Systematic Review of the Functional Movement Screen™

Mike Kulju, BS,* Dustin Collins, BS,** Andy Swentik, MD*** and J. Bryan Dixon, MD****

[Abstract]

Context: The Functional Movement Screen™ (FMS™) is a quantitative movement analysis system created to evaluate athletes for potential injury risk. The FMS™ has become a popular assessment in fitness, sports and rehabilitation settings. In addition, the FMS™ has been increasingly integrated into the sports pre-participation physical exam process by team physicians.

Objective: Provide a systematic review of the published literature on the FMS™. Review and evaluate the literature regarding the reliability and validity of the FMS™ for injury prevention and sport performance.

Data Sources: PubMed, Google Scholar and Ovid were searched.

Study Selection: Studies that sought reliability or validity of the FMS™ in assessing movement quality, injury risk, or sports performance. Peer review publication was needed for inclusion and publications exclusively as posters were excluded.

Results: Overall, 24 studies were reviewed. The FMS™ was found to have good to excellent reliability. The FMS™ appears to be a valid measure to detect deficits in gross movement quality and to identify movement asymmetries, but does not seem to have the ability to predict sports performance. FMS™ appears to be a valid measure of injury risk when a cutoff score of 14 is used.

Conclusions: As a screening test for injury risk, the FMS™ has the potential to be an improvement on the standard practice. It has shown the ability to highlight abnormal movement patterns in athletes and be used as a predictor of injury in certain clinical settings. Its use to predict performance is not supported by this review. Although the FMS™ shows promise in sports medicine applications, further research is warranted prior to broad implementation in sports medicine settings.

Keywords: injury prevention; reliability; validity; Functional Movement Screen

*Michigan State University, College of Human Medicine
**Northern Michigan University
*** University of Missouri
****Marquette General Hospital/Duke Lifepoint
Each year in the United States, an estimated 180 million people participate in sports or recreational activity. (1) Sports and recreational physical activity have many health benefits, but also carry a risk for injury. Although the total extent of sports and recreational injuries are unknown, injury surveillance studies suggest that there are 4.5 million sports- and recreation-related injuries annually in the United States. (2) Although injury causation is complex, risk factors are likely to include intrinsic factors of the individual participant. (3) Persons engaging in sports or recreational physical activity should be aware of their individual risks for injury. A practical screening test to predict musculoskeletal injury in the individual would be a valuable tool in sports, fitness and health settings to inform participants and potentially reduce morbidity associated with sports and recreational activity.

The clinical value of a screening test is established by its ability to reduce morbidity and mortality. Valuable musculoskeletal injury screening must predict injury risk beyond those inherent in the activity, or identify conditions that will negatively affect the athlete. In addition to these categorical requirements, the test must be sufficiently sensitive and specific, while also being practical and affordable.

Current efforts at musculoskeletal (MSK) screening in sports medicine are evident in the pre-participation examination (PPE) process. The fourth edition of the PPE monograph is a dedicated effort to provide increasing standardization and evidence-based support for the PPE process. The MSK screening component of the PPE includes three components. First, historical information is acquired from a standardized questionnaire. The patient questionnaire identifies most musculoskeletal conditions relevant to safety for participation in sports. (4, 5, 6, 7, 8) (level of evidence = B according to PPE). Work by Gomez (8) has shown the history by itself to be 92% sensitive in detecting significant MSK injuries. In addition to the MSK questionnaire, the PPE monograph discusses two distinct but inter-related approaches to MSK screening with physical exam. First is the general screening examination, and second, a joint-specific examination. These tests can be used in conjunction or alone to screen for musculoskeletal conditions. (Preparticipation Physical Evaluation, Fourth Edition. Editors: Bernhardt DT, Roberts WO. 2010)

Despite its widespread use and the general recommendation for MSK screening exams in the PPE, the value of the current MSK screening examination appears to be low in predicting injury and insensitive in identifying significant MSK conditions. (8, 9) For these reasons, alternative approaches for MSK screening are desired. The Functional Movement screen™ (FMS™) represents a potential candidate for replacing the current general MSK screening exam or augmenting the joint-specific exams. The current MSK screening exam seeks to evaluate gross asymmetries, imbalances, and movement restrictions that may predispose to injury. (9, 10)

These objectives reflect the belief that movement proficiency and quality are an integral component of safe sports participation and that asymmetries in movement are a risk factor for injury.
The Functional Movement Screen™ (FMS™) is a quantitative movement analysis tool created in 1997. FMS™ was created to fill the need for a standardized, objective and reproducible screen that could be completed by both lay individuals and trained sports and medical professionals. The FMS™ incorporates seven movements on a 0-3 (or, more recently, a 0-100) scoring system to grade the quality of movement. As the name implies, the FMS™ was intended to serve as a screen to identify individuals with functional movement deficits that could indicate an increased risk of injury.

Over the past 15 years the FMS™ has been gaining favor and use by exercise scientists, physical therapists, and clinicians alike. Some health professionals note that the gain in popularity of the FMS™ has been more driven by marketing and less by research-proven results. This systematic review of FMS™ literature will evaluate and critique the collection of peer-reviewed research on the FMS™ to help assess its current value as a screening tool by looking at data on reliability and validity.

Figure 1 (Illustration of the FMS movements)

Figure 2 (Description of FMS scoring)

Figure 3 (FMS scoring sheet)

**Literature Review Methods and Article Identification**

The databases of PubMed, Google Scholar, and Ovid were searched using the terms *Functional Movement Screen, functional movement,* and *FMS*. Literature searches were performed both electronically and by manual evaluation of bibliographic references from these articles obtained. The search was limited to articles in the English language. This process resulted in the return of 247
published articles. Subsequently, abstracts and non-peer-reviewed publications were excluded.

![Diagram](image.png)

**Criteria for Article Selection**

Articles were included in the systematic review if they addressed FMS™ reliability or validity. Articles addressing validity are grouped according to topic as noted below. Reliability was defined as inter-rater, intra-rater or test-retest reliability. Validity has no consensus definition when used for medical research purposes. In common usage, validity describes how well the objective corresponds to reality. This is distinct from the concept of formal validity, which is restricted to evaluations of form only and not content. In this article, validity is used in the more general sense; that is, to evaluate the evidence that the FMS™ corresponds to real-world applications in which it has been studied. Our literature review included articles that addressed three areas of application of the FMS™, namely, movement quality, injury risk, and sports performance. In some cases, articles evaluated more than just those aspects of the FMS™.

**Results**
Our goal was to review the literature on the FMS™ in order to evaluate its utility in the clinical setting. We hoped to review its strengths and weaknesses to determine its true value as a screening exam. Key areas reviewed, and areas with the most literature, included the reliability and validity of the FMS™. Other studies reviewed addressed the FMS™ ability to predict performance in some way.

**Reliability of the FMS™**

Multiple studies have indicated substantial to excellent reliability. This is despite the fact that many of the articles have vast differences in their testing methods, scorer training, test viewing methods, and even in the statistical analyses used to report the data. When using the 21-point FMS™, Schneider et al recorded an inter-rater Kappa score of <0.70 for two FMS™ trained scorers watching previously recorded movement videos (12). Frohm et al recorded an intra-rater Kappa of 0.80 using eight live-scoring physiotherapists (14). Onate et al recorded a strong intra-class correlation coefficient (ICC) of 0.98 for inter-rater reliability when using live scoring with one certified FMS™ specialist and one FMS™ novice, whereas Teyhen et al recorded a moderate inter-rater ICC of 0.74 when the scoring was done by four FMS™ trained physical therapy doctoral students (22) (26). Studies by Smith and Gribble et al both attempted to look at the correlation between FMS™ scores by raters of differing backgrounds and experience levels. Smith in his study used two sessions where the FMS™ was scored live; the study recorded ICCs of 0.89 and 0.87 (27). In Gribble’s study the exam was scored using three video-taped subjects who were reviewed in random order a week later. A strong intra-rater ICC (0.96) was recorded by trained athletic trainers (ATC) with greater than six months’ experience with FMS™, a moderate intra-rater ICC (0.77) with trained ATCs with less than six months of FMS™ experience, and a poor intra-rater ICC (0.37) with ATC students. However, when comparing the scores across all participants, without classification of experience, Gribble’s study produced a moderate intra-rater ICC (0.75) (34). One study by Anstee et al recorded inter-rater ICCs of 0.63 and 0.66, which are only considered moderate, but that same study did record intra-rater ICC of 0.98 (29).

Some argue that the range of possible scores for the 21-point FMS™ (0-3) is limited in its ability to truly separate those at risk, and is almost “inherently reliable.” A study by Butler et al that used the newer 100-point FMS™ scoring system recorded inter-rater ICCs ranging from 0.91 to 1.00, with a composite ICC score of 0.99 (31). The new 100-point system has yet to be accepted as widely as the current scoring system, and would need further research before we could recommend its use.

Based on our review, we believe the FMS™ has good to excellent reliability. As with most screening exams, it seems experience does increase the reliability of scores, but the FMS™ has been shown to be reliable with minimal experience. The vast array of study designs and reported statistical analysis does make direct comparison of the literature difficult, but it does support the FMS™ overall as a screening tool that can be implemented in many ways and in many situations and still be considered reliable.

**Validity of the FMS™**

*Validity of the FMS™ in assessing movement quality*
Objective evaluation of movement quality is the criterion on which the FMS™ was developed. Cook and Burton developed the FMS™ as a screening tool to “determine whether the athlete has the essential movements needed to participate in sports activities with a decreased risk of injury.”

Validity of the FMS™ in assessing injury risk

Using the FMS™ to assess the injury risk in athletes requires a defined failure, or cut-off score. The most widely accepted failure or cut-off score referenced in the literature, when using the 21-point FMS™, is a composite FMS™ score of 14. This referenced value comes from a study by Kiesel et al conducted in 2007, using a cohort of professional football players (18). The cut-off score of 14 is supported by an O’Connor et al study of a large number of marine officer candidates (28), a study by Chorba et al with female collegiate athletes, and in a Lisman et al study also using marine officers (33). Even though 14 is the widely excepted cut-off score, there was a Kiesel et al study that used a group of firefighters that reported a failure score of 13 (24). In contrast to the above studies, a study using firefighters by Peate et al used a cut-off score of 16 (19).

The predictive value of the cut-off score is what determines the test’s true value. Whether reported in terms of specificity or an odds ratio, the reported results published to date have shown promise. Kiesel et al’s study of professional football players leads the way with a specificity of 0.91 (18). The study by O’Connor et al evaluating marine officer candidates recorded an injury prediction specificity of 0.71 (30). Our review found studies with reported odds ratios ranging from 1.68 to 11, and one that recorded a relative risk of injury of 1.5 for those who failed to score above the study-specific cut-off score (19) (28).

While the published data to this point has been encouraging, it still should be viewed in the context in which it has been found. There have been significant differences in the definition of injury used in each study. Chorba et al defined injury as a musculoskeletal injury that occurred as the result of participation in an organized intercollegiate practice or competition setting, for which an athlete sought advice or required medical attention from a physician, certified athletic trainer, or athletic training student (16). In contrast, Kiesel et al defined injury as players being placed on the injured reserve list who lost three weeks of time due to a serious injury (18). Lisman et al defined injury in their marine officers as any injury, which included over-use injuries and traumatic injuries (33). Another factor that should be taken into consideration is the sizable differences between powers of the studies. The study by O’Connor et al (28) tested a much larger cohort compared to the other studies. This study should be used as the “gold standard” when considering failing score injury risk, due to its overwhelming power and heterogeneity of participants, when compared to the others. It is also highly likely that the existence of only six studies published to date is an indication of publication bias.

The data we were able to review are encouraging in their support of the FMS™ as a screening test. Each of the six main studies was able to establish a cut-off score that demonstrated an increased risk of injury for any participant who did not meet that threshold. The predicative value noted in each study did vary, and there are multiple deficiencies as mentioned above about the overall body of literature that must be taken into consideration; however, most current screening exams are also limited in their predictive
value. If used as a part of the overall pre-participation screening exam, we feel the FMS™ can be utilized to help predict injury to those participants who score below the current cut-off recommendation of 14.

Validity of the FMS™ in assessing sports performance

Aside from its proposed injury prediction ability, the FMS™ has been researched in connection to athletic performance. There are no current studies that show a relationship between athletic performance variables and an FMS™ score. A study using a cohort of NCAA Division I golfers showed that there is no significant correlation between FMS™ and Club Head Swing Velocity, Vertical Jump, Sprints of 10 and 20m, or the Agility T-test (11). Okada et al studied recreational athletes and found no correlation between FMS™ and core stability, and concluded that there was no relationship between their performance variables and FMS™ scores (15). Clifton et al found that FMS™ scores showed no relationship to changes in static balance after the completion of an exercise routine, in a cohort of young adults (32). Lisman et al found no correlation between the Marine PFT (physical fitness test) consisting of abdominal crunches, 3-mile run and pull-ups, and FMS™ score (33).

Validity of the FMS™ Other Studies

Kiesel et al found that using an off-season intervention program for professional football players improved the number of players who tested above the injury threshold compared with before the intervention (χ² = 164.9, P<0.01)(13). That study also showed that 50% of players were free from asymmetry at the end of the study as compared with 32% at the beginning of the study (χ² = 7.8, P=0.01) (13). The Peate et al article also showed a reduction in injuries post-intervention by 44% (19). The Frost et al article, conversely, showed no significant differences in the total FMS™ scores for any of their tested groups post-training (30).

The Kiesel et al study findings could possibly be put under scrutiny, as no control group was used and players may have improved their scores simply through the test-retest phenomenon. The finding that patients can actually improve FMS™ scores, and therefore reduce injury risk, is important particularly from a clinical perspective. If supported by future research it may provide an opportunity to screen participants for injury, and then design specific training programs to decrease these participants’ risk of future injury. The intervention following the discovery of a failing score is a component of FMS™ that is in need of further research.

In a study comparing normal weight and overweight/obese children by Duncan et al, it was shown that Total FMS™ score was significantly, negatively correlated with BMI. That study also showed that children of normal weight scored significantly higher on the FMS™ (average 15.5) than did overweight/obese children (average 10.6) (25). This study seems to indicate that body composition is related to injury risk.

Discussion

With the scope of current FMS™ literature, one should be hesitant to use the FMS™ heavily, or to consider the FMS™ a concrete clinical screening tool; instead, one should be open to using the FMS™ in
certain fields and for certain populations. The FMS™ shows promise in certain cohorts in its ability to predict injury, but has not proven significantly that it is an effective screening tool in large, generalized athletic populations. There is reason to question whether these current studies show appropriate external validity, as most of them have small numbers and are in very specific cohorts.

The correlation of low FMS™ scores to injury is something that is externally consistent across the current peer-reviewed literature, but the magnitude of this correlation is debatable. Of the reviewed literature, the results from the O’Connor et al study would be most representative of FMS™ injury prediction validity due to the large power of the study in comparison to others, and it represents the most externally valid study. An area of concern for FMS™ injury prediction validity is the fact that there are at least four non-peer-reviewed studies that indicated no correlation between FMS™ score and injury potential. Since these studies do not fit our inclusion criteria, their specifics are not noted in this review. Their results, however, bring into question whether the FMS™ results are subject to publication bias.

Many other questions still remain unanswered, such as the most valid way to score the test, how results should be interpreted, whether it is possible to correct failing scores, what populations can be scored effectively, and what the true implications of a qualifying failing score might be. With the extremely heterogeneous group of known studies, it may be difficult within the scope of the current literature to conclude any compelling relationship even between seemingly similar results. Many of the studies use such different criteria for defining injury that the injury prediction ability of the FMS™ may simply be correlated to each particular study’s specific injury definition.

Aside from qualifying what defines injury, it may be difficult to determine the likelihood of an athlete to report an injury that is not obvious, but would still fit the given study's injury criteria. If there are external pressures on the athlete, such as financial issues or team performance issues, an athlete may be driven to not report injury, therefore skewing the potential for quantifying “true” injury prediction. With the number of FMS™ studies done involving professional or collegiate cohorts, this may be a serious unaddressed confounder. Another issue in the defining or categorizing of injuries is how to deal with sporadic sport-specific injuries that do not or would not have a relationship to the athlete’s FMS™ score, such as a broken foot from a slap shot. Such injuries are a chance occurrence and are not predictable by any means.

As noted in Schneiders et al, to help improve the injury prediction validity, it may be useful to test FMS™ results in athletes from some sports with differing movement patterns such as dancers, gymnasts, or martial artists (12). The specific training in such sports may improve particular movement patterns or elevate normative FMS™ scores, possibly skewing the relation of low FMS™ to injury.

The reliability is often the most criticized aspect of the FMS™, but ironically seems to show the most congruence across the literature. We were unable to find any published peer-reviewed studies that showed the FMS™ scoring to be unreliable, even when considering the newer 100-point method. The current literature seems to suggest strongly that the FMS™ can be performed in a reliable, reproducible manner across trained scorers of many differing backgrounds. FMS™ scoring reliability across the literature has shown to be good to moderate when considering inter-rater, intra-rater or test-retest
reliability across most scorer training and experience levels, but does consistently show greater reliability when considering those scorers with more clinical experience. This would indicate the FMS™ would be very reliable when used by skilled, experienced health professionals.

As mentioned in Frost et al, one criticism of the FMS™ scoring in terms of scoring reliability is that when using the 0-3 scoring method, some tests have unclearly defined criteria of mid-range performance (a score of 2), which allows for a broad range of movement patterns to fall into this category (30). This makes it more likely that multiple “different” movement patterns will be assigned a score of 2, which is demonstrated by the fact that 2 tends to be the most common score across most movements of the seven tests. Broad groupings of movement patterns such as this would increase the reliability across scorers because it is “easier” to give a subject a 2.

**Conclusion**

FMS™ is a screening tool that has been shown to have good to excellent reliability and has potential application to a variety of sports medicine settings. The validity of the FMS™ in grading the quality of movement is moderate, but tends to lack measures to account for multi-planar or complex movement patterns. Other movement analysis methods, such as the Y-Balance test, may be better suited for those types of movements. The validity of the FMS™ in assessing injury risk is moderate to good, but is in need of additional population-specific studies and evaluation of apparent bimodal risk patterns. The validity of the FMS™ as an assessment of sports performance is poor, and hasn’t yet been shown in any published studies.

The FMS™ shows promise in sports medicine applications, but care must be taken to understand the limits, implications and consequences of FMS™ testing prior to use in sports medicine settings. The need for more definitive findings in future FMS™ research, particularly regarding effectively testable populations and the Functional Movement Screen’s™ true prediction potential, should be a concern for those who tend to apply the FMS™ in a clinical setting. When compared to the MSK screen, the FMS™ seems to have greater predictive ability, but may not be as useful as a generalized pre-participation screen. Due to the need for the users of the FMS™ to have special training, and the somewhat lengthy screen time, the FMS™ seems to have more use in an exercise science capacity than in a clinical setting. Clinical use of the screen may be appropriate in specialized patient populations.

Future research to discover effective interventions and corrective protocols for subjects with a failing score are needed. To maximize validity, they should utilize larger sample groups and replicate the results in more than one study, and the cohorts should be compared against a negative control group to better demonstrate that the FMS™ is, in fact, data that can be corrected. Although there are many cautions to be applied when deciphering the implications of an FMS™ score, even a modest predictive ability in a screening tool like the FMS™ can be useful in many sport-specific and medical settings. The FMS™ can be considered a useful tool with at least moderate injury prediction ability. It has the benefit of being reliable, easy to teach, and inexpensive. The FMS™ can be considered useful in certain specified applications.
References