World Journal of Sport Research
Print ISSN: 2959-9946
Online ISSN: 2959-9954
DOI: https://doi.org/10.61784/wjsr3019

FMS PREDICTIVE VALIDITY & STABILITY: CRITERION (76 BASKETBALL ATHLETES) & TEST-RETEST (60 EXERCISE PARTICIPANTS)

Rui He[1], WeiBo Shang[1], Chen Zhao[2], JunJie Zheng[3], RongJin He[4], FangHao Li[5*]
[1]*Shanghai Shuangyou Es Sports Culture Development Co., Ltd., Shanghai 200000, China.*
[2]*Northwestern Polytechnical University, Xi' an 710000, Shanxi, China.*
[3]*Department of Physical Education, Neijiang Normal University, Neijiang 641100, Sichuan, China.*
[4]*Super Deer Sport, Xiamen 361000, Fujian, China.*
[5]*Faculty of Sport Science and Coaching, Universiti Pendidikan Sultan Idris, Tanjong Malim 35900, Perak, Malaysia.*
[*]*Corresponding Author: FangHao Li*

Abstract: Purpose. The present study investigated predictive validity of the Functional Movement Screen (FMS) in predicting basketball performance and its test-retest reliability and intervention sensitivity. Methods. The first study comprised of 76 basketball athletes who played the national-level 1 and above; the correlation between the total score of FMS and four performance indicators was examined and equivalence-tested. The second study randomized 60 exercise subjects (mean age = 20.8 years, SD = 1.8 years) into an intervention group (corrective education and exercises) or a control group with measurements taken one week apart and the assessors blinded. Results. The results indicated that Study 1: all correlations between FMS and performance were insignificant (|r| < 0.06; equivalence was established). Study 2: controls also had a very high reliability (r = 0.91, ICC = 0.90, change -0.1 ± 1.3, 93% within 2 points); the intervention group had a much lower correlation (r = 0.47 compared with r = 0.91, p < .001) and more variance in changes (SD = 2.9 compared with SD = 1.3, p < .001), and individual changes ranged between -5 and +6. Conclusion. FMS is not predictive of the basketball performance and is associated with low stability in the conditions of short intervention. The corrective education can change the meaning of scoring, which will compromise test-retest reliability. It is recommended that practitioners do not use FMS as the only decision-making or evaluation measure.
Keywords: Functional movement screen; FMS; Predictive validity; Test-retest reliability; Practice effect; Basketball

1 INTRODUCTION

The Functional Movement Screen (FMS) is a set of seven basic movements that are used to evaluate stability and mobility, detect compensations and asymmetries, and consequently predict the likelihood of sports-related injury [1]. Over the past few years, FMS has been adopted on a large scale among athletes, and some of the practitioners have started extrapolating the FMS score to denote athletic performance or using it as the criterion of athlete selection or effectiveness of training [2]. Alongside its mass spread in practice, the dissonance between the scientific basis of the FMS and commercialization has become more intense and even a sociological criticism has been leveled at it as a breeding ground of pseudoscience because the excessive commercialization outweighs its modest empirical basis [3]. However, the usage of any measuring device implies that the presence of sufficient reliability and validity data is necessary [4]. The ability to predict athletic performance with FMS scores (predictive validity) and the stability of scores across time (test-retest reliability) are the fundamental psychometric characteristics of the FMS as a measure.
The current cross-sectional researches have shown that the relationship between the sum of the score on the FMS and power-related measurements of performance, which are counter-movement jump and sprint, is very weak (r < 0.2) [5-6]. Nevertheless, the majority of the studies involved general university student populations, used relatively limited sets of performance indicators, and did not use equivalence tests in making statistical inferences, thus denying them methodologically sound confirmation of “meaningfully negligible relationships”. Most importantly, performance in basketball is strongly determined by various capacities, such as muscle strength (e.g. the capacity to perform a weighted squat), power output (e.g. counter-movement jump height and explosive movement speed), agility and linear sprint speed. As yet, there has been no research involving gold-standard apparatus to simultaneously record the above core indicators and assess the predictive validity of the FMS. In order to fill this gap, Study 1 recruited 76 national-level 1 and above basketball players and measured four indicators simultaneously, i.e. peak velocity of 30 percent 1RM loaded squat jump, counter-movement jump height, lane agility test, and ¾ court sprint, using force platform and linear position transducer and photocell timing system. The severity of the associations was strictly judged through the use of two one-sided test (TOST) to determine equivalence with a predetermined equivalence limit of r being less than or equal to 0.30.
Moreover, the results on the test-retest reliability of the FMS have also been inconsistent: some studies reported ICCs of 0.7-0.9 [7], other studies have found a substantial practice effect on initial testing, with subsequent systematic increases in scores on retesting [8]. Also, short educational or feedback interventions have also been found to have dramatic

effects on FMS scores which indicates that the FMS might not be robust and that its test-retest reliability might be threatened [9].

Basketball requires extreme physical power, movement velocity and agility of its players. Will FMS scores be sufficient indicators of such abilities? Are FMS scores constant in both non-intervention and intervention circumstances? These queries were examined through two sub-studies of the current research. Study 1 made use of a highly controlled cross-sectional methodology, which used the gold-standard instrument and equivalence testing framework that is mentioned above, to confirm the predictive validity of the FMS total scores with respect to multidimensional basketball-specific performance indicators. Study 2 used a randomized controlled test-retest design to explore the natural stability of the FMS over a brief period and its reactivity to one session of standardized movement education. Our hypothesis was that the size of correlations between FMS and every performance measurement will be less than a moderately strong correlation ($r < 0.30$) and that the intervention will cause significant changes in FMS scores, thus compromising its claim of being a reliable screening tool.

According to the methodology of diagnostics, any screening tool should be promoted and implemented based on the well-established validation of the tool in comparison to the gold standards. To illustrate it using splenic tuberculosis (a disease not common in clinics), Liu et al. (2025) proved conclusively that using non-specific clinical symptoms or standard laboratory tests is strongly associated with the likelihood of misdiagnosis or late treatment, and that correct diagnosis can only be made by objective diagnostic tools like pathological or imaging tests [10]. On similar grounds, in the field of evaluation of athletic performance, assuming that FMS total scores will serve as a screening indicator of athletic capacity, the power of its correlation with the gold-standard measurements, such as force platforms and photocell timing systems, should be examined thoroughly. Without this validation, the extensive use and implementation of the FMS in the sports environment can be subject to validity danger comparable to the one experienced in clinical misdiagnosis.

2 METHOD

2.1 Study 1: Predictive Validity

The subjects: The study involved 76 male basketball players (national-level 1 and higher; McKay classification: Tier 3) who had not suffered any recent injuries, and they all gave their written informed consent to take part in the research. Descriptive statistics included: age 22.9 ± 3.65 years, height 185.8 ± 8.39 cm, body mass 86.9 ± 15.31 kg.

Test Procedures: All the testing was done in an indoor standard basketball court. Subjects initially participated in the FMS standardized testing battery. Certified professionals who have attained FMS Level 2 Advanced certification marked all of the FMS assessments, as well as senior students in the College of Physical Education and Sports Science (inter-rater reliability: $\kappa = 0.91$). After adequate rest, participants completed the following in this sequence of tests: ¾-court sprint (photocell timing, SmartSpeed), countermovement jump (CMJ; Kistler force platform, 1000 Hz), peak velocity on 30 percent 1RM loaded squat jump (GymAware RS linear position transducer), and Lane Agility Test (photocell timing, SmartSpeed). Three trials of each test were done and the best trial used in the analysis. Every test was followed by a 3-5 min break. Table 1 shows the measurement variables and units.

Table 1 Descriptive Statistics for Performance Variables (n = 76)

Variable	Mean ± SD	Unit
Peak Velocity of Squat Jump (30% 1RM)	2.48 ± 0.29	m/s
CMJ Height	41.7 ± 12.3	cm
Lane Agility Test	12.5 ± 1.4	s
¾Court Sprint	3.52 ± 0.58	s

Pearson correlation coefficients (r) and their 90% confidence intervals (Fisher's z transformation) were computed between FMS total scores and each of the four performance indicators. Equivalence testing was implemented through the use of two one-sided test (TOST) procedure where the previously specified equivalence limit was $|r| = 0.30$. Less than a moderate magnitude correlation was defined as when the 90% CI lay completely inside the [-0.30, 0.30] range and when both one-sided *p*-values of TOST were < 0.05.

2.2 Study 2: Test-Retest Reliability and Intervention Effect

Subjects: In total, 60 exercise subjects (20.8 ± 1.8 years) who had not previously undergone FMS testing were put into either an intervention group (n = 30) or a control group (n = 30) at random. There are no significant differences in the two groups in terms of age, BMI, and baseline FMS total scores.

Experimental Design:

Pre-test (T1): On the first day, all subjects were subjected to the FMS testing battery, which was applied and scored by the same rater with the same qualifications as described in Study 1.

Intervention: After one week, the intervention group viewed together a 20 minutes standardized FMS movement education video providing clear direction in scoring criteria and typical compensatory movement patterns on each item. It was then followed by a 15 minute unloaded movement practice session that was similar to the test movements led by a physical therapist. The control group was given an equivalent-duration basketball highlight compilation to watch simultaneously without engaging in any form of physical activity.

Post-test (T2): After the intervention (intervention group) or a time of seated rest (control group), a second FMS evaluation was given immediately, and the test protocol was the same as that used in T1. Rater was blinded to the group allocation during the posttest.

Analysis of statistics: Mean and standard deviation values of FMS total scores at T1 and T2 were determined in every group. Pearson correlation coefficients (r), and intraclass correlation coefficients (ICC, two-way random, single measure, absolute agreement) were used to evaluate test-retest reliability by group. The mean change scores (T2-T1) between the two groups were compared using independent-samples t-test. In order to compare the homogeneity of variance of change scores, Levene's test was used. Fisher’s z-transformation was also used to determine whether the difference between the correlation coefficients of the two groups was statistically significant. The effect sizes were reported in terms of Cohen d. The significance level was defined as $\alpha = 0.05$.

3 RESULT

3.1 Study 1: Correlation Between FMS and Performance

The average score of FMS in total was 14.5 (SD = 3.0), with a range of 8-21. The Pearson correlation coefficients between FMS total scores and each of the four measures of athletic performance are reported in Table 2 (ranging between -0.04 and 0.05 and all p-values exceeding 0.05). All of the correlations had 90 percent confidence intervals that were both completely within the pre-determined equivalence limit of ± 0.30, and TOST procedure maximum one-sided p-values were all lower than 0.05, which means that FMS total scores had not significant correlation with such basketball-based athletic performance measures (no valid evidence).

Table 2 Correlation Coefficients and Equivalence Tests Between FMS Total Score and Performance Indicators (n = 76)

Performance indicator	Pearson r	90% CI (Lower, Upper)	Equivalence bound (±)	TOST p (max onesided)	Conclusion
Peak Velocity of Squat Jump (30% 1RM)	0.03	(-0.16, 0.22)	0.3	0.04	Equivalent noncorrelation
CMJ Height (cm)	0.05	(-0.14, 0.24)	0.3	0.03	Equivalent noncorrelation
Lane Agility Test (s)	-0.04	(-0.23, 0.15)	0.3	0.04	Equivalent noncorrelation
3/4Court Sprint (s)	0.01	(-0.18, 0.20)	0.3	0.04	Equivalent noncorrelation

3.2 Study 2: Stability of FMS Scores

Comparison of the baseline: There was no statistically significant difference between baseline FMS total scores of the intervention group and the control group (14.2 ± 2.9 versus 14.4 ± 3.0; $t = 0.26$, $p = 0.80$).

The control group showed a large test retest correlation of the total scores of FMS ($r = 0.91$, ICC = 0.90) and average change score was -0.1 ± 1.3 and 93 per cent of the participants had changes less than 2 points. However, the intervention group had a considerably lower test retest correlation than the control group ($r = 0.47$ vs. 0.91, Fisher’s $z = 3.84$, $p < 0.001$; ICC = 0.45) with a mean change score of 0.2 ± 2.9. It is important to note that the standard deviation of the changes scores of the intervention group were significantly higher (2.9) as compared to the control group (1.3) (Levene s’ $F = 12.34$, $p < 0.001$), where the individual change ranged between - 5 and + 6. The results are given in more detail in Table 3.

Table 3 Pre- and Post-Test FMS Total Scores and Change Scores of the Two Groups

Group	n	T1	T2	Change (T2–T1)	Pre-post r	ICC
Intervention	30	14.2 ± 2.9	14.4 ± 3.6	0.2 ± 2.9	0.47	0.45
Control	30	14.4 ± 3.0	14.3 ± 2.9	-0.1 ± 1.3	0.91	0.90

4 DISCUSSION

The current research examined the FMS based on two main factors, which are predictive validity and test-retest reliability. Study 1 showed that absolute correlation coefficients between FMS total score and sprint time on ¾-court, counter-movement jump height, highest velocity during loaded squat jump, and pro-agility test in 76 national-level 1 and above basketball athletes were less than 0.10. The equivalence testing (with a predetermined threshold of $|r| = 0.30$) indicated that all relationships were below a moderate effect level (TOST one-sided $p < 0.01$). These results provide a very strong empirical indication that the FMS does not predict basketball-specific performance, as was found by Parchmann et al. and Silva et al. [5-6]. The physiological reason behind this null result is that the FMS evaluates slow, self-paced movements that are controlled by low-threshold motor control and flexibility, and explosive performance requires maximum power production and stretch-shortening cycle operation at high speeds - two types of neuromuscular control that are essentially incompatible [11]. This statement is also supported by sport-specific training literature; for example in judo, various kinds of weight-bearing training have very specific transfer effects on explosive performance, suggesting that there is no simple direct linear transformation between fundamental screening movements and sport-specific movement patterns [12].

The second study uncovered another serious problem with the FMS scores: their variability. The control group had a high level of test-retest reliability (ICC=0.90) since they did not get any treatment, and thus, the FMS score was consistent under the same testing conditions. Nevertheless, a combination of one educational video lasting 20 minutes and 15 minutes of practice on the test movements minimized the test-retest correlation in the intervention group by 0.47, where the standard deviation of change scores rose to 2.9 - a considerable increase compared to the control group. It means that FMS scores are extremely vulnerable to recent training or education programs which cause significant swings in the absence of fundamental changes to the participants ability to move. This learning effect or intervention responsiveness is a critical threat to the validity of the FMS as a reliable screening tool, since clinicians cannot easily determine if a change in the score indicates a real movement improvement or simply an adaptation to the testing processes or instruction effects. The same issues are common in clinical exercise interventions; e.g., studies on the optimization of individualized aerobic exercise in type 2 diabetes mellitus patients have shown that credible attribution of intervention effects requires dynamic modeling using highly stable physiological baseline indicators [13], and if the baseline measures themselves are highly variable, then the interpretation of later intervention effects becomes unfounded. Simple instructions of movement also improved FMS scores at once, as reported by Frost et al., in support of the results of the current study [9].

Both the integrated results of the two studies have significant practical implications. Firstly, coaches and medical practitioners in sports must not assess the explosive power, speed and technical ability of athletes through the FMS, nor can they eliminate an athlete on the basis of the low score in the FMS. Secondly, due to the volatility of FMS scores as a screening measure of injury risk, it would be prudent to average several test outcomes or carefully control the previous exposure and training of the participants before testing them. Further research is needed in order to incorporate injury surveillance data by determining if the size of individual score variability, not the absolute FMS score, could be used as a more sensitive indicator of injury risk, and to create evaluation procedures that are more resistant to effects of practice or learning. The future assessment models could go beyond a screening score towards more ecologically valid models based on functionality; e.g., a recent proposal of a lifelong fitness model combining exercise snacks with neuromuscular fall prevention approaches focuses on evaluating changes in the real capacity of movement over time instead of using fixed and one-time scores [14].

A number of constraints in the present research must be noted. The cross-sectional nature of Study 1 does not allow making any causal inferences; moreover, the sample size being a pilot study was quite small and more large-scale experimental groups need to be conducted to confirm the weak or null relationship between the FMS scores and athletic performance. The intervention used in Study 2 was short-term education only without long-term follow-up and the subjects were active sportsmen and women, not elite athletes so the results cannot be generalized to high-level athlete populations. Furthermore, the fact that FMS scores are an unstable injury screening tool means that one test score can be inaccurate, which is especially problematic in adolescent athletes, who need very specific injury prevention measures as the training monitoring and rehabilitation decision-making in this population rely highly on the stability and reliability of assessment tools [15]. Also, the current research did not assess directly the presence of injuries, so it was impossible to confirm the clinical effects of score instability in a direct way.

5 CONCLUSION

The Functional Movement Screen (FMS), which has no predictive validity on basketball player performance, has a total score that has insignificant relationship with the overall physical conditioning attributes such as speed, power and agility, and also sport-specific technical skills. Moreover, the FMS scores are extremely vulnerable to the significant changes after the short educational interventions in physically active groups, where the test-retest reliability is greatly affected by the intervention. The presence of these two core flaws implies that the FMS should have no place in the choice of athletes or in assessing their performance, and that it relies on the testing context under which it was administered as a screening device. This result is not based only on its low psychometric characteristics but can also be considered in line with the sociological criticism of the FMS as a constructed assessment tool [3]. It is recommended that practitioners be careful in interpreting FMS findings and opt to use other more rigorous methods of pre-exercise assessment.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

ACKNOWLEDGEMENT

We gratefully acknowledge the academic guidance provided by the Chinese TikTok account “Strength and Condtioning Pope”.

REFERENCES

[1] Cook G, Burton L, Hoogenboom B J. Pre-participation screening: The use of fundamental movements as an assessment of function—Part 1. North American Journal of Sports Physical Therapy, 2006, 1(2): 62-72.

[2] Kiesel K, Plisky P J, Voight M L. Can serious injury in professional football be predicted by a preseason functional movement screen? North American Journal of Sports Physical Therapy, 2007, 2(3): 147-158.

[3] Li F, Huang Y, Zhou M, et al. Ritual, commodity, and incubator of pseudoscience: A sociological critique on the proliferation of the Functional Movement Screen (FMS). Frontiers in Arts, Humanities & Social Sciences, 2026, 1(3).

[4] Skorupiński P M. American Educational Research Association, American Psychological Association, National Council on Measurement in Education, standards for educational and psychological testing. Kwartalnik Pedagogiczny, 2015, 238(4): 201-203.

[5] Parchmann C J, McBride J M. Relationship between functional movement screen and athletic performance. The Journal of Strength & Conditioning Research, 2011, 25(12): 3378-3384.

[6] Silva B, Clemente F M, Martins F M. Associations between functional movement screen scores and performance variables in surf athletes. The Journal of Sports Medicine and Physical Fitness, 2017, 58(5): 583-590.

[7] Smith C A, Chimera N J, Wright N J, et al. Interrater and intrarater reliability of the functional movement screen. The Journal of Strength & Conditioning Research, 2013, 27(4): 982-987.

[8] Onate J A, Dewey T, Kollock R O, et al. Real-time intersession and interrater reliability of the functional movement screen. The Journal of Strength & Conditioning Research, 2012, 26(2): 408-415.

[9] Basar M J, Stanek J M, Dodd D D, et al. The influence of corrective exercises on functional movement screen and physical fitness performance in army ROTC cadets. Journal of Sport Rehabilitation, 2019, 28(4): 360-367.

[10] Liu P, Liu N, Li F. Diagnosis and management of splenic tuberculosis: A case report and literature review. Frontiers in Medicine, 2025, 12: 1622794.

[11] Suchomel T J, Nimphius S, Stone M H. The importance of muscular strength in athletic performance. Sports Medicine, 2016, 46(10): 1419-1449.

[12] Li F. Comparison of the effects of the different forms of the weight bearing training on the explosive power migration: A case study of Judo. СОВРЕМЕННАЯ АЗИЯ: ПОЛИТИКА, ЭКОНОМИКА, ОБЩЕСТВО, 2026, 1(10): 47.

[13] Li F, Zeng Z. Personalized optimization of aerobic exercise for type 2 diabetes mellitus patients based on improved Q-learning and dynamic physiological modeling. Fifth International Conference on Information Technology and Contemporary Sports (TCS 2025), SPIE, 2026, 14115: 156-165.

[14] Li F, Li Z, Wang L, et al. Redefining lifelong fitness: A practical framework combining exercise snacks and neuromuscular fall prevention strategies. Frontiers in Arts, Humanities & Social Sciences, 2026, 1(3).

[15] Li F. Research on the prevention and rehabilitation of sports injuries in adolescent sports training. Learning Weekly-Teaching and Research, 2025(15).