Evaluation of the Functional Movement Screen as an Injury Prediction Tool Among Active Adult Populations: A Systematic Review and Meta-analysis

Bryan S. Dorrel, PhD, ATC,*† Terry Long, PhD,† Scott Shaffer, PT, PhD, OCS, ECS,‡ and Gregory D. Myer, PhD, FACSM, CSCS*D§

Context: The Functional Movement Screen (FMS) is an assessment tool for quality of human movement. Research reports a significant difference between FMS scores of subjects who later experienced injury and those who remain uninjured.

Objective: To systematically review literature related to predictive validity of the FMS. From the aggregated data, a meta-analysis was conducted to determine the prognostic accuracy of the FMS.

Data Sources: PubMed, Ebscohost, Google Scholar, and the Cochrane Review databases were searched between 1998 and February 20, 2014.

Study Selection: Identified studies were reviewed in full detail to validate inclusion criteria. Seven of the 11 identified studies were included. Articles were reviewed for inclusion criteria, then bias assessment and critical analysis were conducted.

Study Design: Systematic review and meta-analysis.

Level of Evidence: Level 3.

Data Extraction: Extracted data included the following: study type, methodology, study subjects, number of subjects, injury classification definition, FMS cut score, sensitivity, specificity, odds ratios, likelihood ratios (LR), predictive values, receiver operator characteristic (ROC) analysis, and area under the curve (AUC).

Results: Overall bias for the included 7 studies was low with respect to patient selection. Quality assessment scored 1 study 5 of a possible 7, 2 studies were scored 3 of 7, and 4 studies were scored 2 of 7. The meta-analysis indicated the FMS was more specific (85.7%) than sensitive (24.7%), with a positive predictive value of 42.8% and a negative predictive value of 72.5%. The area under the curve was 0.587 (LR+, 1.7; LR-, 0.87; 95% CI, 0.6-6.1) and the effect size was 0.68.

Conclusion: Based on analysis of the current literature, findings do not support the predictive validity of the FMS. Methodological and statistical limitations identified threaten the ability of the research to determine the predictive validity of FMS.

Keywords: Functional Movement Screen; injury prediction; diagnostic accuracy

From †Northwest Missouri State University, School of Health Science, Maryville, Missouri, ‡Academy of Health Sciences, Graduate School, US Army Baylor Doctoral Program in Physical Therapy, Fort Sam Houston, Texas, and §Division of Sports Medicine, Cincinnati Children’s Hospital Medical Center, Cincinnati, Ohio

*Address correspondence to Bryan S. Dorrel, PhD, ATC, Northwest Missouri State University, School of Health Science, 800 University Drive, Maryville, MO 64468 (email: bdorrel@nwmissouri.edu).

The authors report no potential conflicts of interest in the development and publication of this article.

DOI: 10.1177/1941738115607445

© 2015 The Author(s)
Among collegiate athletes, injuries occur at a rate of 13.8 injuries per 1000 athlete-exposures (AEs) while high school athletic injuries range between 2.51 and 4.36 per 1000 AEs. In 2005, lower extremity injuries among high school athletes were 2298 of a total 4350 injuries, projecting a potential of 807,222 lower extremity injuries nationwide at a rate of 1.33 per 1000 AEs.

As sport-related injuries occur frequently, steps to reduce injury can have an impact on the frequency and associated costs. Researchers in many disciplines dedicate time and resources to record measures and identify associated risk factors for specific injuries, identify those most at risk to sustain injury, and develop interventions that address the identified risks.

While researchers have determined risk factors for some specific injuries, they have not determined a parsimonious set of tests that identify individuals who are predisposed to future injuries. Despite these limitations, a few injury screening measures have demonstrated promise in various populations. The Functional Movement Screen (FMS) is one such assessment tool and is used to assess fundamental movement patterns in a practical and dynamic way. The FMS was specifically designed to bridge the gap between preseason physical examinations and physical performance testing. The intended purposes of the FMS include the following: (1) assessment of stability and mobility within the kinetic chain of full body movements, (2) identification of body asymmetries, and (3) recognition of overall poor quality movement patterns. Specific applications include screening active adults for future injury and establishing a baseline of movement competence to allow comparisons after treatment, rehabilitation, or human performance training.

The FMS comprises 7 individual tests: the deep squat, the in-line lunge, the hurdle step, shoulder flexibility, push-up, straight leg raise, and the rotary trunk stability assessment. Each FMS assessment is scored on a scale of 1 to 3. On completion of all portions of the test, the subject is issued a comprehensive score of 0 to 21. A score of “0” is issued on an individual test if the subject experiences any pain during the assessment process. A score of “1” indicates poor performance, and “3” excellent performance. Preliminary research indicates a significant difference between the comprehensive or individual FMS scores of individuals who were later injured and those who were not. These data provide a foundation of support, indicating that the test may identify those at high risk of sports-related injury. However, predictive validity across multiple active adult populations is currently unknown. The purpose of the current project was to systematically assess and use meta-analysis methodology to evaluate the current literature relative to the efficacy of the FMS for injury prediction in active adult populations. Specifically, we aim to aggregate and examine the existing literature that prospectively evaluated the FMS relative to the association with subsequent injury.

METHODS
Protocol and Registration
The review protocol for the systematic review was based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement for reporting of systematic reviews and meta-analyses for the evaluation of health care interventions. No previous registration of the project was conducted.

Information Sources, Eligibility Criteria, and Study Selection
PubMed, EBSCOhost, Google Scholar, and the Cochrane Review databases were searched between 1998 and February 20, 2014, with the following terms and Boolean phrases: “Functional movement screen” and “Functional Movement Screen” AND “Prediction of Injury.” In addition to searching databases, the reference lists of identified FMS literature were searched to find other potential articles on the predictive validity of the FMS. In addition, other researchers familiar with the FMS were solicited for their knowledge of relevant publications. All studies examining the ability of the FMS to predict injury among active adults (e.g., firefighters, athletes, military) were considered for inclusion. Inclusion was limited to studies published in peer-reviewed journals. The study selection was unblinded and conducted by the primary researcher. All identified studies were read and reviewed in full detail to validate the inclusion criteria (Figure 1).

Data Collection Process
Data were extracted from the studies and compiled in an Excel spreadsheet (Microsoft) by the primary author. Data extracted
included the following: general study type, study methodology, study subjects, number of study subjects, injury classification definition, FMS cut score, sensitivity, specificity, odds ratios, likelihood ratios, predictive values, receiver operator characteristic (ROC) analysis, area under the curve (AUC), and whether the study results demonstrated a significant difference between the FMS scores of the injured and uninjured subjects (Table 1 in the Appendix, available at http://sph.sagepub.com/content/by/supplemental-data).

**Risk of Bias, Quality, and Threats to Validity in Individual Studies**

Risk of bias was completed using the QUADAS-2, a recommended tool for use in systematic reviews of diagnostic accuracy. The QUADAS-2 is used to assess for risk of bias and applicability of articles that may be included as one develops a systematic review. Two members of the research team (B.S.D. & T.L.) reviewed the QUADAS-2 guidelines and independently scored each article. Once complete, the scoring was compared, discussed, and agreed upon. In addition to the QUADAS-2 bias assessment, the perceived study limitations and a quality assessment were conducted for each study based on statistical measures of diagnostic accuracy, systematic reviews, and meta-analysis. The quality assessment was composed of 7 criteria that included prospective nature, blinding of study participants, data collectors (index test), outcome assessors (injury data), ROC curve conducted to determine cut score, AUC reported, and threats to the validity noted in the study (study methodology, statistical methodology, or statistical reporting). A grade of “Yes,” “No,” or “Unreported” was issued in each area, and the total frequency of “Yes” scores were tallied to indicate overall quality.

**Meta-analysis**

A meta-analysis of studies that met the inclusion criteria was conducted using the dr-ROC Summary Meta-Analysis Software program version 2.0 (Diagnostic Research Design & Reporting). Analysis results provided a comprehensive summary of statistics calculated within studies of diagnostic accuracy and included: mean sensitivity and specificity, positive and negative predictive values, effect size, ROC summary, and AUC summary. Positive and negative likelihood ratios (LR+ and LR−, respectively) were calculated by the primary author.

**RESULTS**

**Study Selection**

Eleven potential articles were identified, while 7 studies were selected that met the inclusion criteria. Four studies did not meet the defined criteria (see Figure 1). Of the 4 excluded articles, 1 appraised FMS literature, and in 2 articles, the FMS was not tested alone. The QUADAS-2 bias assessment for the included studies in patient selection scored 3 studies as low risk of bias, 2 studies as high risk, and 2 studies as unclear due to a lack of methodological reporting. Studies as high risk, and 2 studies as unclear due to a lack of methodological reporting. For risk of bias of the reference standard (injury definition) among the included studies, 4 studies were scored as low risk, and 2 studies were scored as unclear due to a lack of methodological reporting. With regard to potential bias for the flow and timing, all 7 included studies were scored as unclear risk because none of the studies reported patient attrition rates or if and how any study subjects were excluded from the data set. With regard to the QUADAS-2 applicability assessment of the included studies, all were scored as low risk for patient selection. For the index test, 2 studies were scored as low applicability concern while 5 studies were scored as unclear. For the reference standard, 2 studies were scored as high applicability concern, while 3 studies as low, and 2 studies as unclear.

**Bias Assessment, Study Quality, and Threats to Validity**

The QUADAS-2 bias assessment for the included studies in patient selection scored 3 studies as low risk of bias, 2 studies as high risk, and 2 studies as unclear due to a lack of methodological reporting. For risk of bias of the index test (FMS) among the included studies, 3 studies were scored as low risk of bias and 4 studies were scored as unclear due to lack of methodological reporting. For risk of bias of the reference standard (injury diagnosis/injury definition) among the included studies, 4 studies were scored as low risk, and 2 studies were scored as unclear due to a lack of methodological reporting. With regard to potential bias for the flow and timing, all 7 included studies were scored as unclear risk because none of the studies reported patient attrition rates or if and how any study subjects were excluded from the data set. With regard to the QUADAS-2 applicability assessment of the included studies, all were scored as low risk for patient selection. For the index test, 2 studies were scored as low applicability concern while 5 studies were scored as unclear. For the reference standard, 2 studies were scored as high applicability concern, while 3 studies as low, and 2 studies as unclear.

**Meta-analysis**

Based on available data, the meta-analysis was limited to 6 of the 7 studies included in the systematic review. One study was excluded because statistics required to conduct a meta-analysis were not reported. Studies were weighted by the dr-ROC software program according to the number of study subjects. The meta-analysis indicated the FMS was more specific (0.85; 95% CI, 0.77-0.91) than sensitive (0.24; 95% CI, 0.15-0.36). Specificity is interpreted as the ability of the test to accurately classify those study subjects who score over the cut score and do not sustain injury. Sensitivity is interpreted as the ability of the test to accurately classify those study subjects who scored on or below the FMS cut score and sustain injury. The positive predictive value is the likelihood that a subject with a positive test actually has the target condition and was 0.42 (95% CI, 0.23-0.64). The negative predictive value is the likelihood that a subject with a negative test is actually negative for the target condition and was 0.72 (95% CI, 0.67-0.76). AUC is the ability of the test to accurately discriminate between those at risk and not at risk and was determined to be 0.58 (95% CI, 0.42-0.77). Likelihood ratios are a combination of sensitivity and specificity values reported as a ratio that can be used to quantify a shift in the posttest probability once a test result is determined. The
positive likelihood ratio (LR+) was calculated to be 1.65 (95% CI, 1.3-2.0), which would alter the probability of a positive test result to a minimal and unimportant degree. The negative likelihood ratio (LR−) was calculated to be 0.87 (95% CI, 0.82-0.92) and would as well provide only a minimal and unimportant change to a negative test result (Table 3). Relative risk was calculated to be 1.5 (95% CI, 1.3-1.7). Effect size was 0.67 (95% CI, -0.38 to 1.72).8,29
**DISCUSSION**

From the meta-analysis, the FMS provides adequate specificity (85%) and low sensitivity (24%), equating an AUC (0.58) that would provide a level of discriminatory accuracy slightly above chance. The positive likelihood ratio (LR+, 1.65) demonstrated a low score that would alter the probability to an insignificant and rarely impactful degree. The negative likelihood ratio (LR–, 0.87) may produce a small and rarely important shift in probability. Based on the various study limitations identified during the systematic review, the primary threats to validity are consistent reference standard definition, consistent data analysis methodology, and reporting that specifically includes the ROC, AUC, LR+, LR–, PV+, PV–, RR, CI, and effect size.

**Inconsistent Reference Standard Definition**

Examination of the current literature reveals differences in the reference standard (ie, definition of injury). All the included studies used the FMS as the index test and injury as the reference standard, but differences existed among the exact definition of injury. Inconsistent definition of the reference standard among current FMS studies may limit insight that can be drawn from aggregated data and is a limitation to the interpretation of the current meta-analysis. The problem is compounded by studies utilizing FMS cut scores recommended by studies utilizing a different reference standard other than their own. For example, the initial study by Kiesel et al. in which the reference standard was defined as injury that caused an athlete to be placed on the injured reserve for at least 3 weeks utilized a reference standard that was drastically different from others in FMS research. The study sample of football players likely sustained other injuries during the study period, many of which would have been identified as injuries in the criteria used in other FMS investigations. A musculoskeletal injury that sidelined a player for 2 weeks would account for a true positive in 6 of the 7 included studies, but not in the study by Kiesel et al. Therefore, the various definitions of injury utilized in the current study may limit the potential to draw conclusions relative to the aggregated data analysis.

**Inconsistent Data Analysis Methods**

Of the selected studies, 4 utilized study-specific data to determine their own respective cut score for the study population, but only 1 study reported the AUC. Two studies utilized the cut score of 14 because this was the score determined in the study by Kiesel et al. One study did not use ROC curve analysis to determine the study cut score but rather linear regression (see Table 1 in the Appendix). By using a cut score optimized to a different reference standard, researchers may fail to identify the optimal cut score for their study context and population, which would limit the potential of the FMS to accurately categorize risk. The use of one cut score may threaten the validity of another study’s results.

The AUC represents the diagnostic accuracy of a test, and failure to report the AUC makes it difficult for researchers to determine the ability of the FMS to effectively predict injury. The only study to report AUC is a good example (see Table 1 in the Appendix). By using a cut score optimized to a different reference standard, researchers may fail to identify the optimal cut score for their study context and population, which would limit the potential of the FMS to accurately categorize risk. The use of one cut score may threaten the validity of another study’s results.

**Methodological Limitations**

The overall quality of the available and included FMS research limits the interpretation of the current meta-analysis results. With
regard to the various methods of blinding used to enhance the validity of a study, most of the included studies fail to mention or discuss any methods used or attempts to blind aspects of their respective studies. In addition, all of the included studies report a 0 dropout rate and fail to discuss methodology utilized to assess or control research subject attrition. This may present another challenge to accurately meta-analyze current FMS research. Overall, the quality of the studies available and included in this systematic review was low and contained significant threats to validity, which renders their respective results relative to associations with injury prediction inconclusive.

CONCLUSION

The current aggregate results demonstrate that the FMS provides low sensitivity and a low AUC for discrimination of high injury risk, which indicates the diagnostic accuracy of the FMS to predict injury is low. In addition, neither LR+ nor LR− produces large, strong shifts in probability. The methodological and statistical limitations identified by this systematic review indicate the predictive validity of the FMS may be limited in the current aggregated analyses.

REFERENCES

1. Armason A, Sigurdsdottir SB, Godmundsson A, Holme I, Engebretsen L, Bahr R. Risk factors for injuries in football. Am J Sports Med. 2004;32(1 suppl):S5-16S.
2. Butler BJ, Contreras M, Burton LC, Plisky PJ, Goode A, Kiesel K. Modifiable risk factors predict injuries in firefighters during training academies. Work. 2015;46:11-17.
3. Chorba RS, Chorba DJ, Bouillion LE, Overmyer CA, Landis JA. Use of a functional movement screening tool to determine injury risk in female collegiate athletes. N Am J Sports Phys Ther. 2010;5:47-54.
4. Cook G, Burton L, Hoogenboom B. Pre-participation screening: the use of fundamental movements as an assessment of function—part 1. N Am J Sports Phys Ther. 2006;1:62-72.
5. Cook G, Burton L, Hoogenboom B. Pre-participation screening: the use of fundamental movements as an assessment of function—part 2. N Am J Sports Phys Ther. 2006;1:132-139.
6. Cook G, Burton L, Kiesel K, Rose G, Milo F. Movement Functional Movement Systems: Screening, Assessment and Corrective Strategies. Santa Cruz, CA: On Target Publication; 2010.
7. Fernandez WG, Yard EE, Comstock RD. Epidemiology of lower extremity injuries among US high school athletes. Acad Emerg Med. 2007;14:641-645.
8. Fritz JM, Wainner RS. Examining diagnostic tests: an evidence-based perspective. Phys Ther. 2001;81:1546-1564.
9. Häggland M, Waldén M, Ekstrand J. Previous injury as a risk factor for injury in elite football: a prospective study over two consecutive seasons. Br J Sports Med. 2006;40:767-772.
10. Hewett TE, Myer GD, Ford KR. Anterior cruciate ligament injuries in female athletes: part 1, mechanisms and risk factors. Am J Sports Med. 2006;34:299-311.
11. Hootman JM, Dick R, Agel J. Epidemiology of collegiate injuries for 15 sports summary and recommendations for injury prevention initiatives. J Athl Train. 2007;42:311-319.
12. Hoover D, Kiesel KB, Butler RJ, Plisky PJ. Prediction of injury by limited and asymmetrical fundamental movement patterns in American football players. J Sport Rehabil. 2014;23:88-94.
13. Junge A, Lamprecht M, Stamm H, et al. Countrywide campaign to prevent soccer injuries in Swiss amateur players. Am J Sports Med. 2011;39:57-63.
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;21:287-292.
15. Kiesel K, Plisky PJ, Voight ML. Can serious injury in professional football be predicted by a preseason Functional Movement Screen? N Am J Sports Phys Ther. 2007;2:147-158.
16. Kiesel KB, Butler RJ, Plisky PJ. Prediction of injury by limited and asymmetrical fundamental movement patterns in American football players. J Sport Rehabil. 2014;23:88-94.
17. Kramer M, Flanagan M, Bruener J, Durlal C. The accuracy of the Functional Movement Screen™ to identify individuals with an elevated risk of musculoskeletal injury. J Sport Rehabil. 2014;23:590-594.
18. Leetun DT, Ireland ML, Wilson JD, Ballantyne BT, Davis JM. Core stability measures as risk factors for lower extremity injury in athletes. Med Sci Sports Exerc. 2004;36:920-934.
19. Lehr ME, Plisky PJ, Butler RJ, Fink ML, Kiesel KB, Underwood FB. Field-expedient screening and injury risk algorithm categories as predictors of noncontact lower extremity injury. Scand J Med Sci Sports. 2013;23:225-e232.
20. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med. 2009;151:W65-W94.
21. Lisman P, O’Connor FG, Deuster PA, Knajpik J. Functional Movement Screen and aerobic fitness predict injuries in military training. Med Sci Sports Exerc. 2013;45:636-643.
22. McKeon PO, Hertel J. Systematic review of postural control and lateral ankle instability, part 1: can deficits be detected with instrumented testing. J Athl Train. 2008;43:293-304.
23. Mohler D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151:264-269.
24. Myer GD, Ford KB, Khoury J, Succop P, Hewett TE. Biomechanics laboratory-based prediction algorithm to identify female athletes with high knee loads that increase risk of ACL injury. Br J Sports Med. 2011;45:245-252.
25. O’Connor FG, Deuster PA, Davis J, Pappas CG, Knajpik J. Functional Movement Screening: predicting injuries in officer candidates. Med Sci Sports Exerc. 2011;43:2224-2230.
26. Orchard JW. Preventing sports injuries at the national level: time for other nations to follow New Zealand’s remarkable success. Br J Sports Med. 2008;42:392-393.
27. Parvin CA. Book, Software, and Web Site Brows: dr-ROC software. Mathew D. Mitchell. Ginspade, PA: DrzConsulting, 2005, CD. Clin Chem. 2006;52:545-544.
28. Peate W, Bates G, Landa K, Franco S, Bellamy K. Cone strength: a new model for injury prediction and prevention. J Occup Med Toxicol. 2007;2:5.
29. Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice. Norwalk, CT: Appleton & Lange; 1993.
30. Rechel JA, Yard EE, Comstock RD. An epidemiologic comparison of high school sports injuries sustained in practice and competition. J Athl Train. 2008;43:197-204.
31. Shankar PR, Fields SK, Collins CL, Dick RW, Comstock RD. Epidemiology of high school and collegiate football injuries in the United States, 2005-2006. Am J Sports Med. 2007;35:1295-1303.
32. Shojaeddin SS, Latafatakfer A, Hadadnejhad M, Dehlhooda MR. Relationship between Functional Movement Screening score and injury of history and identifying the predictive value of the FMS for injury. Int J Inj Contr Saf Promot. 2014;21:355-360.
33. Sugimoto D, Myer GD, Foss KD, Hewett TE. Dosage effects of neuromuscular training intervention to reduce anterior cruciate ligament injuries in female athletes: meta-and sub-group analyses. Sports Med. 2014;44:551-562.
34. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155:529-536.
35. Zazulaik BT, Hewett TE, Reeves NP, Goldberg B, Chodwick J. Deficits in neuromuscular control of the trunk predict crank knee injury risk. Am J Sports Med. 2007;35:1123-1130.

For reprints and permission queries, please visit SAGE’s Web site at http://www.sagepub.com/journalsPermissions.nav.