Logía, educación física y deporte.
2026, 7(1), 70-84.
ISSN: 2695-9305.

Validity and Reliability of using Modern Technologies (Jump MD-Flexometre) in Evaluating footballers' explosive force and flexibility. Field Study on M'sila Football Teams "Seniors"

Validez y fiabilidad del uso de tecnologías modernas (Jump MD-Flexometer) en la evaluación de la fuerza explosiva y flexibilidad de los futbolistas. Estudio de campo sobre los equipos de fútbol M'sila "Seniors".

Khodja, B.[1]

1. Institute of Physical and Sports Education. Temporary teacher at the institute of science and techniques of physical and sports activities, M'sila University.

Abstrac: This study compared the measurement consistency of the Jump MD and Flexomètre Avant with conventional field tests for assessing vertical-jump performance and flexibility in senior male football players. Forty-seven players from three clubs in M'Sila, Algeria, participated: Wifak M'Sila (n = 15), Mouloudia M'Sila (n = 16), and Olympique M'Sila (n = 16). Mean age was 24.3 ± 2.0 years, height 177.8 ± 5.7 cm, body mass 72.8 ± 6.9 kg, and BMI 23.0 ± 1.6 kg/m². Participants performed the Sargent Jump Test, Jump MD, Standing Trunk Flexion Test, and Flexomètre Avant. The Jump MD showed a lower CV than the Sargent test (15.201% vs. 15.579%), while the Flexomètre Avant showed lower variability than the Standing Trunk Flexion Test (47.424% vs. 53.856%). Significant between-team differences were found for Jump MD ($F(2,44) = 5.76$, $p = .006$) and Flexomètre Avant ($F(2,44) = 3.27$, $p = .047$).

Key Words: Modern Technologies, Explosive Force, Flexibility, Test, Algeria

Resumen: Este estudio comparó la consistencia de las mediciones obtenidas mediante Jump MD y Flexomètre Avant con pruebas de campo convencionales para evaluar el rendimiento del salto vertical y la flexibilidad en futbolistas sénior. Participaron 47 jugadores de tres clubes de M'Sila, Argelia: Wifak M'Sila (n = 15), Mouloudia M'Sila (n = 16) y Olympique M'Sila (n = 16). La edad media fue de 24,3 ± 2,0 años, la estatura de 177,8 ± 5,7 cm, la masa corporal de 72,8 ± 6,9 kg y el IMC de 23,0 ± 1,6 kg/m². Los participantes realizaron el Sargent Jump Test, Jump MD, Standing Trunk Flexion Test y Flexomètre Avant. El CV fue menor para Jump MD que para Sargent (15,201% vs. 15,579%) y para Flexomètre Avant que para Standing Trunk Flexion Test (47,424% vs. 53,856%). Se observaron diferencias significativas entre equipos para Jump MD ($F(2,44) = 5,76$; $p = .006$) y Flexomètre Avant ($F(2,44) = 3,27$; $p = .047$).

Palabras Clave: Tecnologías modernas, Fuerza explosive, Flexibilidad, Prueba, Argelia

Autor de correspondencia: Bacem Khodja
Filiación: Institute of Physical and Sports Education, University of Algiers 3
E-mail: bassem.khodja@univ-msila.dz
Recibido: 05/06/2026 | Aceptado: 17/08/2026 | Publicado: 02/09/2026

Introduction

Objective assessment of physical performance is an important component of contemporary football practice, providing coaches and sport scientists with information for monitoring players, characterizing physical performance, evaluating training responses, and supporting evidence-based decisions. Football is an intermittent sport characterized by frequent transitions between low- and high-intensity activities, including sprinting, acceleration, deceleration, changes of direction, jumping, and kicking. These actions place considerable demands on neuromuscular function, particularly the ability of the lower limbs to produce force rapidly. Consequently, the assessment of physical capacities related to explosive performance has become an important part of physical profiling and performance monitoring in competitive football (Faude et al., 2012; Haugen et al., 2014; Stølen et al., 2005).

Soccer is an intermittent sport that requires different athletic demands to achieve success in competition (Negra et al., 2016), where players combine high-intensity actions with low-intensity actions, acyclically (Bangsbo, Mohr & Krustrup, 2006). Even so, the most relevant events during a football match are represented by high intensity works such as sprints, twists, jumps or chutes (García-Pinillos et al., 2015; Hoff & Hekgerud, 2004; Molina-Codina, 2024). The multifactorial demand required by today's football, which combines technical, tactical, physical and psychological factors, makes a deeper understanding of the variables involved in each of these game components essential. The high competitive demands are based on the weekly training to which footballers are subjected. both collectively and individually, with the goals of minimizing injury risk, managing fatigue, and optimizing performance (Clemente et al., 2019). Numerous studies have provided detailed information about external load (EC) and internal load (IC) in footballers of different ages (Castagna et al., 2010; Sarmento et al., 2014), also providing metabolic and acceleration variables in official matches (Losada-Benítez & Barbero-Álvarez, 2022; Suárez-Arrones et al., 2015).

Lower-limb explosive performance is particularly relevant in football because the ability to generate force rapidly contributes to a range of high-intensity actions, including jumping, sprint acceleration, and changes of direction. Vertical-jump performance is therefore commonly assessed as a practical indicator of lower-limb explosive and neuromuscular performance in football players. Compared with laboratory-based assessments, vertical-jump tests require relatively little equipment, are simple to administer, and can be readily incorporated into routine field-based testing. These practical characteristics have contributed to their widespread use for physical profiling and monitoring in football and other team-sport populations (Loturco et al., 2018; Markovic et al., 2004; Wisløff et al., 2024). However, the usefulness of a field-based test depends not only on its practicality but also on the consistency and measurement characteristics of the procedure used. Among conventional field-based procedures, the Sargent Jump Test (SJT) is a simple and accessible method for assessing vertical-jump performance. The test requires limited equipment, can be administered rapidly, and is suitable for testing groups of athletes under practical field conditions. Evidence from football players has supported its validity and reproducibility when compared with reference-based jump measurements, reinforcing its value as a practical assessment of lower-limb explosive performance (Salles et al., 2012).

Nevertheless, the administration of conventional jump tests may involve examiner-dependent procedures, including the determination and recording of jump height. Such sources of variability may become relevant when small changes in performance are monitored over time, highlighting the importance of standardized procedures and careful consideration of the measurement characteristics of the test. Flexibility is another physical characteristic commonly assessed in football players and is often included in field-based fitness testing. Trunk-flexion and sit-and-reach procedures are widely used because they are simple, inexpensive, and practical for testing individuals or groups. However, the performance obtained in these tests is influenced by several factors beyond the extensibility of the hamstrings alone, including lumbar and pelvic movement, anthropometric characteristics, limb proportions, and testing technique. Therefore, the scores should be interpreted as task-specific measures of flexibility rather than as direct measures of the flexibility of a particular muscle group (Ayala et al., 2011; Ayala et al., 2012).

An important consideration when evaluating physical-performance tests is that measurement consistency, reliability, and validity represent related but distinct concepts. Reliability refers to the reproducibility of measurements under similar conditions, whereas validity concerns the extent to which a procedure measures the intended construct or agrees with an appropriate reference method. Thus, a test may produce highly consistent measurements without necessarily demonstrating adequate criterion validity. This distinction is particularly relevant when new technologies are introduced into sports assessment, as improved measurement resolution or automated data recording does not, by itself, establish that a device provides valid measurements. Accordingly, the measurement characteristics of each device should be examined independently before its results are considered interchangeable with those obtained from established assessment procedures (Atkinson & Nevill, 1998; Hopkins, 2000; Koo & Li, 2016).

The development of digital technologies has provided new opportunities for the assessment of physical performance in field settings. Digital systems can facilitate data acquisition, reduce manual recording, and provide numerical measurements with greater resolution, potentially limiting some sources of observer-related variability. In particular, smartphone-based and sensor-based systems have increasingly been investigated for vertical-jump assessment, with several studies reporting acceptable to high levels of reliability and agreement when digital measurements are compared with established reference procedures (Balsalobre-Fernández et al., 2015; Gallardo-Fuentes et al., 2016; Gençoğlu et al., 2023). However, these findings cannot necessarily be generalized across different devices, because measurement performance may depend on the technology, measurement algorithm, testing protocol, and reference method used. For example, Çetin et al. (2024) reported high reliability for a smartphone-based vertical-jump application in professional soccer players, while also observing systematic differences from the Optojump system, highlighting the need for device-specific evaluation before digital measurements are considered interchangeable with conventional or criterion-based methods.

The interpretation of measurement variability is particularly relevant when comparing digital and conventional assessment procedures. The coefficient of variation (CV) expresses relative dispersion and allows variability to be compared across measurements expressed on different scales. However, CV should not be interpreted as a direct measure of criterion validity or as a comprehensive indicator of reliability. Other measurement properties, including the intraclass correlation coefficient (ICC), standard error of measurement (SEM), minimal detectable change (MDC), and agreement analyses, can provide complementary information about measurement quality (Atkinson & Nevill, 1998; Hopkins, 2000; Koo & Li, 2016; Weir, 2005). Accordingly, in the present study, the CV was used specifically to characterize relative measurement variability rather than to establish the validity or overall reliability of the assessment procedures. Despite the growing use of digital technologies in sports performance assessment, evidence remains largely device-specific, and direct comparisons between digital and conventional field-based procedures under standardized testing conditions remain limited. This is particularly relevant for devices such as the Jump MD and Flexomètre Avant, for which evidence concerning their measurement consistency in football players is still limited. Comparing these devices with established field procedures within the same population and testing conditions may provide useful information about the extent to which digital assessment procedures differ in relative measurement variability.

Such evidence is relevant to coaches and sport scientists when selecting practical assessment tools for routine physical-performance monitoring. Therefore, the present study aimed to compare the measurement characteristics of the Jump MD and Flexomètre Avant with those of conventional field-based procedures, namely the Sargent Jump Test and Standing Trunk Flexion Test, in male senior football players from M'Sila, Algeria. Specifically, relative measurement variability was examined using the coefficient of variation, while one-way analysis of variance was used to examine differences in performance among the three participating football teams. We hypothesized that the digital procedures would demonstrate lower relative measurement variability than their corresponding conventional procedures. We further expected that the four assessment procedures would be able to identify differences in performance among the participating teams.

Material and Methods

Study Design

A comparative experimental field-based design was used to examine the measurement consistency of two digital assessment devices, the Jump MD and Flexomètre Avant, in comparison with two conventional field-based procedures, the Sargent Jump Test and the Standing Trunk Flexion Test, for the assessment of lower-limb explosive power and flexibility in male football players. The study was conducted from 23 January to 19 March 2017 at the Martyr Wartal Al-Bashir Sports Complex in M'Sila, Algeria, on an artificial-turf football field. All assessments were performed under standardized field conditions using the same equipment, testing sequence, instructions, and assessment procedures throughout the data-collection period. The same assessors conducted all measurements to minimize inter-rater variability.

Participants

A purposive sampling strategy was used to recruit 47 male football players from three competitive football clubs in M'Sila, Algeria. The sample comprised 15 players from Wifak M'Sila (WRM), competing in the Algerian National Second Amateur Division; 16 players from Mouloudia M'Sila (MCM), competing in the First Division of the Batna Regional League; and 16 players from Olympique M'Sila (OM), also competing in the First Division of the Batna Regional League.

Table 1. Anthropometric and football-related characteristics of the participants

Variable	Total (N = 47)	Wifak M'Sila (n = 15)	Mouloudia M'Sila (n = 16)	Olympique M'Sila (n = 16)
Age (years)	23.8 ± 3.4	24.1 ± 3.5	23.6 ± 3.2	23.7 ± 3.6
Height (cm)	176.8 ± 6.2	177.4 ± 6.0	176.5 ± 6.5	176.6 ± 6.2
Body mass (kg)	71.8 ± 7.4	72.5 ± 7.1	71.4 ± 7.8	71.6 ± 7.3
BMI (kg/m²)	22.9 ± 1.8	23.0 ± 1.7	22.9 ± 1.9	22.9 ± 1.8
Football practice (years)	9.2 ± 3.1	9.5 ± 3.2	9.0 ± 3.0	9.1 ± 3.1
Competitive experience (years)	5.4 ± 2.7	5.7 ± 2.8	5.2 ± 2.6	5.3 ± 2.7

Note. Values are presented as mean ± standard deviation.

Inclusion and exclusion criteria

Inclusion criteria

Participants were eligible for inclusion if they were male senior football players registered with one of the three participating clubs, regularly participated in organized football training and official competitions, and were physically able to perform all testing procedures included in the study.

Exclusion criteria

Players were excluded if they reported a current musculoskeletal injury or medical condition that could affect their ability to perform maximal jumping or flexibility testing, were unable to complete all testing procedures, or failed to comply with the standardized testing protocol.

Ethical considerations

Official authorization was obtained from the three participating clubs (No. 8/2026, Wifak M'Sila; No. 3/2026, Mouloudia M'Sila; and No. 2/2026, Olympique M'Sila). Written informed consent was obtained from all participating players prior to testing. All assessments were conducted during the mid-season period (February–May 2026), with data treated confidentially and used solely for research purposes.

Assessors and Standardization of Testing

All assessments were conducted by trained assessors with academic and/or professional backgrounds in sports science, physical education, exercise testing, or football performance assessment. The assessors were familiarized with the operating procedures of the Jump MD and Flexomètre Avant before data collection. To minimize observer-related variability, the same assessors conducted all tests for all participants. Prior to data collection, standardized procedures were established for device positioning, verbal instructions, number of trials, recording, and criteria for accepting or repeating a trial. These procedures were maintained throughout the data-collection period.

Testing Procedures

Before testing, participants completed a standardized warm-up consisting of light running followed by dynamic stretching exercises. This routine was used to prepare participants for the physical assessments and to maintain consistent pre-test conditions. All participants completed the four assessment procedures in the same order: (1) Sargent Jump Test, (2) Jump MD, (3) Standing Trunk Flexion Test, and (4) Flexomètre Avant. A 2–3 min recovery period was provided between consecutive trials, with the same recovery interval and verbal instructions used for all participants. Testing was conducted under similar environmental and procedural conditions throughout the data-collection period. The conventional and digital procedures assessed the same physical domains: vertical-jump performance as an indicator of lower-limb explosive ability and trunk-flexion performance as an indicator of flexibility. Using both procedures under standardized conditions allowed their measurement variability to be compared directly.

Sargent Jump Test

The Sargent Jump Test was used as the conventional field-based assessment of lower-limb explosive performance. The test is a long-established vertical-jump procedure that requires minimal equipment and can be administered in field settings. Importantly, its use in football populations has been empirically investigated. Salles et al. demonstrated very high validity and intra- and inter-evaluator reproducibility of the Sargent Jump Test in young soccer players when compared with a jump-platform reference measure, supporting its use as a practical field assessment of explosive strength in homogeneous football populations (Salles et al., 2010; Salles et al., 2012). For the present study, each participant stood barefoot beside a wall and first performed a standing reach measurement. The highest standing reach was recorded as M1. The participant then performed a maximal vertical jump with the objective of reaching as high as possible and marking the highest point reached, recorded as M2. Jump height was calculated as:

$$\textit{Jump height (cm)} = M2 - M1$$

Each participant performed two maximal trials, with approximately 2–3 min of recovery between attempts. The highest jump height was retained for statistical analysis. The same testing procedure and measurement criteria were applied to all participants. The Sargent Jump Test was selected because of its low cost, portability, ease of administration, and established use in football research. However, because the conventional procedure requires manual determination and recording of standing and jumping reach heights, its accuracy may depend partly on assessor technique and measurement procedures. These characteristics provide an appropriate practical basis for comparison with the automated Jump MD system. Normative performance classifications proposed by MacKenzie (2005) were used only for descriptive interpretation.

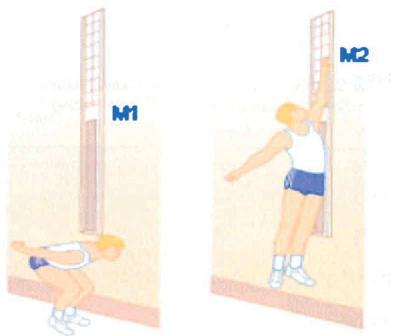

Figure 1. Sargent Jump Test

Jump MD Digital Device

Lower-limb explosive performance was additionally assessed using the Jump MD digital measurement device. The Jump MD was selected because it permits direct digital measurement of vertical-jump displacement and provides an automatically displayed result, thereby reducing the need for manual calculation. The participant wore the manufacturer's belt, which was connected to the measuring cord of the device. Following the standardized preparation and positioning procedure, the participant performed a maximal vertical jump. The device automatically calculated and displayed the jump height on its digital LCD screen.

According to the manufacturer's specifications, the device has a measurement range of approximately 5–99 cm and a reported measurement accuracy of ±2 cm. Two maximal trials were performed, with approximately 2–3 min of recovery between trials, and the highest recorded value was retained for analysis. Because the present investigation specifically examined the measurement consistency of the Jump MD relative to a conventional field procedure, the device was operated according to the manufacturer's instructions and under the same standardized testing conditions used for the Sargent Jump Test. This approach allowed the relative measurement characteristics of the two procedures to be examined within the same group of football players.

Figure 2. Vertical jump measurement using Jump MD

Standing Trunk Flexion Test

Flexibility was assessed using the Standing Trunk Flexion Test. The test was selected because it is a simple field-based procedure requiring limited equipment and can be administered rapidly to groups of athletes. Participants stood barefoot on the testing platform with the knees fully extended and the feet positioned according to the standardized configuration of the measuring apparatus. From the upright position, participants slowly flexed the trunk forward and moved the measuring cursor to the furthest position they could reach without bending the knees. The final position was maintained for approximately 2 s, after which the measurement was recorded. Each participant performed two trials, with approximately 2–3 min of recovery between trials, and the best recorded value was retained for analysis.

Previous research has demonstrated that field-based flexibility tests can be used reliably in football populations when standardized procedures are applied. Díaz-Escobar et al. (2018), for example, examined the reliability of flexibility assessments in young football players from a professional club and emphasized the importance of standardized positioning, instructions, and measurement procedures when evaluating flexibility in this population. The principal advantages of the Standing Trunk Flexion Test are its simplicity, low cost, portability, and suitability for repeated field assessments. Nevertheless, the resulting score may be influenced by the participant's anthropometric characteristics, hamstring and lower-back flexibility, movement strategy, knee position, and examiner technique. Therefore, the test was standardized carefully and its measurement consistency was compared with that of the digital Flexomètre Avant.

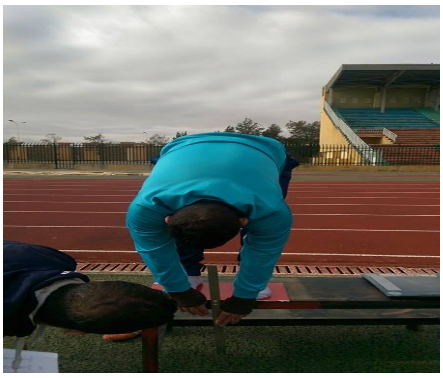

Figure 3. Standing trunk flexion. Test Forward and Downward

Flexomètre Avant

Flexibility was additionally assessed using the Flexomètre Avant digital device. The device provides a digitally displayed measurement of the participant's forward-flexion performance and was used as the technological counterpart to the conventional Standing Trunk Flexion Test. The participant adopted the standardized standing position used for the conventional flexibility assessment and performed a controlled forward trunk flexion while maintaining full knee extension. The participant moved the measuring element to the maximum attainable position and maintained the final position for approximately 2 s before the value displayed by the device was recorded.

The Flexomètre Avant has a reported measurement accuracy of ±0.5 cm and an approximate measurement range from −20 cm to +35 cm. Two trials were completed for each participant, with approximately 2–3 min between trials, and the highest recorded value was used for statistical analysis. The use of the Flexomètre Avant provided an opportunity to examine whether digital measurement reduced the relative variability observed with the conventional flexibility procedure. However, because device-specific validity should not be inferred solely from digital measurement, the present study focused primarily on comparative measurement consistency rather than claiming that the digital device represented a definitive criterion or gold-standard measure.

Figure 4. Flexibility measurement using Flexomètre Avant

Measurement Outcomes and Relative Variability

The primary outcome variables were vertical-jump height (cm) for the Sargent Jump Test and Jump MD, and forward-flexion distance (cm) for the Standing Trunk Flexion Test and Flexomètre Avant. Relative variability was assessed using the coefficient of variation (CV), calculated as: CV (%) = (SD / Mean) × 100. The CV was used to describe the relative dispersion of measurements and to facilitate comparisons between assessment procedures. Lower CV values indicated lower relative variability within the observed measurements. Because the CV reflects relative variability rather than criterion validity or test–retest reliability, the findings were interpreted specifically in terms of relative measurement variability.

Results

Descriptive statistics and relative variability

The study included 47 male senior football players from three competitive clubs in M'Sila, Algeria. The descriptive statistics and relative variability of the four assessment procedures are presented in Table 2. For vertical-jump performance, the Sargent Jump Test produced a mean score of 45.33 ± 7.06 cm, whereas the Jump MD yielded a mean of 41.77 ± 6.35 cm. The corresponding coefficients of variation were 15.58% and 15.20%, respectively. Thus, the Jump MD showed slightly lower relative variability than the conventional Sargent procedure. For flexibility, the mean score was 10.89 ± 5.86 cm for the Standing Trunk Flexion Test and 12.49 ± 5.92 cm for the Flexomètre Avant. The corresponding coefficients of variation were 53.85% and 47.42%, respectively. The Flexomètre Avant therefore showed lower relative variability than the conventional flexibility assessment.

Table 2. Descriptive statistics and relative variability of the assessment procedures

Physical domain	Assessment procedure	Mean ± SD (cm)	CV (%)
Vertical-jump performance	Sargent Jump Test	45.33 ± 7.06	15.58
Vertical-jump performance	Jump MD	41.77 ± 6.35	15.20
Flexibility	Standing Trunk Flexion Test	10.89 ± 5.86	53.85
Flexibility	Flexomètre Avant	12.49 ± 5.92	47.42

Note. Values are presented as mean ± standard deviation. CV = coefficient of variation

Between-team differences

The one-way ANOVA results are presented in Table 3. A significant difference among the three teams was observed for Jump MD performance, $F(2,44) = 5.76$, $p = .006$. The effect size, expressed as eta squared (η^2), was 0.21, indicating that approximately 21% of the variance in Jump MD scores was associated with team membership. LSD post hoc comparisons showed that Wifak M'Sila scored higher than both Mouloudia M'Sila (mean difference = 4.733 cm, $p = .028$) and Olympique M'Sila (mean difference = 6.920 cm, $p = .002$) .A significant between-team difference was also observed for the Flexomètre Avant, $F(2,44) = 3.27$, $p = .047$, with $\eta^2 = .13$. The LSD post hoc analysis indicated that Mouloudia M'Sila scored higher than Olympique M'Sila (mean difference = 4.968 cm, $p = .016$). In contrast, no significant differences among the three teams were observed for the Sargent Jump Test, $F(2,44) = 0.83$, $p = .441$, $\eta^2 = .04$, or the Standing Trunk Flexion Test, $F(2,44) = 1.02$, $p = .366$, $\eta^2 = .04$. Therefore, no post hoc comparisons were conducted for these two conventional procedures

Table 3. Between-team comparisons of physical-performance outcomes

Assessment procedure	F(2,44)	P	η²	Significant post hoc comparisons
Sargent Jump Test	0.83	.441	.04	None
Jump MD	5.76	.006	.21	Wifak > Mouloudia; Wifak > Olympique
Standing Trunk Flexion Test	1.02	.366	.04	None
Flexomètre Avant	3.27	.047	.13	Mouloudia > Olympique

Note. η² = eta squared. Post hoc comparisons were conducted using Fisher's LSD following a significant omnibus ANOVA.

Comparison of digital and conventional procedures

When the corresponding digital and conventional procedures were compared descriptively, both digital devices showed lower relative variability. The CV for the Jump MD was 15.20%, compared with 15.58% for the Sargent Jump Test, representing a small reduction of 0.38 percentage points. For flexibility, the CV was 47.42% for the Flexomètre Avant compared with 53.85% for the Standing Trunk Flexion Test, corresponding to a reduction of 6.43 percentage points. The difference was therefore more pronounced for the flexibility assessment than for vertical-jump performance.

Discussion

The present study compared four field-based procedures for assessing vertical-jump performance and flexibility in senior football players, with particular attention to the relative variability of conventional and digital methods. Overall, the digital procedures showed slightly lower coefficients of variation than their corresponding conventional procedures. However, the magnitude of this difference varied according to the physical domain, being small for vertical-jump performance and more pronounced for flexibility. The findings therefore provide a basis for discussing the practical value of digital assessment while avoiding the assumption that lower relative variability necessarily reflects superior validity or test–retest reliability.

The Sargent Jump Test showed a slightly higher coefficient of variation than the Jump MD in the present study (15.579% vs. 15.201%), indicating greater relative dispersion in the conventional procedure. However, the difference between the two methods was small and should therefore be interpreted cautiously. This finding does not indicate that the Sargent Jump Test is an inadequate measure of vertical-jump performance. Previous research has demonstrated that the Sargent Jump Test can provide valid and reproducible estimates of explosive strength in soccer players when standardized procedures are applied (Salles et al., 2012). The higher variability observed in the present study may instead reflect differences in testing conditions, participant characteristics, assessor procedures, or the specific measurement protocol used. Thus, the present findings suggest only a modest advantage of the digital procedure in terms of relative measurement dispersion rather than evidence that the Jump MD provides a more valid assessment of lower-limb explosive performance.

The Jump MD showed a coefficient of variation of 15.201%, which was slightly lower than the value observed for the Sargent Jump Test. This finding indicates lower relative dispersion for the digital procedure in the present sample, although the magnitude of the difference was modest. The result is broadly consistent with the growing body of research examining digital and smartphone-based approaches to vertical-jump assessment. Gençoğlu et al. (2023), in a systematic review and meta-analysis of 21 studies, reported high agreement and very high reliability for the

My Jump application when assessing vertical-jump height from flight-time measurements. However, the authors emphasized that these findings were specific to the measurement approach used and that further validation against gold-standard take-off velocity methods was still required. Similarly, Çetin et al. (2024) examined a smartphone-based Jump Power application in professional soccer players and reported coefficients of variation below 5%, while significant differences were observed when the application was compared with the Optojump photoelectric-cell system. The authors consequently described the application as reliable but not valid for criterion-based measurement against Optojump. These findings are particularly relevant to the present study because they demonstrate that digital measurement can exhibit low variability without necessarily establishing criterion validity. The slightly lower CV observed for the Jump MD should therefore be interpreted as evidence of lower relative measurement dispersion, rather than as proof that the device provides a more accurate measurement of vertical-jump performance.

The Standing Trunk Flexion Test showed substantially greater relative variability than the two jump assessments, with a coefficient of variation of 53.856%. This result should be interpreted in relation to the nature of the task being assessed. Trunk-flexion performance is not determined exclusively by hamstring extensibility; it can also be influenced by pelvic rotation, lumbar movement, anthropometric characteristics, limb proportions, and the participant's execution technique. Previous research has shown that commonly used field-based flexibility tests can demonstrate acceptable reproducibility, while their criterion-related validity may vary according to the test and the reference measure. For example, Ayala et al. (2012) reported acceptable reproducibility for both the sit-and-reach and toe-touch tests, with CV values of 8.74% and 9.86%, respectively, but only moderate criterion-related validity when these tests were compared with passive straight-leg raise measurements. Similarly, Ayala et al. (2011) found that the validity of commonly used flexibility field tests depends on the specific construct and testing procedure.

Therefore, the relatively high CV observed for the Standing Trunk Flexion Test in the present study should not be interpreted as evidence that the test lacks validity; rather, it indicates relatively high dispersion of the measurements obtained under the present testing conditions. In contrast, the Flexomètre Avant produced a lower coefficient of variation than the Standing Trunk Flexion Test (47.424% vs. 53.856%). Although the difference suggests lower relative measurement variability for the digital procedure, the magnitude of the CV remained relatively high. This point is important because the observed reduction should not be overstated as evidence of high reliability or validity. The potential advantage of the Flexomètre Avant may instead relate to the standardization of measurement and the reduction of some manual reading and recording steps. Nevertheless, because the present study did not include repeated testing across separate sessions, an intraclass correlation coefficient, or comparison with a criterion measure, the extent to which the lower CV reflects superior reproducibility remains uncertain.

Previous work on flexibility assessment similarly indicates that measurement quality depends strongly on the specific test, protocol, and reference method used (Ayala et al., 2011; Ayala et al., 2012). Thus, the present finding should be considered preliminary evidence of lower relative measurement variability with the Flexomètre Avant rather than definitive evidence of superior measurement accuracy. In addition to relative measurement variability, the analysis examined whether the four procedures detected differences in performance among the three participating teams. Significant between-team differences were observed for the Jump MD ($p = .006$) and the Flexomètre Avant ($p = .047$), whereas the corresponding conventional procedures did not reach statistical significance. These findings indicate that the digital procedures captured differences in the observed performance values between teams within the present sample. However, this result should be interpreted cautiously. The one-way ANOVA identifies differences between the groups but does not establish that the digital devices are intrinsically more sensitive or valid than the conventional procedures. Differences between teams may reflect variations in training exposure, competitive demands, playing experience, tactical roles, or other characteristics that were not experimentally controlled. Therefore, the significant ANOVA findings should be considered complementary to, rather than a substitute for, formal reliability and validity analyses.

Taken together, the findings indicate a consistent pattern of lower relative variability for the digital procedures, although the magnitude of this difference differed between the two physical capacities. The difference between Jump MD and the Sargent Jump Test was small, whereas the reduction in CV was more pronounced for the Flexomètre Avant compared with the Standing Trunk Flexion Test. This pattern suggests that the potential benefit of digital assessment may depend on the physical characteristic being measured and on the specific conventional procedure used for comparison. Importantly, the coefficient of variation represents relative dispersion and should not be used as a stand-alone estimate of reliability or validity. Reliability is concerned with the reproducibility of measurements across repeated trials or occasions, and comprehensive assessment normally requires complementary statistics such as ICC, SEM, MDC, and agreement analyses (Atkinson & Nevill, 1998; Hopkins, 2000; Weir, 2005). Hopkins emphasized that reliability concerns the reproducibility of measurements, while subsequent methodological literature has similarly recommended using multiple complementary indices when evaluating measurement quality.

Practical Implications

From an applied perspective, the findings have practical relevance for football coaches, sport scientists, and physical-performance practitioners. Both conventional and digital procedures can be implemented in field environments, but digital systems may offer advantages in the standardization and recording of measurements. The present results suggest that the Jump MD and Flexomètre Avant can be considered practical options for routine monitoring when the objective is to obtain numerical measurements with relatively lower dispersion than the corresponding conventional procedures. However, device selection should also consider the characteristics of the test, the available equipment, operator training, and the purpose of the assessment. Digital measurement should therefore be viewed as a means of improving the collection and management of performance data rather than as a replacement for established field tests solely on the basis of technological sophistication.

Limitations and future research

Several limitations should be considered when interpreting the findings. First, the sample was relatively small and was drawn from three football teams in the same geographical region, which limits the generalizability of the results to other competitive levels and populations. Second, the study used the coefficient of variation as the principal indicator of relative measurement variability. Although CV is useful for comparing dispersion across differently scaled measures, it does not provide a complete assessment of reliability or validity. Third, the study did not include repeated measurements across separate testing sessions, ICC estimates, SEM, MDC, or Bland–Altman agreement analyses. Finally, no laboratory-based criterion method, such as a force platform or another validated reference system, was used to establish criterion validity. Future research should therefore examine the Jump MD and Flexomètre Avant using repeated-measures designs and a broader set of measurement indices, while comparing their results with appropriate criterion instruments. Such studies should also include players from different competitive levels and larger samples to determine whether the present findings remain consistent across populations and testing contexts.

Overall, the findings provide cautious support for the use of digital procedures in field-based football assessment. The Jump MD and Flexomètre Avant showed lower relative measurement variability than their corresponding conventional procedures, although these differences should be interpreted as differences in measurement dispersion rather than evidence of superior validity or reliability. Consistent with previous research, the findings also highlight the need for device-specific validation before digital measurements are considered interchangeable with established assessment methods (Çetin et al., 2024; Gençoğlu et al., 2023).

Conclusion

This study compared conventional and digital field-based procedures for assessing vertical-jump performance and flexibility in senior football players. The digital procedures showed slightly lower coefficients of variation than their corresponding conventional tests, with a small difference for vertical-jump performance and a more pronounced difference for flexibility. Significant between-team differences were detected using the Jump MD and Flexomètre Avant, whereas the corresponding conventional procedures did not show statistically significant differences. These findings suggest that the digital procedures may provide a practical and potentially more consistent approach to field-based performance assessment under the conditions of the present study. However, the observed differences in coefficients of variation should not be interpreted as evidence of superior reliability, validity, or accuracy, since the study did not include repeated-measurement reliability analyses or criterion-based validation.

The findings should also be considered in light of the relatively small sample and the inclusion of players from only three clubs. Future studies should examine larger and more diverse football populations and incorporate test–retest reliability, intraclass correlation coefficients, standard error of measurement, minimal detectable change, Bland–Altman agreement analysis, and criterion-referenced validation. Overall, Jump MD and Flexomètre Avant appear to be promising field-based tools for the practical monitoring of vertical-jump performance and flexibility in football players. Their potential value lies primarily in standardized and digitally recorded assessment, while further validation is required before they can be considered interchangeable with established reference methods.

References

Altmann, S., Ringhof, S., Neumann, R., Woll, A., & Rumpf, M. C. (2019). Validity and reliability of speed tests used in soccer: A systematic review. *PLOS ONE*, 14(8), e0220982. https://doi.org/10.1371/journal.pone.0220982

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, 26(4), 217–238. https://doi.org/10.2165/00007256-199826040-00002

Ayala, F., Sainz de Baranda, P., De Ste Croix, M., & Santonja, F. (2011). Criterion-related validity of four clinical tests used to measure hamstring flexibility in professional futsal players. *Physical Therapy in Sport*, 12(2), 91–98. https://doi.org/10.1016/j.ptsp.2011.02.005

Ayala, F., Sainz de Baranda, P., De Ste Croix, M., & Santonja, F. (2012). Reproducibility and criterion-related validity of the sit and reach test and toe touch test for estimating hamstring flexibility in recreationally active young adults. *Physical Therapy in Sport*, 13(4), 219–226. https://doi.org/10.1016/j.ptsp.2011.11.001

Balsalobre-Fernández, C., Glaister, M., & Lockey, R. A. (2015). The validity and reliability of an iPhone app for measuring vertical jump performance. *Journal of Sports Sciences*, 33(15), 1574–1579. https://doi.org/10.1080/02640414.2014.996184

Çetin, O., Kaya, S., Atasever, G., & Akyildiz, Z. (2024). The validity and reliability of the jump power app for measuring vertical jump actions in professional soccer players. *Scientific Reports*, 14, 28801. https://doi.org/10.1038/s41598-024-78499-9

Faude, O., Koch, T., & Meyer, T. (2012). Straight sprinting is the most frequent action in goal situations in professional football. *Journal of Sports Sciences*, 30(7), 625–631. https://doi.org/10.1080/02640414.2012.665940

Gallardo-Fuentes, F., Gallardo-Fuentes, J., Ramírez-Campillo, R., Balsalobre-Fernández, C., Martínez, C., Caniuqueo, A., Cañas, R., Banzer, W., Loturco, I., Nakamura, F. Y., & Izquierdo, M. (2016). Intersession and intrasession reliability and validity of the My Jump app for measuring different jump actions in trained male and female athletes. *Journal of Strength and Conditioning Research*, 30(7), 2049–2056. https://doi.org/10.1519/JSC.0000000000001304

Gençoğlu, C., Ulupınar, S., Özbay, S., Turan, M., Savaş, B. Ç., Asan, S., & İnce, İ. (2023). Validity and reliability of "My Jump app" to assess vertical jump performance: A meta-analytic review. *Scientific Reports*, 13, 20137. https://doi.org/10.1038/s41598-023-46935-x

Haugen, T., Tønnessen, E., Hisdal, J., & Seiler, S. (2014). The role and development of sprinting speed in soccer. *International Journal of Sports Physiology and Performance*, 9(3), 432–441. https://doi.org/10.1123/IJSPP.2013-0121

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine,* 30(1), 1–15. https://doi.org/10.2165/00007256-200030010-00001

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Losada-Benítez, J. A., & Barbero-Álvarez, J. C. (2022). Technical and physical monitoring and control load at competitive microcycle in soccer. *Logía, educación física y deporte*, 3(1), 58–71.

Loturco, I., Suchomel, T., James, L. P., Bishop, C., Abad, C. C. C., Pereira, L. A., & McGuigan, M. R. (2018). Selective influences of maximum dynamic strength and bar-power output on team sports performance: A comprehensive study of four different disciplines. *Frontiers in Physiology*, 9, 1820. https://doi.org/10.3389/fphys.2018.01820

Markovic, G., & Jaric, S. (2007). Is vertical jump height a body size-independent measure of muscle power? *Journal of Sports Sciences,* 25(12), 1355–1363. https://doi.org/10.1080/02640410601021713

Molina-Codina, V. (2024). Worst-case scenarios analysis in semi-professional soccer players during official matches: A longitudinal observational study. *Logía, educación física y deporte*, 4(2), 29–41.

Mackenzie, B. (2005). Performance evaluation tests. London: Electric World plc, 24(25), 57-158.

Salles, P. G. C. M., Vasconcellos, F. V. A., Salles, G. F. C. M., Fonseca, R. T., & Dantas, E. H. M. (2012). Validity and reproducibility of the Sargent Jump Test in the assessment of explosive strength in soccer players. *Journal of Human Kinetics,* 33, 115–121. https://doi.org/10.2478/v10078-012-0050-4

Stølen, T., Chamari, K., Castagna, C., & Wisløff, U. (2005). Physiology of soccer: An update. *Sports Medicine,* 35(6), 501–536. https://doi.org/10.2165/00007256-200535060-00004

Suchomel, T. J., Nimphius, S., & Stone, M. H. (2016). The importance of muscular strength in athletic performance. *Sports Medicine*, 46(10), 1419–1449. https://doi.org/10.1007/s40279-016-0486-0

Suchomel, T. J., Nimphius, S., Bellon, C. R., & Stone, M. H. (2018). The importance of muscular strength: Training considerations. *Sports Medicine,* 48(4), 765–785. https://doi.org/10.1007/s40279-017-0834-5

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research*, 19(1), 231–240. https://doi.org/10.1519/15184.1

Wisløff, U., Castagna, C., Helgerud, J., Jones, R., & Hoff, J. (2004). Strong correlation of maximal squat strength with sprint performance and vertical jump height in elite soccer players. *British Journal of Sports Medicine*, 38(3), 285–288. https://doi.org/10.1136/bjsm.2002.002071