretical basis upon which outcomes of the test are interpreted (30). Face validity, content validity, and criterion-referenced validity are all subordinate elements of construct validity and provide evidence for it (30). Face validity is a measure of the degree to which a test is perceived by either the subject or casual observers to measure what it sets out to measure (30). Face validity is unusual in that it can be beneficial or harmful to a test. Where maximal performance is required, a high degree of face validity can be beneficial, as it helps to ensure that athletes exert themselves to the greatest extent. However, where a test is performed on a movement under controlled conditions in an attempt to gain an insight by proxy into the same or similar movements that is performed under noncontrolled conditions (as in the FMS), a high degree of face validity is harmful. In this instance, a high degree of face validity means that the subjects can manipulate their performances and alter the outcome of the test. Content validity is defined as the extent to which a test contains all the necessary elements to achieve its stated goals, as assessed by expert opinion (30). Criterion-referenced validity is the extent to which outcomes of the test are correlated with other measures of the same underlying quality and is often subdivided into 4 further subtypes of validity: concurrent, convergent, predictive, and discriminant. Concurrent criterion-referenced validity describes the extent to which a test is correlated with other similar tests (30). Convergent criterion-referenced validity describes the extent to which a test is correlated with other gold standard tests (30). Predictive criterion-referenced validity describes the extent to which a test is able to predict future events that are associated with the outcome measure (30). For example, a test involving a countermovement jump would be expected to be acceptably predictive of sprint running ability (33). Discriminant criterion-referenced validity describes the extent to which a test is not correlated with other dissimilar tests (30). For example, a test with good criterion-referenced discriminant validity for countermovement jump height would not be well correlated with a flexibility test, as the 2 tests are measuring inherently different underlying constructs (lower-body power and flexibility). It is remarkable that, despite the great interest in the FMS and significant uptake by the strength and conditioning community of the test, very little research has been performed into the construct validity of

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

Table 1 Interrater and intrarater reliability of the functional movement screen (6,17,21,28,42,45,48,53,57,59)				
Study	Method	Interrater	Method	Intrarater
Minick et al. (42)	Kappa	Substantial-excellent	NA	No test performed
Schneiders et al. (53)	Kappa	Excellent	NA	No test performed
Frohm et al. (21)	ICC	Good	ICC	Good
Onate et al. (45)	ICC	Good	ICC	Good
Teyhen et al. (59)	ICC	Moderate-good	ICC	Moderate-good
Butler et al. (6)	ICC	Good	NA	No test performed
Shultz et al. (55)	K alpha	Poor	ICC	Moderate
Smith et al. (57)	ICC	Good	ICC	Good
Gribble et al. (28)	ICC	Moderate-good	ICC	Varied (poor to good)
Elias (17)	ICC	Good	NA	No test performed
Parenteau-G et al. (48)	ICC	Good	ICC	Good

ICC = intraclass correlation coefficient; NA = not applicable.

the FMS, either directly or by assessing one of the subordinate types of construct validity (face validity, content validity, and criterion-referenced validity).

To assess the overall construct validity of the FMS, a very clear definition of the object of measurement is needed. In the case of many screening tools, the object of measurement is very clearly stated. In a comprehensive review, Dallinga et al. (14) detailed a number of screens that are predictive of lower-body injury. One screen that the reviewers concluded was capable of identifying lower-body injury is the Beighton's test, which assesses the extent of generalized joint laxity by measuring average joint hypermobility over several different joints (14). In contrast, the object of measurement of the FMS seems less easy to define. The designers of the test have stated that the purpose of the FMS is to "determine whether the athlete has the essential movements needed to participate in sports activities with a decreased risk of injury" (11) and that athletes who score poorly on an individual test are "utilizing compensatory movement patterns during their activities" (11). It therefore seems that the object of measurement of the FMS is intended to be compensatory movement patterns as they are performed during sporting activities. Indeed, the designers of the test state that the FMS movements "tie all sports together because they are fundamental and representative of human movement" (9). This also suggests that the movement patterns tested in the FMS are similar to or the same as those in sporting movements. Researchers investigating the FMS have made the same assumption. For example, Chorba et al. (8) state that the "the tool was designed to challenge the interactions of kinetic chain mobility and stability necessary for performance of fundamental, functional movement patterns. Such movements require controlled neuromuscular execution in a variety of occupational and athletic tasks." This is a key assumption for assessing construct validity; if the FMS is found to identify compensatory movement patterns in movements that are dissimilar from those performed in sport, then it is unclear how identifying such compensatory movement patterns can validly relate to the prediction of injury or performance during sporting activities. After all, sporting activity movement patterns are the real subject of interest, being the environment in which athletes are injured and in which athletes perform.

The FMS has been found to display poor construct validity when assessed directly rather than by one of the subtypes of construct validity (face validity, content validity, and criterion-referenced validity). Frost et al. (24) reported that similar exercises to those used in the FMS tests displayed different movement characteristics when performed faster or under greater load. For this study, Frost et al. (24) recruited 52 professional male firefighters and asked them to perform 5 low-load and low-movement speed, whole-body movements. The subjects were then asked to perform the same movements but modified in 3 ways: with increased movement speed, with increased external load, and with increased movement speed combined with increased external load. The 5 tasks were as follows: (a) a lift from the ground to waist height, (b) a bodyweight squat to self-selected depth, (c) a lunge, (d) a cable press with the right arm while the left leg was forward, and (e) a cable pull with the right arm while the left leg was forward. Increasing both load and speed led to changes in variables in

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

each movement but the changes were not always similar across the 5 selected movements nor were they similar between load and speed. For example, when performing the lifting task with a heavier load, subjects used a more upright trunk posture. However, when using a faster movement speed for the same lifting pattern, the subjects used a less upright posture and shifted to a hip-dominant pattern. Similar effects of load and speed were found for the squat. These findings suggest that any compensatory movement patterns displayed by athletes during the FMS tests might differ from those displayed during sporting movement at sports-specific velocities.

The FMS has been found to display high and (in the context of the test) undesirable levels of face validity. Frost et al. (23) found that knowledge of the FMS test criteria was able to affect the outcome of the test significantly, which may imply that performance of the test can be influenced by the athlete irrespective of the underlying compensation patterns that may or may not be present. Frost et al. (23) administered the FMS to a group of professional firefighters without first providing the subjects with information about the objectives of the test and without giving any feedback during the test. After 3 minutes, the researchers asked the subjects to perform the FMS a second time. However, before the subjects performed the FMS for a second time, the criteria used to grade each of the 7 individual components of the FMS were described to them. The mean overall FMS score increased significantly from 14.1 ± 1.8 points to 16.7 ± 1.9 points after the subjects were provided with knowledge of the scoring criteria. This suggests that there is likely to be a discrepancy between the FMS test movement patterns as they are performed in front of a rater and during sporting movement, when the criteria are not being assessed. It also suggests that participants are capable of influencing their FMS score deliberately, either consciously or subconsciously.

Although no studies have explored the concurrent, convergent, or discriminant criterion-referenced validity of the FMS, one study has explored the relationship between the individual components of the overall FMS test, which can be described as a measure of latent validity. Ben Kazman et al. (1) investigated the factor structure of the FMS test because the FMS overall score involves the summation of the individual test scores. According to classical test theory, when a score is derived from multiple tests, the overall score is assumed to be a test of the same latent variable (1). Latent variables are assumed to have the same properties as measurements of the tested variables. If they do not have the same properties, then the meanings of both the overall score and individual tests are deemed to be ambiguous and are therefore said to lack construct validity. One way of assessing whether this is the case is to measure the correlation between the individual test results. Ben Kazman et al. (1) reported that there was poor correlation between the outcomes of the various tests, meaning that the validity of combining the scores of each test to produce a sum score to predict injury risk is questionable.

Many studies have explored the predictive criterion-referenced validity of the FMS by investigating its ability to predict injury risk. The detailed results of these trials are set out in the following section. However, it is important to note that although there is some evidence that the FMS has a degree of predictive capability, it has very poor sensitivity scores. Studies that have measured the sensitivity of the FMS using a cut-off of 14 points have reported that sensitivity ranges widely from 16 to 84% (median 54%), meaning that the FMS is able to detect correctly around 54% of people who are likely to become injured. This low median sensitivity is problematic for the overall predictive criterion-referenced validity of the FMS as it indicates that it is only marginally better than a coin toss for identifying individuals who are genuinely at risk of injury.

CAPABILITY OF THE FUNCTIONAL MOVEMENT SCREEN TO PREDICT INJURY RISK

One of the original principles of the FMS is that it measures the prevalence of compensation patterns that are thought to be injurious (9). At least 8 peer-reviewed studies have assessed whether the overall FMS score can predict the incidence of injury (5,8,36,38,41,43,50,54). Within these 8 studies, 5 unique studies have compared individuals achieving an FMS score of ≤14 points with individuals achieving an FMS score of >14 points (5,8,36,38,43). Two further trials have compared individuals achieving an FMS score of ≤17, with individuals achieving an FMS score of >17 (50,54). The cut-off value of 14 points was in certain studies identified by means of a statistical method known as a ROC curve (5,38,43). This technique allows researchers to identify the numerical score that maximizes the correct prediction of injury classification (66). Not all studies that performed this statistical method found that the cut-off value of 14 points maximized the correct prediction of injury classification (54). In some studies (8,36), the researchers simply adopted the cut-off value of 14 points based on the findings of previous studies. Although this may not maximize the predictability of the cut-off point in those individual studies that elected not to use a ROC curve, it does have the advantage of enhancing comparability between trials.

Each study exploring the capability of the FMS to predict injury risk provided either odds ratios or relative risks or both. Odds ratios provide a measure of the probability of an event occurring relative to the probability that it does not occur (2). Relative risk is the multiple of the risk of an outcome occurring in one group as a result of an exposure compared with the risk of the same outcome in another group (65). By dividing the subjects into the groups based on a cut-off point, and

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

because a risk of injury could be calculated for each group, the researchers were able to calculate odds ratios and also measures of relative risk. In the 5 studies that used an FMS score of ≤14 points as a cut-off, the odds ratio was between 2.00 and 11.67 and the relative risk was between 1.55 and 4.20, as shown in Table 2. These findings indicate that an overall cut-off point of ≤14 points FMS may well differentiate between individuals who are at a greater or lesser risk of injury. However, it is also important to note that the FMS may not be able to identify individuals who have previously suffered significant traumatic lower-body injury. For example, Chorba et al. (8) found that the FMS is not able to differentiate between individuals with and without prior anterior cruciate ligament (ACL) reconstruction. Because prior injury is a significant risk factor for various lower-body noncontact injuries, including hamstring strains (4) and ACL rupture (18,61), this is an important deficiency of the test. Moreover, it is noteworthy that ACL injury is known to alter movement patterns (58), and therefore, it is even more surprising that the FMS was not able to detect this difference between injured and noninjured groups. However, many tests that can identify individuals who have suffered an ACL injury are performed under situations of high demand, such as drop jumps (52), and because Frost et al. (24) have found that movement patterns change with load and speed, it could be that this failure of the FMS to detect differences resulting from ACL injury results from the differences in demand between the tests.

SENSITIVITY OF THE FUNCTIONAL MOVEMENT SCREEN

Predictive value is just one way of measuring the effectiveness of a screen or test (19). Another key measure is sensitivity. Sensitivity refers to the percentage of subjects who do actually have the target condition (e.g., compensatory movement patterns, as in the FMS) and who correctly give positive test results (19). Sensitivity is therefore sometimes called a test of "true positives" because it rates a test solely based on its ability to flag positive results. Where sensitivity is high or perfect (i.e., 100%), there is little chance of failing to detect the target condition. However, high or perfect sensitivity is also a feature of a counterfeit test that is designed to always give a positive result, irrespective of whether the subject actually has the condition or not. A test with good sensitivity can still therefore be very poor at clearing subjects who do not have the condition. In studies that have measured the sensitivity of the FMS using a cut-off of 14 points, the results range widely from 16 to 84% (median 54%), meaning that the FMS is able to detect correctly between 16 and 84% (median 54%) of people who are likely to become injured (5,43). This low median sensitivity (being close to 50%) is considered problematic: researchers working in the area have commented that sensitivity >50% is a key for identifying those at risk of injury (54). Indeed, it is important to note that a test that has sensitivity close to 50% is in reality no better than a coin toss in this respect. These findings suggest that the FMS using 14 points as a cut-off is relatively poor at detecting individuals who actually have compensation patterns. This indicates that the FMS test may risk missing individuals who are actually at risk of injury, if compensation patterns are indeed a risk factor for injury.

Table 2
Ability of the functional movement screen to predict injury risk (5,8,38,43,50,54)

	Cut-off	Relative risk	Odds ratio	Sensitivity	Specificity
Kiesel et al. (38)	14	4.20	11.67	0.54	0.91
Butler et al. (5)	14	Not provided	8.31	0.84	0.62
Chorba et al. (8)	14	1.89	3.85	0.58	0.74
Kiesel et al. (36)	14	1.87	2.33	0.27	0.87
O'Connor et al. (43)	14	1.55	2.00	0.16	0.92
Shojaedin et al. (54)	17	1.81	4.70	0.65	0.78
Peate et al. (50)	17	1.29	1.43	0.36	0.71

SPECIFICITY OF THE FUNCTIONAL MOVEMENT SCREEN

Predictive value is just one way of measuring the effectiveness of a screen or test (19). Another key measure is specificity. Specificity refers to the percentage of subjects who genuinely do not have the target condition (e.g., compensatory movement patterns, as in the FMS) who correctly give negative test results (19). Specificity is therefore sometimes called a test of "true negatives" because it rates a test solely based on its ability to flag negative results. Where specificity is high or perfect (i.e., 100%), there is little chance of failing to clear a subject who does not have the condition. However, high or perfect specificity is also a feature of a counterfeit test that is designed to always give a negative result, irrespective of whether the subject actually has the condition or not. A test with good specificity can therefore still be very poor at identifying subjects who actually have the condition. In studies that have measured the specificity of the FMS using a cut-off of 14 points, the results range from 62 to 92% (median 87%), meaning that the FMS is able to avoid falsely giving a positive test rest to between 62 and 92% (median 87%) of people who take the test (5,43). This means that between 8 and 38% (median 13%) of people taking the FMS test and using 14 points as a cut-off will be falsely informed that they are at higher risk of injury when they are not.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

POSITIVE AND NEGATIVE PREDICTIVE VALUES

Taking the sensitivity and specificity information one step further can be done by using positive predictive value (PPV) and negative predictive value (NPV). The PPV is the probability that an individual with a positive screening result does in fact have the condition that is being tested for. The PPV is derived from the sensitivity but also takes the overall prevalence of the condition into account. The NPV is the probability that an individual with a negative screening result does not in fact have the condition that is being tested for. The NPV is derived from the specificity but also takes the overall prevalence of the condition into account. Unfortunately, only 1 study (5) reported the PPV and NPV of the FMS test, noting that the values were 86 and 58%, respectively. These findings are in contrast to the above pool of results based on 6 studies, as they suggest that the FMS is better at identifying potentially compensatory movement patterns in individuals who have them than it is at clearing individuals who do not have them.

CAPABILITY OF THE FUNCTIONAL MOVEMENT SCREEN TO PREDICT ATHLETIC PERFORMANCE

The FMS is thought to measure the prevalence of potentially compensatory movement patterns that are believed to be inefficient (9). It is therefore feasible that the FMS could be predictive of a reduced level of performance. At least 5 studies (7,20,29,44,47) have assessed correlations between athletic performance and FMS sum scores and none found any correlation between athletic performance and FMS score. However, it is interesting that Chapman et al. (7) found that athletes with a FMS score of >14 points had a significantly better change in athletic performance from one season to the next ($+0.41 \pm 2.50\%$) compared with those athletes who had an FMS score of <14 points ($-0.51 \pm 2.30\%$). In this study, athletic performance was assessed by reference to the individual athletes' best time or mark in their primary event in each season. Whether this was because of the ability of the FMS to predict injury risk and it was injury that led to the reduced performance increment or whether this is reflective of some other predictive ability of the FMS is unclear. Moreover, we cannot infer any direct causation from the correlations observed in this study.

ABILITY OF EXERCISE TRAINING TO IMPROVE FUNCTIONAL MOVEMENT SCREEN SCORES

It is thought that exercise training can be performed on the basis of an FMS test, which may result in a reduced risk of injury (9). Indeed, if the FMS is to be useful, it is important that action can be taken based on the results of a baseline test that then improves the score on a subsequent test. At least 7 studies have assessed the ability of different exercise intervention programs to improve the FMS score in various populations (3,13,22,27,37,39,46). Of these 7 studies, 6 reported that exercise of some kind was able to improve FMS score. Only Frost et al. (22) reported that exercise did not improve FMS score in relation to a control group. However, in the 2 studies that compared corrective or functional exercise with traditional resistance training (22,46), the researchers found in both cases that there was no significant difference in the improvement in FMS score between the 2 approaches. This is problematic, as part of the rationale for performing the FMS is to identify potentially corrective exercises that improve FMS score and thereby lead to a reduced injury risk and improved performance. Moreover, it is important to note that no studies have yet been performed that have tested whether improving an FMS score, whether by specific corrective exercises, functional exercises, or by traditional resistance-training, does in fact lead to a reduced risk of injury.

EXTENT TO WHICH BODY MASS INDEX AND DEVELOPMENT AFFECT FUNCTIONAL MOVEMENT SCREEN SCORE

Higher body mass index (BMI) seems to be associated with a lower overall FMS score. At least 4 studies have reported on the effect of BMI on FMS score (15,16,41,51). All these studies found that higher BMI scores are associated with lower FMS scores. Exactly why there is an inverse correlation of this kind is unclear. It may be the case that high BMI measurements are acting as a proxy for levels of physical activity and exercise participation. Both Duncan and Stanley (15) and Perry and Koehle (51) found that FMS scores are positively related to exercise participation and physical activity; it may be the case that the inverse correlation between BMI and FMS scores could be mediated by the tendency for overweight/obese individuals to perform less physical activity or to participate in less exercise. However, it may be the case that obesity affects the biomechanics of normal movement. Several researchers have reviewed the literature in this respect (32,62) and noted that there are specific biomechanical differences between individuals with low and high BMI values. For example, obese children tend to have different foot structure from normal weight children, display poorer performance in movements involving bodyweight (such as jumps), walk with greater gait asymmetry, and have longer stance times during the gait cycle in conjunction with overall slower self-selected walking speeds (32). Also, more pertinently for the FMS, researchers have reported that obese individuals display a different movement pattern during sit-to-stand movements. Sibella et al. (56) found that obese individuals perform a much more knee-dominant sit-to-stand movement characterized by a more upright torso and a less vertical tibia. This led to the obese subjects displaying a greater knee extension moment and a smaller hip extension moment than the normal-weight individuals tested. Galli et al. (25) reported similar results. In respect of other movements, Gilleard and Smith (26) observed differences in both kinetics and kinematics between obese and normal-weight individuals even during a simple standing working task. Xu et al.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

(64) observed differences in trunk kinematics between obese and normal-weight individuals during a series of lifting tasks. Such biomechanical differences between obese and normal-weight individuals may be caused by differences in the relative sizes and weights of the various body segments and by differences in respect of the relative strength of the musculature in relation to those segments. Additionally, Paszkewicz et al. (49) reported that the stage of development also affected the outcome of the FMS, which may imply that the test is not applicable to adolescent athletes or may at least be expected to display a different normal score in adolescent populations in comparison with adults.

DIRECTIONS FOR FUTURE RESEARCH

As noted above, it is remarkable that, despite the great interest in the FMS and significant uptake by the strength and conditioning community of the test, very little research has been performed into the construct validity of the FMS. Little research has been performed either directly or by assessing one of the subordinate types of construct validity (which include construct validity, face validity, content validity, and criterion-referenced validity). Moreover, no research has been performed in respect of content validity and criterion-referenced validity, with the exception of predictive criterion-referenced validity. Yet, what little research has been performed indicates that there the FMS may in fact lack sufficient construct validity to be a usable test by strength and conditioning professionals. We therefore suggest that rather than focus on performing additional trials assessing predictive criterion-referenced validity, future research should focus on further assessments of face validity, as well as first assessments of content validity and those areas within criterion-referenced validity that have not been researched (i.e., concurrent, convergent, and discriminant criterion-referenced validity). Given the current widespread use of the FMS, such research would seem to be required urgently.

CONCLUSIONS

In summary, the FMS is a preparticipation screening tool comprising 7 individual tests for which both individual scores and an overall score are given. The overall test seems to be reliable both interrater and intrarater. However, the test has recently been subject to strong challenge. In fact, it is possible to suggest that the FMS may in fact lack sufficient construct validity to be a usable test by strength and conditioning professionals. Although research has not been performed into each main subordinate type of construct validity, the FMS has been found to display high and (in the context of the test) undesirable levels of face validity (23), poor construct validity when assessed directly (24), poor predictive criterion-referenced validity when assessed by reference to sensitivity when using a cut-off of 14 points (the median sensitivity of 6 studies is just 54%, which is very close to the desirable threshold of 50%), and poor latent validity (1). Despite these very serious problems with construct validity, the FMS seems to have some degree of predictive ability for identifying athletes who are at an increased risk of injury. Exactly how this is possible is unclear. Moreover, despite the apparently good predictive capability, sensitivity is poor (median: 54%). Specificity is better (median: 87%), suggesting that the FMS may be better at clearing individuals who do not have potentially compensatory movement patterns than it is in detecting individuals who actually have them. Additionally, contrary to claims, the FMS seems to have no relation to athletic ability between individuals. A poor score on the FMS does not preclude athletes from competing at the highest level nor does it differentiate between athletes of differing abilities. However, a high BMI seems to be correlated with poor performance on the FMS. Because FMS scores are positively related to exercise participation and physical activity, it may be the case that the inverse correlation between BMI and FMS scores could be mediated by the tendency for overweight/obese individuals to perform less physical activity or to participate in less exercise. However, it may be the case that obesity affects the biomechanics of normal movement. Finally, it seems likely that greater levels of exercise, training, and physical activity (which may also lead to lower BMI) can improve FMS scores. However, corrective exercise or functional training does not seem to be any better than traditional resistance training methods at improving FMS scores over a period.

Conflicts of Interest and Source of Funding: The authors report no conflicts of interest and no source of funding.

Chris Beardsley *is the Director of Strength and Conditioning Research Limited.*

Bret Contreras *is currently pursuing his PhD at AUT University.*

REFERENCES

1. Ben Kazman J, Galecki J, Lisman P, Deuster PA, and O'Connor FG. Factor structure of the functional movement screen in marine officer candidates. *J Strength Cond Res* 28: 672–678, 2014.
2. Bland JM and Altman DG. Statistics notes: The odds ratio. *BMJ* 320: 1468, 2000.
3. Bodden JG, Needham RA, and Chockalingam N. The effect of an intervention program on functional movement screen test scores in mixed martial arts athletes. *J Strength Cond Res*, 2013. Epub ahead of print.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

4. Brockett CL, Morgan DL, and Proske U. Predicting hamstring strain injury in elite athletes. *Med Sci Sports Exerc* 36: 379, 2004.
5. Butler RJ, Contreras M, Burton LC, Plisky PJ, Goode A, and Kiesel K. Modifiable risk factors predict injuries in firefighters during training academies. *Work* 46: 11–17, 2013.
6. Butler RJ, Plisky PJ, and Kiesel KB. Interrater reliability of videotaped performance on the functional movement screen using the 100-point scoring scale. *Athl Train Sports Health Care* 4: 103–109, 2012.
7. Chapman RF, Laymon AS, and Arnold T. Functional movement scores and longitudinal performance outcomes in elite track and field athletes. *Int J Sports Physiol Perform* 9: 203–211, 2014.
8. Chorba RS, Chorba DJ, Bouillon LE, Overmyer CA, and Landis JA. Use of a functional movement screening tool to determine injury risk in female collegiate athletes. *N Am J Sports Phys Ther* 5: 47, 2010.
9. Cook G. *Athletic Body in Balance—Optimal Movement Skills and Conditioning for Performance*. Champaign, IL: Human Kinetics, 2004. pp. 26–28, 2004.
10. Cook G, Burton L, Fields K, and Kiesel K. *The Functional Movement Screen*. Danville, VA: Athletic Testing Services Inc, 1998.
11. Cook G, Burton L, and Hoogenboom B. Pre-participation screening: The use of fundamental movements as an assessment of function—Part 1. *N Am J Sports Phys Ther* 1: 62, 2006.
12. Cook G, Burton L, and Hoogenboom B. Pre-participation screening: The use of fundamental movements as an assessment of function—Part 2. *N Am J Sports Phys Ther* 1: 132, 2006.
13. Cowen VS. Functional fitness improvements after a worksite-based yoga initiative. *J Bodyw Mov Ther* 14: 50–54, 2010.
14. Dallinga JM, Benjaminse A, and Lemmink KA. Which screening tools can predict injury to the lower extremities in team sports? *Sports Med* 42: 791–815, 2012.
15. Duncan MJ and Stanley M. Functional movement is negatively associated with weight status and positively associated with physical activity in British primary school children. *J Obes* 2012: 697563, 2012.
16. Duncan MJ, Stanley M, and Wright SL. The association between functional movement and overweight and obesity in British primary school children. *BMC Sports Sci Med Rehabil* 5: 11, 2013.
17. Elias JE. The inter-rater reliability of the functional movement screen within an athletic population using untrained raters. *J Strength Cond Res*, 2013. Epub ahead of print.
18. Faude O, Junge A, Kindermann W, and Dvorak J. Risk factors for injuries in elite female soccer players. *Br J Sports Med* 40: 785–790, 2006.
19. Florkowski CM. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: Communicating the performance of diagnostic tests. *Clin Biochem Rev* 29 (Suppl 1): S83, 2008.
20. Fox D, O'Malley E, and Blake C. Normative data for the Functional Movement Screen™ in male gaelic field sports. *Phys Ther Sport* 27: 458–462, 2013.
21. Frohm A, Heijne A, Kowalski J, Svensson P, and Myklebust G. A nine-test screening battery for athletes: A reliability study. *Scand J Med Sci Sports* 22: 306–315, 2012.
22. Frost DM, Beach TA, Callaghan JP, and McGill SM. Using the Functional Movement Screen™ to evaluate the effectiveness of training. *J Strength Cond Res* 26: 1620–1630, 2012.
23. Frost DM, Beach TA, Callaghan JP, and McGill SM. FMS™ scores change with performers' knowledge of the grading criteria—Are general whole-body movement screens capturing "dysfunction"? *J Strength Cond Res*, 2013. Epub ahead of print.
24. Frost DM, Beach TA, Callaghan JP, and McGill SM. The influence of load and speed on individuals' movement behavior. *J Strength Cond Res*, 2013. Epub ahead of print.
25. Galli M, Crivellini M, Sibella F, Montesano A, Bertocco P, and Parisio C. Sit-to-stand movement analysis in obese subjects. *Int J Obes Relat Metab Disord* 24: 1488–1492, 2000.
26. Gilleard W and Smith T. Effect of obesity on posture and hip joint moments during a standing task, and trunk forward flexion motion. *Int J Obes (Lond)* 31: 267–271, 2006.
27. Goss DL, Christopher GE, Faulk RT, and Moore J. Functional training program bridges rehabilitation and return to duty. *J Spec Oper Med* 9: 29, 2009.
28. Gribble PA, Brigle J, Pietrosimone BG, Pfile KR, and Webster KA. Intrarater reliability of the functional movement screen. *J Strength Cond Res* 27: 978–981, 2013.
29. Grygorowicz M, Piontek T, and Dudzinski W. Evaluation of functional limitations in female soccer players and their relationship with sports level—A cross sectional study. *PLoS One* 8: e66871, 2013.
30. Harman E and Garhammer J. Principles of test selection and administration. In: *Essentials of Strength Training and Conditioning*. Baechle TR and Earle RW, eds. Champaign, IL: Human Kinetics, 2008. pp. 239–240.
31. Hayes AF and Krippendorff K. Answering the call for a standard reliability measure for coding data. *Commun Methods Measures* 1: 77–89, 2007.
32. Hills AP, Hennig EM, Byrne NM, and Steele JR. The biomechanics of adiposity—Structural and functional limitations of obesity and implications for movement. *Obes Rev* 3: 35–43, 2002.
33. Kale M, Asçi A, Bayrak C, and Açikada C. Relationships among jumping performances and sprint parameters during maximum speed phase in sprinters. *J Strength Cond Res* 23: 2272, 2009.
34. Karras DJ. Statistical methodology: II. Reliability and validity assessment in study design, Part A. *Acad Emerg Med* 4: 64–71, 1997.
35. Karras DJ. Statistical methodology: II. Reliability and validity assessment in study design, Part B. *Acad Emerg Med* 4: 144–147, 1997.
36. Kiesel KB, Butler RJ, and Plisky PJ. Limited and asymmetrical fundamental movement patterns predict injury in American football players. *J Sport Rehabil* 23: 88–94, 2014.
37. Kiesel K, Plisky P, and Butler R. Functional movement test scores improve following a standardized off-season intervention program in professional football players. *Scand J Med Sci Sports* 21: 287–292, 2011.
38. Kiesel K, Plisky PJ, and Voight ML. Can serious injury in professional football be predicted by a preseason functional movement screen?. *N Am J Sports Phys Ther* 2: 147, 2007.
39. Klusemann MJ, Pyne DB, Fay TS, and Drinkwater EJ. Online video-based resistance training improves the physical capacity of junior basketball athletes. *J Strength Cond Res* 26: 2677–2684, 2012.
40. Lehr ME, Plisky PJ, Butler RJ, Fink ML, Kiesel KB, and Underwood FB. Field-expedient screening and injury risk

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

algorithm categories as predictors of noncontact lower extremity injury. *Scand J Med Sci Sports* 23: e225–e232, 2013.

41. McGill S, Frost D, Andersen J, Crosby I, and Gardiner D. Movement quality and links to measures of fitness in firefighters. *Work* 43: 357–366, 2012.
42. Minick KI, Kiesel KB, Burton L, Taylor A, Plisky P, and Butler RJ. Interrater reliability of the functional movement screen. *J Strength Cond Res* 24: 479–486, 2010.
43. O'Connor FG, Deuster PA, Davis J, Pappas CG, and Knapik JJ. Functional movement screening: Predicting injuries in officer candidates. *Med Sci Sports Exerc* 43: 2224–2230, 2011.
44. Okada T, Huxel KC, and Nesser TW. Relationship between core stability, functional movement, and performance. *J Strength Cond Res* 25: 252–261, 2011.
45. Onate JA, Dewey T, Kollock RO, Thomas KS, Van Lunen BL, DeMaio M, and Ringleb SI. Real-time intersession and interrater reliability of the functional movement screen. *J Strength Cond Res* 26: 408–415, 2012.
46. Pacheco MM, Teixeira LA, Franchini E, and Takito MY. Functional vs. strength training in adults: Specific needs define the best intervention. *Int J Sports Phys Ther* 8: 34, 2013.
47. Parchmann CJ and McBride JM. Relationship between functional movement screen and athletic performance. *J Strength Cond Res* 25: 3378–3384, 2011.
48. Parenteau-G E, Gaudreault N, Chambers S, Boisvert C, Grenier A, Gagné G, and Balg F. Functional movement screen test: A reliable screening test for young elite ice hockey players. *Phys Ther Sport* 5: 169–175, 2014.
49. Paszkewicz JR, McCarty CW, and Van Lunen BL. Comparison of functional and static evaluation tools among adolescent athletes. *J Strength Cond Res* 7: 2842–2850, 2013.
50. Peate WF, Bates G, Lunda K, Francis S, and Bellamy K. Core strength: A new model for injury prediction and prevention. *J Occup Med Toxicol* 2: 1–9, 2007.
51. Perry FT and Koehle MS. Normative data for the functional movement screen in middle-aged adults. *J Strength Cond Res* 27: 458–462, 2013.
52. Ristanis S, Stergiou N, Patras K, Vasiliadis HS, Giakas G, and Georgoulis AD. Excessive tibial rotation during high-demand activities is not restored by anterior cruciate ligament reconstruction. *Arthroscopy* 21: 1323–1329, 2005.
53. Schneiders AG, Davidsson Å, Hörman E, and Sullivan SJ. Functional movement screen™ normative values in a young, active population. *Int J Sports Phys Ther* 6: 75, 2011.
54. Shojaedin SS, Letafatkar A, Hadadnezhad M, and Dehkhoda MR. Relationship between functional movement screening score and history of injury and identifying the predictive value of the FMS for injury. *Int J Inj Contr Saf Promot*, 2013. Epub ahead of print.
55. Shultz R, Anderson SC, Matheson GO, Marcello B, and Besier T. Test-retest and interrater reliability of the functional movement screen. *J Athl Train* 48: 331–336, 2013.
56. Sibella F, Galli M, Romei M, Montesano A, and Crivellini M. Biomechanical analysis of sit-to-stand movement in normal and obese subjects. *Clin Biomech (Bristol, Avon)* 18: 745–750, 2003.
57. Smith CA, Chimera NJ, Wright NJ, and Warren M. Interrater and intrarater reliability of the functional movement screen. *J Strength Cond Res* 27: 982–987, 2013.
58. Stergiou N, Ristanis S, Moraiti C, and Georgoulis AD. Tibial rotation in anterior cruciate ligament (ACL)-deficient and ACL-reconstructed knees: A theoretical proposition for the development of osteoarthritis. *Sports Med* 37: 601, 2007.
59. Teyhen DS, Shaffer SW, Lorenson CL, Halfpap JP, Donofry DF, Walker MJ, Dugan JL, and Childs JD. The functional movement screen: A reliability study. *J Orthop Sports Phys Ther* 42: 530–540, 2012.
60. Viera AJ and Garrett JM. Understanding interobserver agreement: The kappa statistic. *Fam Med* 37: 360–363, 2005.
61. Waldén M, Hägglund M, and Ekstrand J. High risk of new knee injury in elite footballers with previous anterior cruciate ligament injury. *Br J Sports Med* 40: 158–162, 2006.
62. Wearing SC, Hennig EM, Byrne NM, Steele JR, and Hills AP. The biomechanics of restricted movement in adult obesity. *Obes Rev* 7: 13–24, 2006.
63. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 19: 231–240, 2005.
64. Xu X, Mirka GA, and Hsiang SM. The effects of obesity on lifting performance. *Appl Ergon* 39: 93–98, 2008.
65. Zhang J and Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. *JAMA* 280: 1690, 1998.
66. Zweig MH and Campbell G. Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clin Chem* 39: 561–577, 1993.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.