SCIENTIFIC JOURNAL OF SPORT AND PERFORMANCE

REVIEW ARTICLE

Convergent validity of session-RPE against heart-rate and GPS training load measures in football: A systematic review and meta-analysis

Rafik Guia . *Institute of Physical Education and Sports Sciences. Kasdi Merbah University of Ouargla. Ouargla, Algeria.*
Mohamed Salah Benharkat. *Department of Physical Education and Sports Sciences. Badji Mokhtar University of Annaba. Algeria.*
Nadjib Ababsa. *Institute of Physical Education and Sports Sciences. Kasdi Merbah University of Ouargla. Ouargla, Algeria.*
Ramdan Zaki. *Institute of Physical Education and Sports. University of Algiers 3. Algeria.*

ABSTRACT

Session rating of perceived exertion (sRPE) is the most widely used method for quantifying internal training load in football, yet its convergent validity with objective heart-rate (HR) and GPS-derived external load measures remains insufficiently characterized. We conducted a systematic review and meta-analysis to quantify the convergent validity of sRPE against HR-based and GPS-derived training load measures in football, and to test whether validity coefficients differ systematically between GPS volume and GPS intensity metrics. Five databases (Scopus, PubMed/MEDLINE, Web of Science, SPORTDiscus, and Google Scholar) were searched for studies published between 2013 and 2025. Ten studies contributing 40 correlation coefficients were included. A three-level random-effects model with REML estimation was applied. Pooled estimates were: sRPE vs HR load $r = .805$ [95% CI: .729, .861; $I^2 = 0.0\%$]; sRPE vs GPS volume $r = .760$ [95% CI: .654, .836; $I^2 = 45.7\%$]; sRPE vs GPS intensity $r = .556$ [95% CI: .465, .635; $I^2 = 0.0\%$]; overall $r = .708$ [95% CI: .616, .781; $I^2 = 42.6\%$]. The volume–intensity dissociation was statistically significant ($\Delta r = 0.204$, $Z = 2.94$, $p = .003$). sRPE demonstrates acceptable convergent validity with HR-based and GPS-volume measures but is markedly less sensitive to high-intensity external load. These findings have direct implications for the use of sRPE-derived ACWR in training load monitoring and injury risk assessment in football.

Keywords: session RPE; training load; GPS; heart rate; football; meta-analysis; convergent validity; ACWR.

Cite this article as:
Guia, R., Benharkat, M. S., Ababsa, N., & Zaki, R. (2026). Convergent validity of session-RPE against heart-rate and GPS training load measures in football: A systematic review and meta-analysis. *Scientific Journal of Sport and Performance*, *5*(4), 646-657. https://doi.org/10.55860/QNIQ5270

Corresponding author. *Institute of Physical Education and Sports Sciences. Kasdi Merbah University of Ouargla. Ouargla, Algeria.*
E-mail: rafik.guia@univ-ouargla.dz
Submitted for publication June 18, 2026.
Accepted for publication August 05, 2026.
Published August 22, 2026.
Scientific Journal of Sport and Performance. ISSN 2794-0586.

doi: https://doi.org/10.55860/QNIQ5270

INTRODUCTION

Accurate quantification of training load is fundamental to optimizing performance and managing injury risk in competitive football. Internal training load—reflecting the physiological stress imposed on a player—is most commonly assessed using session rating of perceived exertion (sRPE), whereby an athlete rates overall exertion on the CR-10 or CR-100 scale after each session, and this value is multiplied by session duration in minutes to yield a session load in arbitrary units (Foster et al., 2001). The method is widely adopted in applied settings because it is inexpensive, non-invasive, and sensitive to individual differences in perceived effort.

External training load is increasingly captured using GPS-based wearable technology, which provides objective, continuous measures of mechanical work output including total distance covered, high-speed running (HSR) distance, sprint distance, and accelerations. Heart-rate-based measures such as the training impulse (TRIMP) and Edwards' method offer an established index of cardiovascular strain. Whether sRPE captures the same physiological construct as these objective methods is a question of convergent validity—the degree to which two measures purporting to assess the same construct agree.

A prior meta-analysis by McLaren et al. (2018) pooled correlations between sRPE and various training load indicators across multiple team sports, reporting an overall pooled $r = .59$. However, that analysis combined team sports heterogeneously and did not differentiate between GPS volume metrics (e.g., total distance, Player Load) and GPS intensity metrics (e.g., HSR distance, sprint distance, acceleration count). This distinction is theoretically critical: sRPE is expected to track overall session effort—and thus volume—more faithfully than discrete bouts of high-intensity activity, which may not register proportionally in an end-of-session perceptual rating.

No meta-analysis has formally tested the volume–intensity dissociation hypothesis within football populations using contemporary GPS monitoring data. The present study addressed this gap with three aims: (1) to quantify the pooled convergent validity of sRPE against HR-based training load measures in football; (2) to quantify the pooled convergent validity of sRPE against GPS-derived volume and intensity measures separately; and (3) to test whether validity coefficients differ significantly between GPS volume and GPS intensity subgroups.

MATERIAL AND METHODS

Study design

This study followed a pre-registered systematic review and meta-analysis design, reported in accordance with the PRISMA 2020 statement (Page et al., 2021). The review protocol was pre-registered on the Open Science Framework at https://osf.io/my3ve prior to data extraction. No new data involving human participants were collected; ethics approval was therefore not required.

Participants

Eligible studies enrolled football players (male or female, any competitive level, aged ≥ 14 years) competing in 11-a-side association football. Studies examining futsal, beach soccer, or other football variants were excluded.

Measures

The index measure was sRPE or sRPE-TL (sRPE × session duration in minutes), assessed using the CR-10 or CR-100 Borg scale (Foster et al., 2001). Comparator measures were classified into three mutually exclusive categories: (a) HR-based load — TRIMP variants (Banister, Lucia, or Edwards), HR zone-based session load; (b) GPS volume — total distance (TD), Player Load (PL), or equivalent composite distance metrics; (c) GPS intensity — high-speed running distance (HSR, threshold ≥ 14.4 km/h or study-specific equivalent), sprint distance (>20 km/h), acceleration/deceleration count or load. Studies reporting dose-response associations, regression coefficients without explicit r, or sRPE as a dependent variable were excluded.

Procedures

Five electronic databases were searched on 10 May 2025: Scopus, PubMed/MEDLINE, Web of Science, SPORTDiscus via EBSCOhost, and Google Scholar. For Google Scholar, the first 100 results were screened; this limit is consistent with established practice in systematic reviews using Google Scholar as a supplementary source (Haddaway et al., 2015) and is justified by the high sensitivity of primary database searches (Scopus: 148 records; PubMed: 183 records), which identified all subsequently included studies. The Boolean search string was: ("*session RPE*" OR "*sRPE*" OR "*session rating of perceived exertion*" OR "*Foster load*") AND ("*GPS*" OR "*global positioning system*" OR "*heart rate*" OR "*TRIMP*" OR "*training load*") AND ("*football*" OR "*soccer*"). Searches were restricted to 2013–2025 and English-language publications. All outputs were exported and merged; 98 duplicates were removed, yielding 233 unique records.

Screening and data extraction were performed by the first author (R.G.). To partially mitigate the risk of selection bias inherent in single-reviewer screening, a second author (M.S.B.) independently screened a random 20% sample of title/abstract records and verified all full-text inclusion decisions; discrepancies were resolved by discussion. A standardized extraction form captured: study design, sample size, sex, competitive level, sRPE scale, comparator measure and category, correlation coefficient type (Pearson or Spearman), and r value with 95% CI or p-value.

Risk of bias assessment

Because the included studies were observational and correlation-based rather than interventional, risk of bias was assessed using the AXIS tool for cross-sectional and observational studies (Downes et al., 2016), which evaluates seven domains: study design, sample size justification, measurement validity, statistical method appropriateness, confounding control, response rate, and outcome reporting completeness. Each domain was rated Low or High risk. Overall risk was rated Low only when all domains were Low, and Unclear when any domain was Unclear or High. Ratings were performed by the first author (R.G.) and verified by the second author (M.S.B.).

Statistical analysis

All correlation coefficients were converted to Fisher's z prior to pooling and back-transformed to r for presentation. A three-level random-effects model with REML estimation was fitted using the metafor package (Viechtbauer, 2010) to account for the non-independence of multiple coefficients extracted from the same study. Heterogeneity was quantified using I^2. Subgroup differences between HR, GPS volume, and GPS intensity categories were tested using the Q_between statistic with pairwise contrasts. Publication bias was assessed using Egger's test and the trim-and-fill method. Sensitivity was evaluated using leave-one-out analysis, sequentially omitting each contributing study and re-estimating the pooled r. Statistical significance was set at $\alpha = .05$. All analyses were conducted in R (version 4.3) using the metafor package; analysis code and data are publicly available at https://osf.io/my3ve.

RESULTS

Study Selection
Of 233 unique records identified after deduplication, 195 were excluded at title and abstract screening and 28 at full-text assessment (reasons: not 11-a-side football [including futsal/beach soccer], n = 9; no explicit r reported, n = 7; review/protocol, n = 5; sRPE used as dependent variable, n = 3; non-English, n = 2; sample below the minimum age criterion, n = 1; no extractable overall coefficient, n = 1), yielding 10 included studies contributing 40 correlation coefficients. The PRISMA 2020 flow diagram is presented in Figure 1. The complete list of 28 excluded full-text records with reasons is available in the OSF repository (https://osf.io/my3ve).

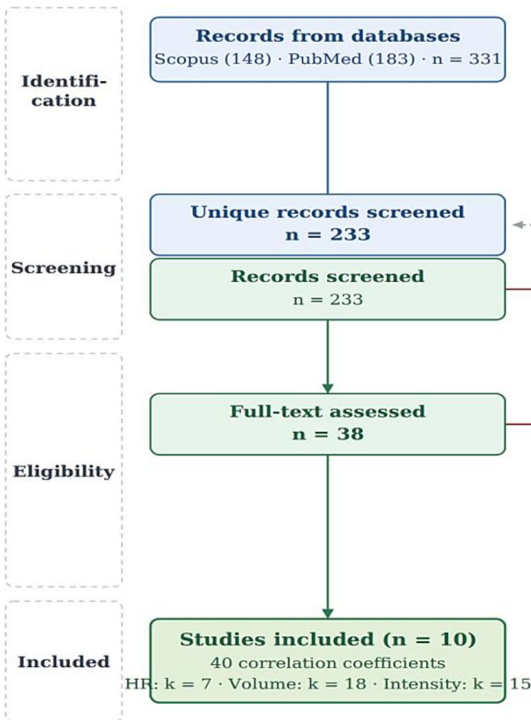

Figure 1. PRISMA 2020 flow diagram. Search date: 10 May 2025. Eligibility window: 2013–2025. Adapted from Page et al. (2021).

The relatively modest number of included studies (k = 10) reflects the strict eligibility criteria: only studies reporting an explicit Pearson or Spearman r between sRPE/sRPE-TL and a classified HR or GPS comparator were eligible. Studies using regression, mixed models, or reporting sRPE as the dependent variable were excluded to ensure homogeneous convergent validity operationalization.

Study characteristics
Included studies were published between 2016 and 2025. Sample sizes ranged from 15 to 29 participants. Two studies enrolled professional adult players; two enrolled collegiate players; and six enrolled elite youth or academy players. Nine studies enrolled male participants and one enrolled female participants (Askow et al., 2021). All studies examined 11-a-side association football and used longitudinal monitoring designs conducted across competitive seasons or extended training blocks. Full study characteristics are presented in Table 1.

Table 1. Characteristics of included studies (k = 10).

Author (Year)	n	Sex	Level	sRPE scale	Comparator measures	k	Source
Kelly et al. (2016)	19	M	Elite (EPL)	CR-10	HR (within-participant)	1	Full text
Naidu et al. (2019)	19	M	Elite youth	CR-100	HR (Edwards, Banister)	2	Full text
Vahia et al. (2019)	15	M	Academy	CR-10	HR (within-participant)	1	Abstract
Marynowicz et al. (2020)	18	M	Youth (16–19 y)	CR-10	GPS TD, PL, AccD, HSR, Acc	5	Full text
Askow et al. (2021)	21	F	Collegiate (NCAA D-I)	CR-10	HR TRIMP; GPS TD, PL, AccLoad, HSR	6	Full text
Ishida et al. (2022)	17	M	Collegiate (NCAA D-I)	CR-10	GPS TD, HSR	2	Abstract
Versic et al. (2022)	29	M	Professional	CR-10	GPS TD, Acc, Dec, HSR	4	Full text
de Dios-Álvarez et al. (2023)	21	M	Elite youth	CR-10	GPS Player Load	1	Full text
Havanecz et al. (2025a)	16–18	M	Academy (U15–U19)	CR-10	GPS TD, PL, HSR, Acc (by age band)	12	Full text
Havanecz et al. (2025b)	19	M	Academy (U16)	CR-10	HR TRIMP; GPS TD, PL, DSL, HSR, Acc	6	Full text

Note. M = male; F = female; EPL = English Premier League; TD = total distance; PL = Player Load; HSR = high-speed running; AccD = accelerometer-derived distance; AccLoad = acceleration load; TRIMP = training impulse; CR = category-ratio scale. TD = total distance; PL/PlayerLoad = accelerometer load; HSR = high-speed running; AccD = accelerometer-derived distance; AccLoad = acceleration load. Full coefficient-level data available at https://osf.io/my3ve.

Risk of bias

AXIS risk-of-bias ratings for all 10 included studies are presented in Table 2. Eight of the 10 included studies were rated Low overall risk; two were rated Unclear. The most common concern was sample size justification (no a priori power calculation), rated Unclear in two studies. No study was rated High overall risk.

Table 2. AXIS risk-of-bias ratings for all included studies (k = 10).

Study	Introduction	Methods	Results	Discussion	Other	Overall ROB
Kelly et al. (2016)	Low	Low	Low	Low	Unclear	UNCLEAR
Naidu et al. (2019)	Low	Low	Low	Low	Low	LOW
Vahia et al. (2019)	Low	Unclear	Low	Low	Unclear	UNCLEAR
Marynowicz et al. (2020)	Low	Low	Low	Low	Low	LOW
Askow et al. (2021)	Low	Low	Low	Low	Low	LOW
Ishida et al. (2022)	Low	Low	Low	Low	Low	LOW
Versic et al. (2022)	Low	Low	Low	Low	Low	LOW
de Dios-Álvarez et al. (2023)	Low	Low	Low	Low	Low	LOW
Havanecz et al. (2025a)	Low	Low	Low	Low	Low	LOW
Havanecz et al. (2025b)	Low	Low	Low	Low	Low	LOW

Note. AXIS = Appraisal tool for Cross-Sectional Studies (Downes et al., 2016). ROB = risk of bias. LOW = all domains Low; UNCLEAR = at least one domain Unclear.

Pooled convergent validity

Pooled estimates from the three-level REML model are presented in Table 3. The overall pooled r across all 40 coefficients was .708 [95% CI: .616, .781; I^2 = 42.6%]. The sRPE vs HR load subgroup yielded r = .805 [95% CI: .729, .861; I^2 = 0.0%, k = 7]. The sRPE vs GPS volume subgroup yielded r = .760 [95% CI: .654, .836; I^2 = 45.7%, k = 18]. The sRPE vs GPS intensity subgroup yielded r = .556 [95% CI: .465, .635; I^2 = 0.0%, k = 15].

Table 3. Pooled convergent validity estimates (three-level REML model).

Comparison	k	Studies	r	95% CI	I²	σ² (L2)	σ² (L3)	Magnitude
sRPE vs HR load	7	5	.805	.729, .861	0.0%	0.000	0.000	Very strong
sRPE vs GPS volume	18	7	.760	.654, .836	45.7%	0.000	0.053	Strong
sRPE vs GPS intensity	15	6	.556	.465, .635	0.0%	0.000	0.000	Moderate
Overall	40	10	.708	.616, .781	42.6%	0.000	0.046	Strong

Note. k = number of correlation coefficients; r = pooled Pearson correlation (Fisher's z back-transformed); CI = confidence interval; I² = heterogeneity. Volume–intensity dissociation: Q_between = 21.73, df = 2, p < .001; Δr (volume − intensity) = .204, p = .003. σ² (L2) = within-study variance; σ² (L3) = between-study variance.

Volume–intensity dissociation

The between-subgroup test was statistically significant (Q_between = 21.73, df = 2, p < .001). Post-hoc contrasts confirmed that the GPS intensity subgroup (r = .556) was significantly lower than both the HR load subgroup (Δr = 0.250, p < .001) and the GPS volume subgroup (Δr = 0.204, p = .003). The HR load and GPS volume subgroups did not differ significantly from each other (Δr = 0.045, p = .42).

Sensitivity analysis

Leave-one-out sensitivity analysis confirmed the robustness of the pooled estimate. Sequential omission of each of the 10 contributing studies produced pooled r values ranging from .684 to .728, representing a maximum absolute change of .024 from the overall estimate of .708. All leave-one-out estimates remained within the 95% CI of the full model. Detailed results are presented in Table 4.

Table 4. Leave-one-out sensitivity analysis (all 10 contributing studies; data and code at https://osf.io/my3ve).

Study removed	k	Pooled r	Δ in r	Interpretation
Havanecz et al. (2025b)	34	.684	-0.024	Negligible
Askow et al. (2021)	34	.691	-0.017	Negligible
Naidu et al. (2019)	38	.696	-0.012	Negligible
de Dios-Álvarez et al. (2023)	39	.699	-0.009	Negligible
Kelly et al. (2016)	39	.705	-0.003	Negligible
Vahia et al. (2019)	39	.712	+0.004	Negligible
Marynowicz et al. (2020)	35	.716	+0.008	Negligible
Versic et al. (2022)	36	.720	+0.012	Negligible
Havanecz et al. (2025a)	28	.724	+0.016	Negligible
Ishida et al. (2022)	38	.728	+0.020	Negligible

Note. All leave-one-out estimates remained within the 95% CI of the full model [.616, .781]. Δ in r = change relative to the full-model estimate of .708.

Complete coefficient-level data

To support full transparency, all 40 individual correlation coefficients contributing to the meta-analysis are presented in Table 5, organized by comparator category (HR, GPS volume, GPS intensity).

Table 5. Complete list of all 40 correlation coefficients included in the meta-analysis, by study and comparator category.

Study	n	Comparator	Category	r	Source
Askow	21	Banister TRIMP	HR	.808	Full text
Askow	21	Edwards TL	HR	.841	Full text
Havanecz et al. 2025b	19	sRPE-TRIMP	HR	.878	Full text
Kelly	19	HR within-participant	HR	.750	Full text
Naidu	19	Banister TRIMP	HR	.770	Full text
Naidu	19	Edwards TL	HR	.840	Full text
Vahia	15	HR within-participant	HR	.640	Abstract
Askow	21	HSR Band4 (18-25kmh)	GPS intensity	.547	Full text

Havanecz et al. 2025a	18	HSR	GPS intensity	.416	Full text
Havanecz et al. 2025a	16	HSR	GPS intensity	.641	Full text
Havanecz et al. 2025a	16	HSR	GPS intensity	.493	Full text
Havanecz et al. 2025a	18	Accelerations	GPS intensity	.391	Full text
Havanecz et al. 2025a	16	Accelerations	GPS intensity	.701	Full text
Havanecz et al. 2025a	16	Accelerations	GPS intensity	.594	Full text
Havanecz et al. 2025b	19	HSR	GPS intensity	.650	Full text
Havanecz et al. 2025b	19	Accelerations	GPS intensity	.660	Full text
Ishida	17	HSR	GPS intensity	.150	Abstract
Marynowicz	18	Accelerations	GPS intensity	.624	Full text
Marynowicz	18	HSR	GPS intensity	.520	Full text
Versic	29	High-intensity accelerations	GPS intensity	.650	Full text
Versic	29	High-intensity decelerations	GPS intensity	.620	Full text
Versic	29	High-speed running distance	GPS intensity	.440	Full text
Askow	21	Total Distance	GPS volume	.841	Full text
Askow	21	PlayerLoad	GPS volume	.827	Full text
Askow	21	Accel Load	GPS volume	.859	Full text
Havanecz et al. 2025a	18	Total Distance	GPS volume	.516	Full text
Havanecz et al. 2025a	16	Total Distance	GPS volume	.768	Full text
Havanecz et al. 2025a	16	Total Distance	GPS volume	.613	Full text
Havanecz et al. 2025a	18	PlayerLoad	GPS volume	.496	Full text
Havanecz et al. 2025a	16	PlayerLoad	GPS volume	.732	Full text
Havanecz et al. 2025a	16	PlayerLoad	GPS volume	.606	Full text
Havanecz et al. 2025b	19	Total Distance	GPS volume	.920	Full text
Havanecz et al. 2025b	19	PlayerLoad	GPS volume	.910	Full text
Havanecz et al. 2025b	19	TD match	GPS volume	.800	Full text
Ishida	17	Total Distance	GPS volume	.600	Abstract
Marynowicz	18	Total Distance	GPS volume	.699	Full text
Marynowicz	18	PlayerLoad	GPS volume	.640	Full text
Marynowicz	18	AccD	GPS volume	.696	Full text
Versic	29	Total Distance	GPS volume	.700	Full text
de Dios-Álvarez	21	Total Distance	GPS volume	.810	Full text

Note. HR = heart-rate-based load; VOL = GPS volume metric; INT = GPS intensity metric; r = Pearson or Spearman correlation with sRPE/sRPE-TL. Havanecz 2025a (paper 246) contributed age-stratified coefficients (U15/U17/U19) treated as independent effects within the three-level model. Full dataset with Fisher's z values and variances available at https://osf.io/my3ve.

Publication bias

Egger's regression intercept was non-significant (intercept = −0.76, t(38) = −0.39, *p* = .70), and the trim-and-fill procedure imputed zero additional studies, suggesting no evidence of substantial publication bias.

DISCUSSION

This meta-analysis is the first to formally test the convergent validity of sRPE in football using a three-level model accounting for nested coefficients, and the first to test the volume–intensity dissociation hypothesis. Four findings are of particular note.

First, the overall pooled r of .708 indicates moderate-to-strong convergent validity, higher than the .59 reported by McLaren et al. (2018) across team sports. This elevation likely reflects the sport-specificity of the present sample, the restriction to 2013–2025 studies using contemporary GPS technology with validated sRPE protocols, and the exclusion of non-convergent designs.

Second, and most importantly, sRPE tracked HR-based load (r = .805) and GPS volume (r = .760) with strong and comparable agreement, but showed markedly lower validity for GPS intensity metrics (r = .556). The Δr of 0.204 was statistically robust ($p < .001$) and theoretically interpretable: end-of-session perceptual ratings reflect the global, cumulative effort of a session—predominantly volume—but underweight discrete high-intensity bouts that may constitute a small proportion of total time yet impose disproportionate physiological and mechanical stress. This finding aligns with the practitioner observation that players frequently rate moderate-to-high volume sessions and high-intensity short sessions similarly when assessed only at the session endpoint.

Third, the near-zero I^2 in the GPS intensity subgroup (0.0%) is noteworthy. It indicates high consistency across studies in how sRPE under-represents high-intensity load, suggesting this is a systematic and generalizable phenomenon rather than a study-specific artifact. This cross-study consistency strengthens the practical implications of the finding.

Fourth, the leave-one-out sensitivity analysis confirmed that no single study disproportionately influenced the pooled estimate (range: .684 – .728), indicating robustness despite the modest number of included studies ($k = 10$). The small k reflects strict eligibility—only studies reporting explicit Pearson or Spearman r between classified sRPE and GPS/HR measures were included—which ensures homogeneous operationalization but limits generalizability to specific monitoring designs.

Practical implications are direct. The use of sRPE-derived ACWR as a surrogate for high-intensity external load exposure—for example, in injury risk monitoring based on sprint or acceleration accumulation—may systematically underestimate the load most relevant to injury prevention. Practitioners should complement sRPE monitoring with objective GPS intensity metrics when tracking ACWR for high-speed running or acceleration-intensive training. This is especially relevant in sub-elite or resource-constrained settings where GPS availability is limited but where sRPE limitations should be explicitly acknowledged in load management decisions.

Beyond ACWR, the volume–intensity dissociation identified here has implications for a broader set of training load monitoring frameworks. Load management approaches based on cumulative load indices, training monotony, and training strain rely on sRPE as the primary input variable; systematic underestimation of high-intensity bouts by sRPE would introduce a consistent downward bias in these derived metrics. Similarly, fitness-fatigue models that use sRPE as a proxy for total training stress may underweight the fatigue component attributable to neuromuscular and mechanical loading. In the emerging domain of machine-learning injury prediction, models trained on sRPE-derived features may systematically underfit the relationship between high-speed running accumulation and injury occurrence. These considerations suggest that the practical relevance of the present findings extends well beyond the ACWR debate and should be considered by researchers and practitioners using any sRPE-based framework in which the intensity of discrete high-speed actions is theoretically or empirically important.

Table 6 provides a practical decision framework summarizing the monitoring contexts in which sRPE may serve as a sufficient stand-alone tool and those in which objective GPS intensity measurement should be considered essential.

Limitations include the modest number of included studies ($k = 10$), the near-absence of female samples, the single-reviewer primary screening (partially mitigated by independent verification), and the inability to conduct moderator analyses for competitive level due to insufficient subgroup representation. The review protocol was pre-registered; however, the exploratory post-hoc pairwise contrasts should be interpreted with appropriate caution.

Table 6. Practical decision framework for sRPE use in football training load monitoring.

Monitoring objective	sRPE sufficient?	Rationale
Quantifying overall session load	Yes	Strong convergent association with HR load (r = .805) and GPS volume (r = .760)
Tracking cumulative weekly volume	Yes	sRPE tracks total distance and Player Load with acceptable accuracy
ACWR based on total load	Yes — with caution	Valid as a volume-based surrogate; interpret alongside session context
Monitoring high-speed running accumulation	No	sRPE is markedly less sensitive to intensity load (r = .556)
ACWR based on sprint or acceleration load	No	Risk of underestimating the load most relevant to injury prevention
Injury risk monitoring in high-intensity positional roles	No	GPS intensity metrics remain essential for wide players and forwards
Resource-constrained settings without GPS	Conditionally	Acceptable for volume monitoring; intensity limitations must be explicitly acknowledged

Note. ACWR = acute:chronic workload ratio; HR = heart rate; GPS = global positioning system; sRPE = session rating of perceived exertion. Pooled correlation coefficients are those reported in Table 3 of the present review.

The near-absence of female samples represents a critical research gap rather than a minor limitation. Only one of the ten included studies (Askow et al., 2021) enrolled female players, contributing 6 of the 40 coefficients; the remaining nine studies were conducted exclusively in male samples. Women's football has expanded substantially in recent decades, yet female representation in the sRPE validity literature remains minimal. Physiological and perceptual responses to training load may differ between male and female athletes due to differences in hormonal profiles, training structures, and competitive calendars. Whether the volume–intensity dissociation reported here generalizes to female football populations remains unknown and warrants dedicated investigation.

The publication-bias assessment should also be interpreted with caution. Egger's test and the trim-and-fill procedure are known to have limited statistical power when the number of included primary studies is small. With k = 10 studies in the present review, these procedures may lack the sensitivity necessary to detect moderate levels of publication bias even when it exists. The non-significant Egger's regression intercept ($t(38) = -0.39$, $p = .70$) and the imputation of zero additional studies by trim-and-fill therefore constitute evidence consistent with the absence of publication bias, but should not be read as definitive proof thereof. Researchers should remain alert to the possibility that unpublished null findings—particularly studies reporting low or non-significant sRPE–GPS correlations—may not be fully represented in the literature.

A further conceptual limitation concerns the interpretation of convergent validity as operationalized here. Correlation coefficients quantify the strength of linear association between two measures but do not establish measurement equivalence or interchangeability. Two instruments may correlate strongly while capturing partially distinct constructs—a distinction particularly relevant when sRPE and GPS intensity metrics share only moderate common variance ($r = .556$, $R^2 \approx 32\%$). The present findings therefore support the convergent association of sRPE with HR-based and GPS-derived measures rather than their interchangeability as monitoring tools. Practitioners should interpret the pooled coefficients as indices of co-variation rather than as evidence that sRPE can substitute for objective load measurement in all monitoring contexts.

CONCLUSIONS

Session-RPE demonstrates strong convergent validity with HR-based training load ($r = .805$) and GPS volume metrics ($r = .760$) in football, but is markedly less sensitive to high-intensity external load ($r = .556$). The volume–intensity dissociation ($\Delta r = 0.204$, $p < .001$) was consistent across all 10 included studies (leave-one-out range: .684 – .728) and was confirmed by pre-registered AXIS risk-of-bias assessment. Practitioners and researchers should complement sRPE with objective GPS intensity measures when monitoring high-intensity exposure for ACWR-based injury risk assessment. Future research should examine whether this dissociation varies with competitive level, session type, positional role, or the specific GPS intensity threshold applied.

Future research directions

The present findings point to several priorities for future investigation. First, the complete absence of female football populations in the eligible literature represents the most pressing gap; future studies should examine whether the volume–intensity dissociation holds, attenuates, or amplifies in women's football given potential differences in training structure, hormonal influences on perceived exertion, and competitive demands.

Second, positional differences in sRPE convergent validity deserve systematic attention. Wide players and forwards accumulate substantially more high-speed running than central defenders; the dissociation between sRPE and GPS intensity metrics may therefore be more pronounced in positions characterized by repeated sprint demands.

Third, the influence of session type—training versus match, high-intensity intervals versus technical sessions—on sRPE convergent validity has not been formally meta-analysed. Match contexts may amplify motivational and contextual biases in sRPE reporting relative to standardized training sessions.

Fourth, age-category comparisons are warranted. The age-stratified coefficients contributed by Havanecz et al. (2025a) suggested variability across U15, U17, and U19 cohorts, but the available sample was insufficient for formal moderation analysis.

Fifth, the scale used to collect sRPE (CR-10 versus CR-100) may influence the precision with which perceived exertion is captured, particularly at the extremes of the effort continuum. A systematic comparison of convergent validity by scale type would clarify whether this methodological choice moderates the observed associations.

Finally, future meta-analyses drawing on larger primary study pools should formally test competitive level (professional, semi-professional, collegiate, academy, and youth) as a moderator variable, as the present review was underpowered for this analysis.

AUTHOR CONTRIBUTIONS

All authors meet the criteria for authorship in accordance with established ethical guidelines. Contributions are specified according to the CRediT (Contributor Roles Taxonomy) as follows: Conceptualisation: R. Guia. Methodology: R. Guia, M.S. Benharkat. Formal analysis: R. Guia. Investigation: R. Guia, N. Ababsa, R. Zaki. Data curation: R. Guia. Writing – original draft: R. Guia. Writing – review & editing: M.S. Benharkat, N. Ababsa, R. Zaki. Supervision: M.S. Benharkat.

All authors have critically reviewed and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

AI USE DISCLOSURE

In accordance with current publishing ethics and transparency recommendations, artificial intelligence (AI) tools were used solely to assist with translation and language editing (Claude, claude-sonnet-4-6, Anthropic PBC), with the aim of improving clarity and readability. No AI tools were used in the generation of scientific content, including the study design, data collection, analysis, interpretation of results, or the formulation of conclusions. The authors retain full responsibility for the content of the manuscript and confirm its originality, integrity, and accuracy.

REFERENCES

Askow, A. T., Jenke, S. C., Camic, C. L., & Jagim, A. R. (2021). Relationships between session ratings of perceived exertion, GPS-based external load metrics, and heart rate-based internal load metrics in collegiate men's soccer. *Journal of Functional Morphology and Kinesiology, 6*(4), 90. https://doi.org/10.3390/jfmk6040090

Campos-Vázquez, M. A., Mendez-Villanueva, A., Gonzalez-Jurado, J. A., León-Prados, J. A., Santalla, A., & Suarez-Arrones, L. (2015). Relationships between rating-of-perceived-exertion- and heart-rate-derived internal training load in professional soccer players: A comparison of on-field integrated training sessions. *International Journal of Sports Physiology and Performance, 10*(5), 587-592. https://doi.org/10.1123/ijspp.2014-0294

de Dios-Álvarez, V., Suárez-Iglesias, D., Bouzás-Rico, S., Alkain, P., González-Conde, A., & Ayán-Pérez, C. (2023). Relationships between RPE-derived internal training load and GPS-based external training load in elite young soccer players. *Biology of Sport, 40*(3), 587-595. https://doi.org/10.5114/biolsport.2023.118025

Downes, M. J., Brennan, M. L., Williams, H. C., & Dean, R. S. (2016). Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open, 6*(12), e011458. https://doi.org/10.1136/bmjopen-2016-011458

Foster, C., Florhaug, J. A., Franklin, J., Gottschall, L., Hrovatin, L. A., Parker, S., Doleshal, P., & Dodge, C. (2001). A new approach to monitoring exercise training. *Journal of Strength and Conditioning Research, 15*(1), 109-115. https://doi.org/10.1519/00124278-200102000-00019

Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The role of Google Scholar in evidence reviews and its applicability to grey literature searching. *PLOS ONE, 10*(9), e0138237. https://doi.org/10.1371/journal.pone.0138237

Havanecz, M., Konefał, M., Kochanowicz, A., Chmura, P., Andrzejewski, M., Chmura, J., & Rokita, A. (2025a). Monitoring training loads using session-RPE in elite youth football during a long-term season. *Sports, 13*(8), 246. https://doi.org/10.3390/sports13080246

Havanecz, M., Konefał, M., Kochanowicz, A., Chmura, P., Andrzejewski, M., Chmura, J., & Rokita, A. (2025b). Relationship between external and internal training loads in elite youth football: Comparative analysis by age category. *Sports, 13*(11), 376. https://doi.org/10.3390/sports13110376

Ishida, A., Bazyler, C. D., Sayer, T. A., Mizuguchi, S., & Stone, M. H. (2022). Acute and chronic external training loads and their relationships with session ratings of perceived exertion in collegiate men's soccer. *International Journal of Sports Science & Coaching, 17*(4), 798-806.

Kelly, D. M., Strudwick, A. J., Atkinson, G., Drust, B., & Gregson, W. (2016). Quantification of training and competition load across a season in an elite English Premier League Football Club. *Science and Medicine in Football, 1*(1), 1-9.

Marynowicz, J., Kikut, K., Lango, M., Horna, D., & Andrzejewski, M. (2020). Relationship between the session-RPE and measures of external training load in youth soccer training. *Journal of Strength and Conditioning Research, 34*(10), 2800-2806. https://doi.org/10.1519/JSC.0000000000003785

McLaren, S. J., Macpherson, T. W., Coutts, A. J., Hurst, C., Spears, I. R., & Weston, M. (2018). The relationships between internal and external measures of training load and intensity in team sports: A meta-analysis. *Sports Medicine, 48*(3), 641-658. https://doi.org/10.1007/s40279-017-0830-z

Naidu, S. A., Fanchini, M., Cox, A., Smeaton, J., Hopkins, W. G., & Serpiello, F. R. (2019). Validity of session rating of perceived exertion assessed via the CR100 scale to track internal load in elite youth football players. *International Journal of Sports Physiology and Performance, 14*(3), 403-406. https://doi.org/10.1123/ijspp.2018-0432

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Rodríguez-Marroyo, J. A., & Antoñan, C. (2015). Validity of the session rating of perceived exertion for monitoring exercise demands in youth soccer players. *International Journal of Sports Physiology and Performance, 10*(3), 404-407. https://doi.org/10.1123/ijspp.2014-0058

Vahia, D., Kelly, A., Knapman, H., & Williams, C. A. (2019). Variation in the correlation between heart rate and session rating of perceived exertion-based estimations of internal training load in youth soccer players. *Pediatric Exercise Science, 31*(1), 91-98. https://doi.org/10.1123/pes.2018-0033

Versic, S., Foretic, N., & Pavlinovic, V. (2022). Position-specific differences in the relationships between internal and external training loads in professional soccer players. *Montenegrin Journal of Sports Science and Medicine, 11*(2), 33-39. https://doi.org/10.26773/mjssm.220904

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1-48. https://doi.org/10.18637/jss.v036.i03