World Journal of Clinical Sciences
Print ISSN: 2960-0278
Online ISSN: 2960-0286
DOI: https://doi.org/10.61784/wjcs3021

COMPENSATION–IRRITATION MATRIX (CIM): RECONSTRUCTION AND MATHEMATICAL AND ALGORITHMIC VALIDATION OF THE FUNCTIONAL MOVEMENT SCREEN (FMS) SCORING SYSTEM

Chen Zhao[1], Rui He[2], WeiBo Shang[2], JunJie Zheng[3], RongJin He[4], FangHao Li[5*]
[1]*Northwestern Polytechnical University, Xi'an 710000, Shaanxi, China.*
[2]*Shanghai Shuangyou Es Sports Culture Development Co., Ltd., Shanghai 200000, 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: The Functional Movement Screening (FMS) composite score (0-21 points) is achieved merely by adding up the scores of seven tests, and by implication, it assumes that all deductions (pain, severe restriction, compensation) are equal in numerical terms. Nevertheless, pain is an indicator of local tissue irritation, and compensation is a sign of movement control impairment, which is a significantly different state when it comes to the direction of clinical intervention. Putting them together as one score reduces interpretability, and can also reduce predictive validity. The proposed study introduces the Compensation-Irritation Matrix (CIM) that divides the raw FMS data into two independent dimensions: the Compensation Index (C, the level at which the rest of the movements are preserved) and the Irritation Index (I, the number and location-weighted load of pain). We confirm the theoretical benefits of CIM over the conventional total-score approach by mathematically counting the full 16,384 possible theoretical score combinations. It is found that given the traditional total-score approach, two random people with the same total score have a 71.4% likelihood of having various numbers of painful items (i.e., same score, but varying quality). Contrastingly, the CIM four-quadrant classification (green-yellow-orange-red) makes perfect sense in every combination with no classification issues. Two case examples indicate that two people with the same total score of 14 are identified as yellow (C = 2.00, I = 0, extensive compensation) and orange (C = 2.80, I = 2.0, localized irritation with high movement quality) on CIM, which implies totally opposite intervention orientation. CIM addresses the dimensional confusion brought about by combining the concepts of compensation and irritation in the traditional total-score method providing a logically consistent alternative model of clinical application of the FMS.
Keywords: Functional Movement Screen; Compensation-irritation matrix; Injury risk prediction; Scoring algorithm; Measurement

1 INTRODUCTION

Functional Movement Screen (FMS) is among the most popular standardised movement quality assessment instruments in sports medicine, strength and conditioning, and military physical training. Seven standardised movements are included in the test: Deep Squat, Hurdle Step, In-Line Lunge, Shoulder Mobility, Active Straight Leg Raise, Trunk Stability Push-Up, and Rotary Stability, which are scored on an ordinal scale of 0-3 points (0 = pain, 1 = severe restriction, 2 = compensation, 3 = full performance). All seven scores are mathematically totaled to provide a combined total score between 0 and 21. It is also commonly applied in adolescent sports training to screen injuries and monitor the quality of training [1].

FMS was developed to represent an individuals' overall movement quality by a single numeric value and thus predict the risk of injury. The cut-off of the total score of ≤ 14 has been extensively researched. According to a systematic review by Bonazza et al., low-scoring group had significantly greater injury rates compared to the high-scoring control group [2]. In a longitudinal study involving 2,154 male soldiers, Bushman et al. discovered a relative risk (RR) of 1.60 in individuals who scored ≤ 14 [3]. Nevertheless, the meta-analysis by Moran et al. (24 studies) specifically stated that the relationship between the FMS composite score and the future injury cannot be strong enough to justify the use of the FMS composite score as an independent tool in predicting injuries [4]. Trinidad-Fernández et al. stated a pooled odds ratio of 1.86 with high heterogeneity (I^2 = 70%) [5]. Sensitivity, specificity, and AUC were reported as 24.7%, 85.7%, and 0.587 respectively in the meta-analyses of Dorrel et al. [6]. These findings suggest that predictive validity of the FMS total score is extremely debatable. Apart of the instability of predictive validity, other recent authors have also indicated that the accelerated growth of the FMS across the globe does not have adequate theoretical backing, and that the excessive commercialization and instrumentalisation of the same could have gone against its initial aim as a movement quality measurement tool [7].

Besides the questionable predictive validity, the additive scoring system of the FMS has a more fundamental psychometric problem: all deductions are assumed to be equal. When using the FMS testing, scores of 0 (pain), 1 (very restricted) and 2 (compensation) are all considered as deficit points and summed up. Nevertheless, the clinical significance of these three conditions is entirely distinct - 0 points mean pain (to be evaluated in the rehabilitation process), whereas 1 and 2 points are compensations (to be treated with corrective exercise). Combining such qualitatively different sorts of information into a combined additive formula significantly reduces the usefulness of the overall score in making clinical decisions.

Available empirical data also indirectly confirms this concern. In a prospective study of 2,154 male soldiers, Alemany et al. observed that people who had pain on the FMS were 1.62 times more likely to be injured compared to those having a total score of 14 or lower (RR = 1.58) and the quantity of painful items displayed a dose-response association with the incidence of injury [8]. Fuller et al., who examined 439 junior Australian football players, provided more direct evidence: 38.7% of players said they experienced pain when performing the FMS; pain also had a significant impact on the composite score (original scoring 14.9 versus traditional scoring 13.6, $p < .001$); however, the composite score was not significantly correlated with later injury risk ($p \geq .500$) [9]. They indicate that the dimension of injury risk is not related to the concept of movement quality but rather to pain. Jointly with compensation, it makes contamination of the purity of the notion of compensation and dilution of the predictive value of the notion of irritation.

According to the aforementioned analysis, we suggest the Compensation–Irritation Matrix (CIM) as a framework of the reconstruction of the FMS scoring system. The main concept is to disentangle the raw FMS scores as two separate dimensions: (1) the Compensation Index (C) - indicating the true level of compensation during movement without pain; and (2) the Irritation Index (I) - representing the count of painful elements and local weight. This paper mathematically identifies all theoretically possible score combinations in order to examine the theoretical benefits of CIM compared with the conventional total-score approach in terms of cross-situational consistency, retention of information and clinical interpretability. We predict that CIM will remove the dimensional confusion generated by combining both the concepts of compensation and irritation in the traditional additive scoring and will offer a logically consistent new paradigm of clinical interpretation of the FMS.

2 METHODS

2.1 Study Design

It is a theoretical derivation and mathematical model research with no use of human participants or data collection through experiments. The full enumeration approach was applied to all theoretical score combinations of the seven FMS subtests, and the performance of the traditional total-score method and CIM were analyzed in a common mathematical context.

2.2 FMS Scoring Rules and Variable Definitions

Based on the FMS manual of standard operation, all seven subtests (Deep Squat, Hurdle Step, In-Line Lunge, Shoulder Mobility, Active Straight Leg Raise, Trunk Stability Push-Up, Rotary Stability) are evaluated separately, and the individual scores of the items $x_i \in \{0, 1, 2, 3\}$ ($i = 1, 2, ..., 7$). The composite total score is defined as $S = \sum_{i=1}^{7} x_i$, which takes values in the range of 0 to 21. Suppose k is the count of zero-scored items (i.e. the count of painful items).

2.3 Construction of the Compensation–Irritation Matrix (CIM)

2.3.1 Compensation index (C)

Let us suppose that a person has k items whose scores are 0 because of pain ($k \geq 0$) and the rest, i.e., $7 - k$, are scored at least 1. The Compensation Index is defined as:

$$C = \frac{S}{7-k} \tag{1}$$

When $k = 7$ (every item painful), the value of C cannot be defined (zero denominator); therefore, this person ought to be immediately labeled as systemic irritation.

The value of C is between 1 and 3, and larger values indicate that the movements that could be performed have less compensation and higher quality of movement. The main benefit of C is the absence of painful items in the denominator as C is simply the indicator of quality of the remaining movements. Our cut-off level of the high-quality movement in this study was $C = 2.5$, i.e., the average score of the remaining items was 2.5 or more, which clinically means that at least half of the items scored the best possible score of 3 and the other scores were no lower than 2, which gives an intuitively clinical significance.

2.3.2 Irritation index (I)

The Irritation Index comprises two dimensions:

The amount of painful items (k): the summed number of items that are scored 0, which is the range of items being 0-7.

The site-specific weight (w): There is biomechanical and epidemiological evidence that pain at various sites can have dissimilar injury risk implication. We use the weighting system given in Table 1 as per the results of Alemany et al. and Fuller et al. [8-9].

Table 1 Site Classification, Included FMS Tests, Weight Coefficients, and Rationale

Site category	Included FMS tests	Weight coefficient (w)	Rationale
Core / spine	Trunk Stability Push-Up, Rotary Stability	1.2	Core control deficits associated with multi-site injuries
Lower limb	Deep Squat, Hurdle Step, In-Line Lunge, Active Straight Leg Raise	1.0	Irritation of weight-bearing joints of the lower extremity
Upper limb	Shoulder Mobility	0.8	Relatively lower injury risk for the upper extremity

The site-weighted Irritation Index is defined as:

$$I = \sum_{j=1}^{k} w_{(j)} \tag{2}$$

where $w_{(j)}$ denotes the site weight of the j -th painful item. The I range is between 0 and 8.4 (the highest possible when all the seven items are painful and all core sites: 7 * 1.2 = 8.4).

2.3.3 Two-dimensional risk matrix

The combination of C (horizontal axis, 1-3) and I (vertical axis, 0) gives us four-quadrant risk matrix as shown in Table 2.

Table 2 CIM Four-Quadrant Risk Matrix

Quadrant	Compensation Index (C)	Irritation Index (I)	Risk Level	Clinical implication
Green	$C \geq 2.5$	$I = 0$	Low	High movement quality, no pain
Yellow	$C < 2.5$	$I = 0$	Low–Middle	Widespread compensation without local irritation → basic conditioning/movement control intervention
Orange	$C \geq 2.5$	$I > 0$	Middle–High	Local tissue irritation but other movements perfect → rehabilitation/medical referral
Red	$C < 2.5$	$I > 0$	High	Widespread compensation combined with local irritation → comprehensive intervention

2.4 Mathematical Enumeration and Data Generation

In order to confirm the logical characteristics of CIM, we have conducted an exhaustive list of all theoretical combinations of scores. All seven tests are taken at one of four levels (0-3) and this creates a total of 4^7 = 16,384 combinations. To every combination, we wrote down the seven raw scores, S, k, and derived C, I, and the four-quadrant classification. Traditional total score classification ($S \geq 14$ is low risk, $S < 14$ is high risk) was also recorded to evaluate information retention and classification consistency across both methods.

2.5 Evaluation Metrics and Comparison Framework

When comparing the theoretical performance of the traditional total-score method and CIM we have identified three dimensions of evaluation:

The information retention rate: The traditional total-score technique reduces seven item scores to one scalar (0-21) and thus eliminates the dimensional data of compensation-pain. Quantification: we computed, over all 16,384 profiles, the chance that two randomly chosen profiles with the same total score will have different k values (the probability of having the same score but differing in quality). A higher probability means more information loss.

(2) Consistency across situations: With the traditional approach, the requirements to achieve the same overall score are not the same in the various contexts of k; CIM will automatically ensure cross-situational consistency by stipulating that $C = S / (7 - k)$.

(3) Clinical interpretability: By using case analyses, we were able to compare if the two methods would give logically consistent risk classifications, and offer distinct intervention directions.

3 RESULTS

3.1 Basic Distribution of All Theoretical Score Combinations

This was based on the full enumeration where out of 16,384 theoretical profiles, pain-free situations (k = 0) made up 2,187 or 13.35% of the population whereas painful situations ($k \geq 1$) made up 14,197 or 86.65%. The distribution of the subgroups is presented in Table 3.

Table 3 Distribution of Theoretical Score Profiles across Different Pain Scenarios

Scenario	Number of zero-score items (k)	Formula for number of combinations	Number of profiles	Percentage (%)
Pain-free	0	3^7	2,187	13.35
Painful	1	$C_7^1 \times 3^6$	5,103	31.15
Painful	2	$C_7^2 \times 3^5$	5,103	31.15
Painful	3	$C_7^3 \times 3^4$	2,835	17.30
Painful	4	$C_7^4 \times 3^3$	945	5.77
Painful	5	$C_7^5 \times 3^2$	189	1.15
Painful	6	$C_7^6 \times 3^1$	21	0.13
Painful	7	$C_7^7 \times 3^0$	1	0.006

Theoretical space is mostly occupied by painful scenarios (86.65%) and combined with the significant presence of pain in real clinical practice (38.7% according to Fuller et al), it reinforces the necessity of separating pain and compensation [9].

3.2 The "Same Score, Different Quality" Phenomenon in the Traditional Total-Score Method

The traditional approach will always lead to the loss of information when compressing seven scores into one scalar. We computed the probability that two randomly chosen profiles of equal total score do not have the same *k* values. Out of 16,384 profiles, when grouped by total score S, 9,234 pairs of profiles shared the same total score. Of this number, 6,595 pairs had different *k* values, or:

$$P\ (\text{same score, different quality}) = \frac{6{,}595}{9{,}234} \times 100\% = 71.4\% \qquad (3)$$

It implies that in the conventional approach, two randomly chosen people with equal overall scores have greater than 70 percent probability of having entirely different pain burdens (*k* values). As an example, the most frequent total score of 14 has 393 profiles with $k = 0$, 210 profiles with $k = 1$, and 36 profiles with $k = 2$, three distinct clinical situations defined as low risk (14) but having completely different clinical states (see Section 3.5 to see examples of cases).

3.3 Cross-Situational Consistency of CIM

The CIM framework definition of $C = S / (7 - k)$ by nature negates the denominator-shrinkage effect. In Table 4, you can see the least overall score needed to reach the same C threshold ($C \geq 2.5$) with various values of k.

Table 4 Minimum Total Score Required to Achieve $C \geq 2.5$ under CIM

k	Remaining items	Minimum total score S required for C ≥ 2.5	Corresponding absolute total score
0	7	$S \geq 17.5 \rightarrow S \geq 18$	18
1	6	$S \geq 15.0 \rightarrow S \geq 15$	15
2	5	$S \geq 12.5 \rightarrow S \geq 13$	13
3	4	$S \geq 10.0 \rightarrow S \geq 10$	10
4	3	$S \geq 7.5 \rightarrow S \geq 8$	8
5	2	$S \geq 5.0 \rightarrow S \geq 5$	5
6	1	$S \geq 2.5 \rightarrow S \geq 3$	3
7	0	Undefined (all painful)	—

Major result: The condition of the compensation level under CIM is entirely uniform when considering various values of k - $C \geq 2.5$ always implies that on average, the rest of the movements come up to 2.5 or higher irrespective of the number of painful items that the person has. It fully eliminates mathematical artifacts in the conventional approach where 14 has various interpretations depending on the situation.

3.4 CIM Four-Quadrant Risk Classification

In CIM it is possible to map all 16,384 hypothetical profiles to one of the four quadrants (the case of $k = 7$ is indicated with a separate label of systemic irritation). Table 5 illustrates the spread of profiles among quadrants.

Table 5 Theoretical Distribution of CIM Four-Quadrant Risk Matrix

Quadrant	Criteria	Number of profiles	Percentage (%)	Risk level
Green	$C \geq 2.5, I = 0$	1,134	6.92	Low
Yellow	$C < 2.5, I = 0$	1,053	6.43	Low–Middle
Orange	$C \geq 2.5, I > 0$	4,267	26.04	Middle–High
Red	$C < 2.5, I > 0$	9,929	60.59	High
Special	k = 7 (all painful)	1	0.006	Systemic irritation
Total	—	16,384	100	—

It can be observed in Table 5 that over 86 percent of all theoretical combinations in the table have at least one painful item (which is consistent with Table 3) and the largest portion of them (60.59 percent) is categorized as red (high risk)

because the two most frequent items (widespread compensation $C < 2.5$) are present simultaneously. There are only 6.92 percent of profiles that are actually low-risk (no pain and high movement quality).
However, the CIM risk classification makes logical sense by itself: the fact that someone has been categorized as orange ($C \geq 2.5$, $I > 0$) implies that other movements are of high quality but that they have local pain, which is a clinically feasible condition (e.g., a person who suffers from subacromial impingement syndrome and experiences pain in Shoulder Mobility but can perform Deep Squat and other lower-limb movements flawlessly). In the conventional approach, the same condition may be incorrectly categorized as high-risk due to the fact that the sum score is less than 14, or low-risk because it is over 14, based solely on the number of painful items (i.e., the denominator).

3.5 Case Comparisons

The Table 6 shows four typical cases that allow comparing the risk classification using the traditional total score approach with the CIM approach directly.

Table 6 Representative Cases: Risk Classification by Traditional Total-Score Method vs. CIM

Case	Seven item scores ($x_1 \ldots x_7$)	k	S	Traditional classification	C	I	CIM classification
A	3,3,3,3,2,2,2	0	18	Low risk (>14)	2.57	0	Green
B	3,3,3,3,2,0,0	2	14	Low risk (=14)	2.80	2.0	Orange
C	2,2,2,2,2,2,2	0	14	Low risk (=14)	2.00	0	Yellow
D	0,0,0,3,3,3,3	3	12	High risk (<14)	3.00	3.6	Orange

Case B vs. Case Cillustrates the core value of CIM:
Case B (Two items are painful and the rest are five including four 3s and one 2, sum of scores is 14): The conventional approach is labeled as low risk (14 or higher), but CIM indicates $C = 2.80$ (minimal compensation) and $I = 2.00$ (two irritation points) - Orange (middle-high risk). Clinical application: this athlete has nearly perfect movements, but with two localized areas of pain; the most important is rehabilitation, not overall fitness training.
Case C (no pain, all seven items scored 2, total score 14): The conventional approach is also considered to be the low-risk category (= 14), but CIM gives $C = 2.00$ (widespread compensation) and $I = 0$ (no irritation) - Yellow (low-mid risk). Clinical implication: this person has general compensatory movement patterns without local tissue irritation; the best intervention is basic movement control training.
The traditional total-score method places Cases B and C into the same risk category (both "low risk" with score ≥ 14), whereas CIM correctly distinguishes their completely different clinical intervention needs. Case D (three painful objects and the rest four all 3s, total score 12): the traditional method will classify it as high risk (< 14), but CIM scores $C = 3.00$ (perfect remaining movements) and $I = 3.6$ (three irritation points and central weighting) → Orange (middle high risk). CIM clearly demonstrates that the intervention direction must be local rehabilitation, not general movement correction because the other movements are already perfect and there is no compensation to correct.

3.6 Summary of Results

These findings mathematically support the fact that: (1) the old fashioned total-score approach has significant vulnerability to the problem of the same score but different quality- people with the same total score have a greater than 71.4% probability of having distinct pain burdens; (2) CIM, as it separates the concepts of compensation and irritation into two separate dimensions, does not confuse dimensions at all, and its four-quadrant categorization is logically sound in all 16, 384 possible theoretical combinations; (3) case analyses reveal that CIM offers different guidance on intervention in clinical scenarios that the conventional approach is unable to differentiate.

4 DISCUSSION

4.1 Main Findings

The current research, by numerically enumerating all 16,384 possible theoretical FMS score combinations, confirmed the theoretical benefits of the Compensation-Irritation Matrix (CIM) in contrast to the conventional composite total-score approach. The main results include:
First, the traditional total-score method suffers from severe "same score, different quality". Of all possible theoretical combinations, there is a 71.4 % chance that any two randomly-chosen profiles with the same overall score will differ in their count of painful items (k). It implies that greater than 70% of people who are identical in score have entirely dissimilar clinical features; some can be low scorers because they are widely compensated (but do not have pain), whereas others can be low scorers because of localized pain (but all other movements are perfect). Their treatment is clinically irrelevant.
Second, CIM completely eliminates dimensional confusion by decoupling "compensation" and "irritation". The equation $C = S / (7 - k)$ is by nature regulating the presence of the number of painful items in the denominator, thus the criterion of degree of compensation is made uniform between different values of k. The Irritation Index I is an independent measure of pain quantity and site, which adds another independent clinical decision-making level to local

tissue irritation. This two types of information are combined into a four-quadrant risk matrix to provide various intervention directions.

Third, CIM is significantly superior to the traditional method in clinical interpretability. The comparison of Cases B and C (with the same total score of 14, but labeled as orange and yellow in CIM respectively) clearly demonstrates the inability of the traditional approach to differentiate between the two fundamentally different clinical conditions, namely, between the so-called local irritation with a high-quality movement and widely spread compensation without pain, whereas CIM does not only discriminate between them but also provides explicit intervention advice concerning each of them.

4.2 Theoretical Foundation of CIM and Consistency with Existing Evidence

Two observations which are theoretically constructed in CIM are well supported by empirical research:

The seven FMS tests are not necessarily unidimensional. Kazman et al. reported a Cronbach alpha of merely 0.39 whereas Koehle et al. noted a bifactor structure [10-11]. As the seven tests are not assessing one underlying factor, summing them is not supported by psychometric principles. Indeed, multi-dimensional decomposition approaches to assess movement quality are often used in sport-specific scenarios and not a composite score. To illustrate, research on the influence of weight-bearing exercise on explosive power transfer has demonstrated important sport-specific distinctions in movement patterns that a one-size-fits-all total score cannot reflect [12]. The CIM framework rejects the indefensible postulate that the single total score is indicative of overall movement quality and instead accepts the idea that the FMS can be used to measure two qualitatively distinct pieces of information, namely quality of movement control (compensation) and condition of tissue irritation (pain) which are processed independently.

(2) Pain as an injury risk factor is a dimension unrelated to quality of movement. Alemany et al., however, showed that the correlation between pain and injury risk (RR = 1.62) was higher compared to low total score (RR = 1.58), and the number of painful items was dose-response related. The FMS composite score was also significantly affected by pain, although the composite score did not relate to injury risk according to the report by Fuller et al. [8-9]. These results indicate very strongly that adding up the pain and compensation data into one additive equation would defile the immaculateness of the word, compensation, and weaken the predictive power of the word, irritation. CIM disunites the two thus addressing these empirical facts directly.

4.3 Clinical and Research Implications

CIM proposal will have the following consequences on clinical practice and further study of the FMS:

Giving different intervention directions to clinical decision-making are provided. (Traditionally, an individual with a low score would be informed that they have an increased injury risk, although the intervention direction would be unknown.) The four-quadrant matrix of the CIM explicitly identifies four different clinical scenarios, namely, basic movement training (yellow), local rehabilitation referral (orange), comprehensive intervention (red), and maintenance training (green), which upgrades the FMS as a risk screening tool to a clinical decision support tool.

(2) Offering testable hypotheses in future studies. The main hypotheses of CIM (a) Compensation Index C is negatively related to injury risk; (b) Irritation Index I is positively related to injury risk; (c) the interaction between C and I (maximum risk in the red quadrant) can also all be empirically examined through prospective cohort studies. In relation to empirical testing methods, the conventional multivariate regression might not sufficiently reflect non-linear interactive relationships between C, I and their trajectories across time. According to recent studies on the optimalisation of personalized exercise prescriptions based on reinforcement learning and dynamic models of physiology, multi-dimensional individual assessment data can be integrated into algorithmic decision systems to produce more accurate risk stratifications and intervention matches [13]. This innovation in methodology offers a viable technical model of the future empirical confirmation of the CIM paradigm, e.g., C, I, and their interaction coefficients could serve as state variables in constructing dynamic injury risk prediction models and, say, Q-learning could be applied to optimize personalization of cut-off values.

(3) Clarifying the reason behind the poor predictive reliability of the conventional total-score approach. We found that the fact that the same total score may be associated with a different quality is mathematically explained by the heterogeneity ($I^2 = 70\%$) noted in the meta-analysis of Moran et al. various studies report diverse pain prevalence rates among their participants, which means that the same total score can represent entirely different clinical features across the samples, which makes it unpredictable.

The predictive validity of CIM has not been validated through future empirical research, although the rigor of its mathematical logic is a sound theoretical basis of clinical use.

(5) In a wider context, the two-dimensional nature of the classification logic of CIM can be incorporated into the life-long health promotion and long-term functional mobility maintenance programmes. The latest reports suggest that the whole-life fitness models which incorporate both the strategy of exercise snacks and neuromuscular fall prevention methods are possible [14]; the specific risk stratification offered by CIM is anticipated to provide an exact entry point to these comprehensive frameworks.

4.4 Limitations

As a theoretical deduction and mathematical simulation study, this work has several limitations:

First, we have taken it as a given that all possible combinations of scores are equally likely. The actual distribution of FMS item scores in clinical practice is not even. Nevertheless, the objective of the present research was not to model realistic score distributions, but rather to examine the logical composition of the scoring system per se-regardless of non-uniform distributions, the algebraic truth of same score, different quality would be preserved.
Second, the weight system used in CIM (site weights w) has not been experimentally adjusted. The proposed weights (core 1.2, lower limb 1.0, upper limb 0.8) were derived using small amounts of epidemiological data, and their optimum values should be confirmed by subsequent large-sample research.
Third, the value of C = 2.5 as a cut-off is also to be tested empirically; there exists an optimal level that can differ among different populations and sports.
Fourth, we have not used real injury outcome data in external validation; predictive validity of CIM remains to be examined in future prospective cohort investigations.
Fifth, The CIM framework only considers the Irritation Index (I) as a measure of local musculoskeletal tissue irritation. Nevertheless, in clinical setting, the pain measured in FMS test might be influenced by the visceral or systemic causes that cause referred pain or movement restriction. As an example, early stage splenic tuberculosis and other abdominal infectious diseases are characterized by vague abdominal or shoulder pain, which can confuse the accurate diagnosis of FMS painful items [15]. Hence, we suggest that prior to using CIM as part of clinical decision making, it would be prudent to integrate a simple physical examination or relevant medical imaging to initially determine the origin of the pain and remove non-musculoskeletal sources that can affect the risk stratification process. It is a realistic threshold which needs to be crossed by this framework in its transition between being a mathematical model and a clinical implementation.

5 CONCLUSIONS

The present study, by means of mathematical numbering of all 16,384 theoretical FMS score combinations, suggests the Compensation-Irritation Matrix (CIM) as the reconstruction model of the FMS scoring system. The main conclusions are:
(1) The conventional FMS composite total score approach is subject to extreme same scores but different quality problems: two persons of the same total score have no less than 71 percent chance of experiencing different levels of pain, which means that the same score may mean quite different things clinically between people.
(2) CIM also separates the concepts of compensation (Compensation Index C) and irritation (Irritation Index I) into two separate dimensions, which completely removes any dimensional ambiguity, and has a four-quadrant classification that is logically coherent in all of the 16,384 hypothetical combinations.
(3) The four-quadrant matrix of CIM (green-yellow-orange-red) gives different directions of intervention in different clinical situations and upgrades the FMS to be a clinical decision support tool rather than a risk screening tool.
(4) The predictive validity of CIM needs to be tested empirically in advance, but the mathematical logic of this approach has already provided a strong theoretical basis to its clinical usage.

COMPETING INTERESTS

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

REFERENCES

[1] Li F. Research on the prevention and rehabilitation of sports injuries in adolescent sports training. Learning Weekly-Teaching and Research, 2025(15).
[2] Bonazza N A, Smuin D, Onks C A, et al. Reliability, validity, and injury predictive value of the functional movement screen: A systematic review and meta-analysis. The American Journal of Sports Medicine, 2017, 45(3): 725-732.
[3] Bushman T T, Grier T L, Canham-Chervak M, et al. The functional movement screen and injury risk: Association and predictive value in active men. The American Journal of Sports Medicine, 2016, 44(2): 297-304.
[4] Moran R W, Schneiders A G, Mason J, et al. Do Functional Movement Screen (FMS) composite scores predict subsequent injury? A systematic review with meta-analysis. British Journal of Sports Medicine, 2017, 51(23): 1661-1669.
[5] Trinidad-Fernandez M, Gonzalez-Sanchez M, Cuesta-Vargas A I. Is a low Functional Movement Screen score (≤ 14/21) associated with injuries in sport? A systematic review and meta-analysis. BMJ Open Sport & Exercise Medicine, 2019, 5(1): e000501.
[6] Dorrel B S, Long T, Shaffer S, et al. Evaluation of the functional movement screen as an injury prediction tool among active adult populations: A systematic review and meta-analysis. Sports Health, 2015, 7(6): 532-537.
[7] 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).
[8] Alemany J A, Bushman T T, Grier T, et al. Functional Movement Screen: Pain versus composite score and injury risk. Journal of Science and Medicine in Sport, 2017, 20: S40-S44.

[9] Fuller J T, Lynagh M, Tarca B, et al. Functional movement screen pain location and impact on scoring have limited value for injury risk estimation in junior Australian football players. Journal of Orthopaedic & Sports Physical Therapy, 2020, 50(2): 75-82.

[10] Kazman J B, Galecki J M, Lisman P, et al. Factor structure of the functional movement screen in marine officer candidates. The Journal of Strength & Conditioning Research, 2014, 28(3): 672-678.

[11] Koehle M S, Saffer B Y, Sinnen N M, et al. Factor structure and internal validity of the functional movement screen in adults. The Journal of Strength & Conditioning Research, 2016, 30(2): 540-546.

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