Sports Medicine - Open

SYSTEMATIC REVIEW

Open Access

Reliability, Device Agreement and Validity of Load–Velocity Profiles: A Systematic Review with Meta-analysis

Nina Claassen[1], Stanislav Dimitri Siegel[2,5,6], Mareike Sproll[3], Niklas Lebelt[2], Alyssa Virginia Bargende[2], Anne Marieke Fasold[2] and Konstantin Warneke[4*]

Abstract

Background For a valid one-repetition maximum (1RM) prediction via load–velocity (LV) relationships, high reliability and accuracy must be assumed.

Objective Since individual study results indicate ambivalent prediction, this systematic review and meta-analysis was designed to provide a updated and comprehensive overview, extending knowledge about the validity and reliability of commercially available velocity sensors in Part I and the validity and reliability of velocity-based 1RM prediction models in Part II.

Methods A systematic literature search was conducted in PubMed/MEDLINE, Web of Science, and Scopus. Validity and/or reliability studies or velocity-based 1RM prediction evaluations were included. Methodological quality was assessed using adapted COSMIN. The analysis was performed for intraclass correlation coefficient (ICC), Lin's concordance correlation coefficient (CCC), and Pearson's correlation coefficient (*r*). The review was preregistered in PROSPERO (CRD42025634595).

Results Sixty-three studies were included for sensor validity and reliability and 38 for 1RM prediction models. Part I: Velocity sensors demonstrated good-to-excellent pooled validity and device agreement (ICC = 0.91–0.92 [0.83–0.97]; $k = 55$ and 439, respectively); intra- and inter-day reliability were classified as good to excellent with ICC = 0.90–0.91 [0.85–0.95] ($k = 228$ and 608, respectively), with sensor technology moderating the results. However, substantial heterogeneity and wide ranges of study-level estimates indicated considerable variability across moderators, linear position transducer (LPT) generally showing more consistent performance than inertial measurement units (IMU). Part II: Velocity-based 1RM prediction showed ICCs = 0.90 [0.83–0.94] ($k = 124$) and ICC = 0.91 [0.72–0.98] ($k = 9$); for reliability and validity, respectively.

Discussion Commercial velocity sensors generally provide high relative validity and reliability. Results varied depending on exercise complexity, intensity, sensor technology, and modeling approach. While velocity-based 1RM prediction demonstrated high average validity, large heterogeneity in lower body exercises significantly biased the results. Furthermore, the dearth of measurement error and agreement analyses prohibits final conclusions.

Conclusion Therefore, velocity-based monitoring and 1RM prediction require cautious interpretation, as sensor- and exercise-specific evidence remains limited.

*Correspondence:
Konstantin Warneke
konstantin.warneke@uni-jena.de
Full list of author information is available at the end of the article

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Key Points

- For a valid one-repetition maximum (1RM) prediction via load-velocity relationships, high reliability and accuracy must be assumed.
- This systematic review included 63 studies that revealed good to excellent overall validity and device agreement as well as reliability.
- Exercise type, velocity sensor, and intensity, among others, moderated results.
- A dearth of systematic and random error evaluations calls for appropriate statistical calculations.

Keywords Velocity-based training, Load–velocity relationship, One-repetition maximum prediction, Movement velocity, Resistance training, Measurement validity, Measurement reliability, Wearable sensors, Meta-analysis

Introduction

The precise determination of maximal strength plays a vital role when designing resistance training programs. Herein, training intensity has been described to significantly moderate morphological and functional adaptations and is commonly calculated as a percentage of the one-repetition maximum (1RM). To achieve muscle hypertrophy, 60–85% of the 1RM [1] were sufficient, while higher intensities of > 85% up to 100% were recommended for maximal strength training [2].

Training control on a microdosing level should consider daily performance variations potentially induced via life stress, training load-induced fatigue or recovery, among others [3, 4]. Since daily maximal strength testing before each training session is considered impractical, load–velocity profiles were frequently highlighted in the literature as a practical and valid alternative [5, 6]. In this approach, the bar velocity is measured when performing submaximal repetitions with, e.g., 50% and 70% of the 1RM to approximate the maximal achievable load by extrapolating the 1RM by assuming a minimal velocity threshold based on a hypothesized inverse linear relationship between load (in kg or %1RM) and movement velocity (mean, peak, or mean propulsive velocity) [7, 8]. Recent practical recommendations suggest that individualized load–velocity relationships can be established using as few as two submaximal loads, provided that one load is relatively heavy (> 75% 1RM) and that the selected loads produce a sufficiently large velocity difference of approximately 0.50 m s^{-1} to ensure a stable regression slope [6]. Such protocols can be implemented into warm-up routines, potentially reducing the need for periodic and time-consuming maximal strength testing [5]. Individual studies indicated strong correlations ($r > 0.9$) with deviations < 10 kg from the measured 1RM, rationalizing sufficient construct validity assumptions [7, 8].

Despite frequent recommendations for practical application [9–13], evidence supporting the valid estimation of maximal strength is ambivalent in quality and results, with studies reporting high inter-individual variability in movement velocity contradicting these recommendations [14, 15]. For instance, serious overestimations of the maximal strength from velocity profiles built in warm-up sets were reported in several studies, with systematic overestimations of up to 30 kg [7]. Although the predictability when using high loads near to the 1RM (e.g., 90% 1RM) and slow velocities improved [7, 16], practically relevant measurement errors called the approach into question [17]. Mean errors of up to 20% [18] were found when compared with the actual 1RM for lower [7, 16, 19–21] and upper body strength tests [19]. Numerous sensor types exist. These include linear position transducers (LPT), e.g., GymAware or Tendo, or inertial measurement units (IMUs) like PUSH or EnodePro, with differences in validity [22, 23], especially during complex, multi-joint movements such as squats or Olympic lifts [24, 25]. Numerous sources of variance stem from chosen exercise [23, 26], sensor placement [27], applied velocity threshold choice (minimal versus 0-velocity assumption in maximal strength test) [17], or the number and range of loads used to construct the load–velocity relationship [15] (two- or multipoint method [18, 25]) that affected the accuracy of the maximal strength prediction.

Beyond these influences causing systematic bias, sensor as well as test reliability and validity are of decisive importance in empirical data collections [28–30] and are crucial to allow reasonable conclusions on prediction errors in velocity-based 1RM estimations. Reliability describes the extent to which measurements can consistently distinguish between individuals or remain stable across repeated trials, whereas validity refers to the extent to which a sensor or prediction model corresponds to an appropriate reference criterion. Agreement, in contrast, describes the absolute closeness between repeated measurements or between sensors and is therefore crucial when evaluating sensor interchangeability. If criteria such as intra- and inter-day reliability and repeatability, objectivity, and validity are not fulfilled, internal data validity

as the prerequisite for external and construct validity is violated [29]. If a data collection performed several times in a row results in different outcomes (repeatability problem), or the test outcome depends on the investigator/device collecting the data (objectivity problem), the results cannot be valid, per se [30]. However, detailed measurement error analyses in reliability and validity analyses accounting for different measurement error sources are scarce.

Although previous reviews have addressed commercially available velocity sensors [31] and LV-based 1RM prediction methods [11], neither the measurement quality of the underlying sensors nor device-to-device agreement was quantitatively synthesized in relation to 1RM prediction. The present review extends this work by first evaluating the validity, reliability, and agreement of commercially available velocity sensors and then examining the reliability and validity of velocity-based 1RM prediction models. In addition, updated evidence and moderator analyses were included for sensor technology, exercise type, intensity, velocity metric, and prediction method.

Methods

The study protocol adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (see checklist in Supplemental Material S21) [32], and the study was conducted under consideration of ethical publishing standards for systematic reviews [33] and was registered in the PROSPERO database (CRD42025634595).

Part I addresses the validity, agreement, and reliability of velocity measurements obtained using commercially available velocity sensors, while Part II evaluates the reliability and validity of various approaches *to predict the 1RM.* Part-specific literature searches, eligibility screening, data extraction, and statistical analysis were performed.

Literature Search

The "Participants Intervention Comparison Outcome" (PICO) framework [34] has limited validity for search term production. The search was performed in the scientific databases PubMed/MEDLINE, Web of Science (Core collection), and Scopus. Four investigators (NL, AB, AF, NC) independently performed the systematic literature search based on the following eligibility criteria for study inclusion:

Participants: healthy participants (no restriction for age, training level, type of sport) all available velocity sensors were included.

Outcome: reliability and validity metrics commonly used in the literature, including correlation coefficients (r) and intraclass correlation coefficients.

Study types: reliability and validity studies.

Studies were excluded if missing specific velocity parameters or lacked reliability metrics.

Specific search terms were created based on these criteria. The following search terms were used for PubMed/MEDLINE:

Part I:

(("validity" OR "agreement" OR "reliability" OR "Intraclass correlation coefficient" OR "ICC" OR "test-retest reliability" OR "measurement reliability" OR "consistency" OR "repeatability")
AND
("velocity-based" OR "force-velocity" OR "load-velocity" OR "speed measurement" OR "velocity measurement" OR "velocity sensor" OR "load-velocity profiling" OR
tracking" OR "motion capture" OR "inertial sensor" OR "displacement transducer" OR "linear position transducer"))

This search term was developed on 15 November 2024 and subsequently updated in April 2025.

Part II:

("load velocity" OR "velocity profile" OR "velocity-based profile*" OR "velocity profiling" OR "load-velocity profiling" OR "velocity-load relationship".*
AND
("1RM" OR "one repetition maximum" OR "one rep max" OR "one-rep max" OR "maximal strength" OR "strength prediction" OR "strength estimation").
AND
("validity" OR "reliability" OR "agreement" OR "measurement" OR "accuracy" OR "consistency" OR "Intraclass correlation coefficient" OR "ICC" OR "test–retest").

For this Part, the term was developed in June 2025. Further terms are provided in the Supplemental Material (S1 Search terms).

Quality Assessment

Methodological study quality was evaluated via the COSMIN Risk of Bias Checklist for measurement instrument evaluations [35], using domains for *reliability* (Box 6) and *measurement error* (Box 7). Relevant design requirements and statistical considerations were simplified into five items. This approach was adopted from previous systematic reviews with similar research questions (e.g., [36]).

Part I:

- I1: Participant Selection: representativeness, comparability, and a sample size of ≥ 30 were consid-

ered a minimum threshold to avoid an inadequate rating according to COSMIN's general guidance for measurement reliability and validity analyses.

- I2: Measurement Protocol: clarity and standardization of sensor placement and test procedures. A study was scored "low" if essential details on sensor positioning, test environment, and movement execution were missing.
- I3: Reliability: specification of ICC or coefficient of variation (CV), with 95% confidence intervals, and appropriate test–retest design (minimum 1-day separation, same protocol conditions). Absence of reliability metrics or use of non-validated metrics (e.g., Pearson's *r* only) downgraded.
- I4: Validity: presence of comparison with a gold standard (e.g., force plate, 3D motion capture), and application of at least one accepted metric such as Bland–Altman Limits of Agreement (LoA), mean absolute error (MAE), mean absolute percentage error (MAPE), standard error of the mean (SEM), or root mean square error (RMSE). If no comparison or inappropriate reference (e.g., self-report) [27] was used, the item was scored "low".
- I5: Statistical Rigor: absence of major statistical flaws, such as missing units, inappropriate test selection (e.g., use of t-test for ICC), or lack of error margins. Ambiguities or partially reported results led to a downgrade.

Part II: For studies evaluating 1RM prediction models, definitions were adjusted:

- I2: Measurement Protocol: emphasized the clarity of procedures used to generate predictor variables (e.g., barbell velocity, submaximal loads) and the standardization of movement execution across conditions.
- I3: Reliability: focused on the test–retest consistency of model predictions and underlying input variables.
- I4: Validity: required a direct comparison between predicted and actual 1RM values, with appropriate statistical analysis (e.g., Bland–Altman (BA) analysis, standard error of estimate, or R^2 values) and use of a true 1RM test as reference.
- I5: Statistical Rigor: additionally considered modeling quality (e.g., overfitting, validation methods, appropriate regression analysis) and reporting completeness (e.g., confidence intervals, error margins).

Each assessment item was rated on a three-point scale, with a maximum possible score of 15 points. Higher scores indicated better methodological quality. The quality assessment was conducted independently by two authors (AB and NL). Disagreements were solved by consulting a third person (NC).

Data Processing and Statistics

The reliability- or/and validity metrics extraction process was conducted independently by four investigators (AF, AB, NL and NC). Missing information was requested from the corresponding author of the respective study via e-mail or ResearchGate. No response after the second reminder led to study exclusion.

Statistical analyses were conducted using RStudio (version 2024.09.1). The analysis leveraged several specialized R packages, including metafor, mixtools, ggplot2, and dplyr [37], and followed methodological approaches from previous high-quality analyses [38]. For reliability studies reporting intraclass correlation coefficients, the ICC variance was computed analytically following the method proposed by Donner [39], defined as:

$$Var\,(ICC) \;=\; \frac{2\,(1\;-\;ICC)^2\,(1\;+\;(n\;-\;1)\,ICC)^2}{k\;*\;n\,(n\;-\;1)},$$

where Var = variance, ICC = intraclass correlation coefficient, k = number of repeated measurements, n = number of participants. This estimator was originally derived for ICC(1,1) but was also applied to ICC(2,1) and ICC(3,1), as primary studies rarely reported the within-subject variance component required for estimating the sampling variance of these ICC formulations [40]. ICCs were analyzed on the logit scale and back-transformed for interpretation.

For validity studies reporting correlation coefficients, Fisher's z-transformation was applied to stabilize the variance [41], and were back-transformed to *r* for interpretation. The corresponding sampling variance was calculated as:

$$Var\,(z) \;=\; \frac{1}{n\;-\;3},$$

where n = number of participants, z = fishers z transformed Pearson correlation coefficient.

To ensure consistency across included studies, only correlation coefficients (e.g., ICC, Pearson's *r*, or Lin's CCC) within the valid range of −1 to +1 and with a sample size greater than three were considered for meta-analytic modeling. The reliability-classification by Koo and Li [42] was applied, defining values ICC < 0.50 = poor, ICC = 0.50–0.75 = moderate, ICC = 0.75–0.90 = good, ICC > 0.90 = excellent.

A multilevel random-effects meta-analysis was conducted using restricted maximum likelihood (REML) estimation, with study entered as a random intercept to account for the non-independence of effect sizes within

studies and between-study heterogeneity. Between-study heterogeneity was assessed using Cochran's Q statistic [43] following Chi-square distributions. A statistically significant Q(df) value indicates the presence of substantial heterogeneity among included effects, suggesting that effect sizes vary across studies due to methodological or contextual differences rather than chance.

To assess the risk of publication bias, funnel plots were visually inspected for asymmetry across all outcome domains [44]. In addition, regression-based tests for funnel plot asymmetry (Egger-type tests) were applied to identify potential small-study effects [45].

To explore sources of heterogeneity, meta-regressions were fitted with the following variables as moderators:

1) Exercise category (accounting for movement complexity/stability (e.g., free-weight vs. machine-based exercises) [7, 27].
2) Measured variable (e.g., mean velocity, peak velocity) [5].
3) Training intensity (e.g., %1RM or categorical levels: low [< 50%], moderate [50–80%], high [> 80%] [2] as a continuous and categorical moderator [46]. The categorical model allows for a more intuitive interpretation of intensity levels.
4) Sensor type (LPTs versus IMUs) [47, 48].
5) Gold standard comparator (e.g., 3D motion capture vs. force plate) [49].

Results

Results Part I: Reliability and Validity of Different Velocity Sensors

Risk of Publication Bias

Funnel plots for all primary meta-analyses can be reviewed in the Supplementary Figures S2–S11. Overall, the risk of publication bias was ambivalent, with some visual inspection indicating asymmetry, while others indicated symmetrical distribution.

For device agreement (S2–S4), Egger's test revealed significant funnel plot asymmetry in ICC [$t(437) = -4.01$, $p < 0.001$] and Lin's CCC analyses [$t(62) = -2.16$, $p = 0.035$], suggesting the presence of small-study effects. In contrast, no statistically significant evidence of funnel plot asymmetry was found for Pearson-based agreement estimates ($p = 0.0937$), validity (S5-S6) via ICC ($p = 0.118$) or Pearson-based analyses ($p = 0.611$).

For interday reliability estimates, the funnel plot (S7-S9) suggested slight asymmetry, which was supported by a statistically significant Egger's test [$t(606) = 3.50$, $p < 0.001$ for ICC-based analysis and $t(42) = 7.34$, $p < 0.001$ for Pearson-based analysis], indicating the presence of small-study effects. Asymmetry assessments with a study number of $n < 10$ (e.g., interday reliability analysis with Lin's CCC and intraday reliability with Pearson's r) require cautious interpretation [50].

For intraday reliability via ICCs (Figure S10), the funnel plot showed a symmetrical distribution of smaller studies, although clustering of effects was evident [$t(226) = -4.23$, $p < 0.001$].

Search Results

The searches in the three databases returned 24,802 hits from which 63 studies were included in the analysis (Fig. 1). Among these, 0 studies included only females, 41 focused exclusively on male participants [22, 23, 26, 27, 46, 47, 49, 51–81] and the remaining 22 studies [24, 25, 82–102] included both sexes (Supplementary Table S15).

LPTs were used in 52 studies, including sensors such as SmartCoach [52], GymAware [22, 25, 46, 47, 49, 55, 59, 66, 69, 74, 78, 82, 85, 88, 90], T-force [22, 23, 53, 58, 61, 64, 66, 67, 71, 72, 81, 84, 87, 90, 94, 99, 102], Chronojump [58, 73, 89, 102], Humac [86, 101], Tendo [22, 62, 63, 80, 92, 95, 97], Vitruve [64, 76, 81], Speed4Lift [65, 87, 94], MuscleLab [79, 94], and 1080 Quantum System [85]. Additionally, 28 studies used IMUs, including Gyko [52], PUSH [25, 26, 47, 53, 57, 58, 71, 79, 85], Myotest [54, 56], Bar Sensei [82], EnodePro [24, 25, 65, 74, 83–85, 94], WIMU [61, 99], Jueying [75], IMS [68], Ergonauta [103] and PowerTool [82]. Six studies validated sensors via 3D motion capture systems, e.g., Vicon [51, 74, 84], Qualisys [49], Velowin [58] and OptiTrack [73]. Systems with combined measurements (e.g., LPTs with integrated force measurement, such as the 1080 Quantum System) were classified under the LPT category.

Most studies focused on lower-body exercises such as:

1. Squats [22, 24, 25, 46, 49, 51, 53, 56, 58, 59, 61–68, 70, 71, 73–76, 79, 80, 82, 84–90, 92, 94, 97, 99],
2. Deadlifts [47, 57, 88],
3. Hip thrusts [65] or
4. Jumps [54, 72, 85, 91, 103].

Twenty-three studies were performed for upper body exercises:

1. Bench presses [23, 24, 27, 46, 52, 58, 62, 64, 71, 74, 81, 87, 92, 94–96, 101, 102],
2. Bench press throw [27],
3. Complex exercises such as snatches [83, 85], landmine punch throws [69] and high pulls [77, 78].

Methodological Quality, Risk of Bias and Quality of Evidence

The studies achieved a mean score of 12.06 ± 1.27 points, suggesting an overall acceptable to good methodological

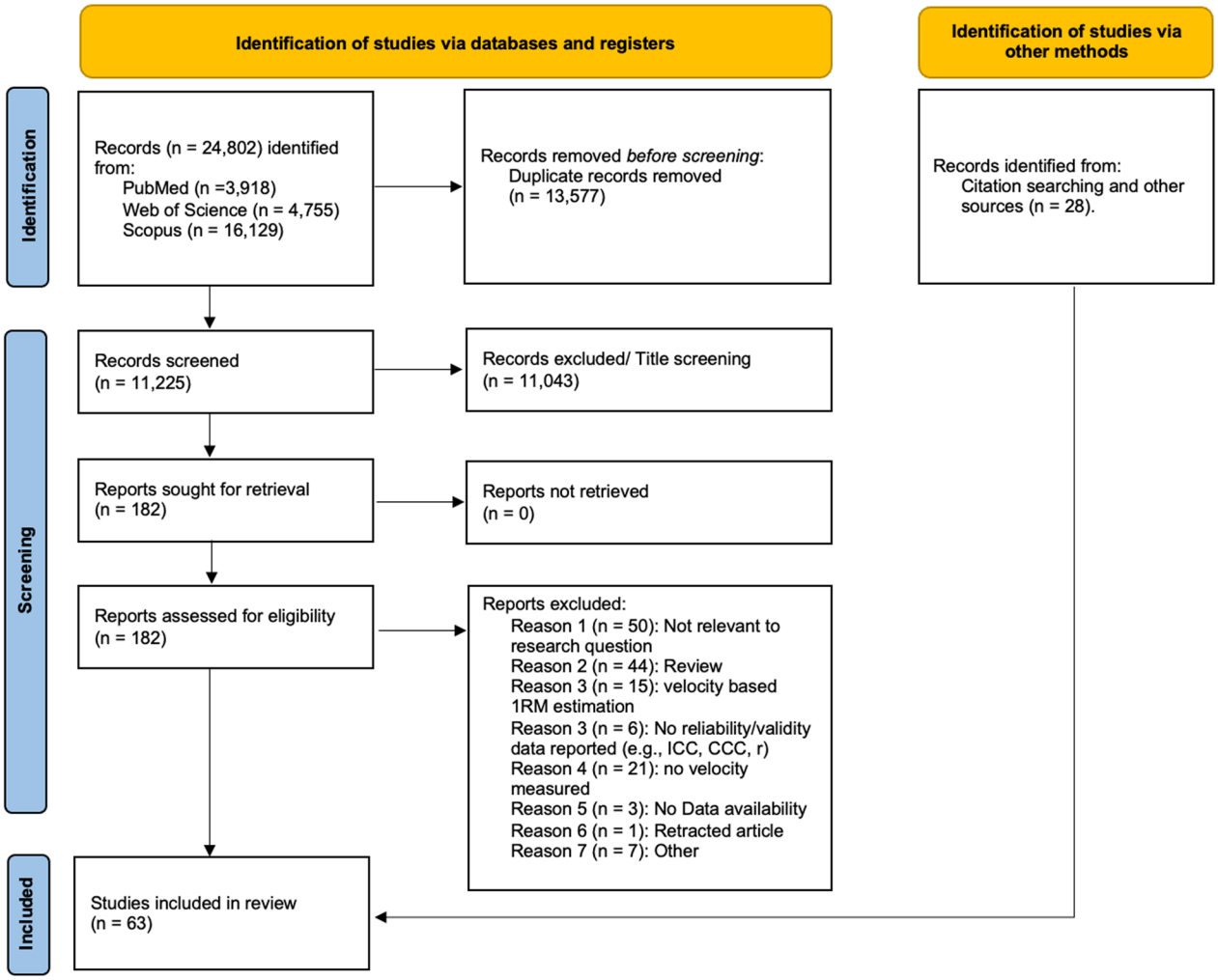

Fig. 1 PRISMA flowchart of the literature search and study selection process for Part I: validity, reliability, and device agreement of commercially available velocity sensors. *1RM* one-repetition-maximum, *ICC* intraclass correlation coefficient, *CCC* Lin's concordance correlation coefficient, *r* Pearson's correlation coefficient

quality [35] with the highest adapted COSMIN score of 14 [24, 25, 62, 64, 84, 92–94, 101, 103] and the lowest score of 9 [77] (Supplementary Table S17).

Validity of Different Velocity Sensors

Overall Effects Device validity ($N=11$) was evaluated by comparing the velocity sensor against a gold-standard system (Motion capture). Device agreement ($N=27$) describes the agreement between the velocity sensor and a criterion-reference system without gold-standard status (e.g., LPT, IMU). Validity based on ICC ($n=5$ studies) [49, 63, 70, 74, 96] demonstrated high validity (ICC = 0.91 [0.87–0.94]) ranging from ICC = 0.65 [74]–0.98 [49]. Additionally, linear association was assessed via Pearson's r in 8 studies [22, 51, 70, 74, 84, 85, 100, 103] ($r=0.95$ [0.82–0.98]) ranging from 0.55 [22] to 0.99 [85], providing supplementary information on correlation but not on absolute agreement.

Since analytical heterogeneity was evident [$Q(54)=689.26$, $p<0.001$], subgroup analyses were performed to identify meaningful moderators.

Moderator Analysis Exercise category ($p=0.196$–0.684), type of gold standard reference ($p=0.130$–0.143), as well as intensity ($p=0.121$) did not reach the level of significance.

Velocity variables significantly moderated validity in correlation statistics ($p<0.001$). However, the beta coefficients for neither peak velocity nor mean velocity were significant ($p=0.201$–0.538).

Sensor type regression ($p<0.001$) and close individual study examination indicated differences between IMU and LPT. Although Quantum, as a bar-mounted hybrid

sensor, is the only sensor with a significant beta value ($\beta = 2.99$, $p = 0.002$) and correlations of $r = 0.97–0.99$, a wide range of validity outcomes was reported with 7 out of 8 IMU studies [22, 51, 70, 74, 84, 85, 96, 100] providing a range of $r = 0.40–0.99$ [22, 51, 70, 74, 84, 85, 100] and ICC = 0.65–0.96 [70, 74, 96]. Most studies indicated moderate to good validity with $r = 0.55–0.61$ [22] in the ballistic squat velocity for the Myotest IMU or $r = 0.40–0.83$ for the EnodePro [84]. In contrast, LPTs showed validity values with $r = 0.77–0.99$ and ICC = 0.87–0.98 [22, 49, 63, 85, 96, 103].

Device Agreement

Overall Effects Across 439 effect sizes ($N = 10$ [24, 62, 64, 73, 79, 80, 82, 86, 92, 101]) device agreement showed good-to-excellent pooled ICC = 0.92 [0.83–0.97] between sensors, individual study estimates ranged from poor to excellent (ICC = 0.19 [52]–0.99 [24, 58, 64, 87, 92, 94]). Lin's CCC was reported in 3 studies [64, 67, 80] with CCC = 0.90 [0.52–0.98]. The lowest CCC was found with CCC = 0.36 [80] and the highest CCC = 0.99 [67]. As Pearson's r reflects association rather than agreement, it was treated as a secondary metric. Pearson correlations were reported in 19 studies [24, 47, 52, 53, 57, 61, 65, 68, 71, 74–77, 79, 81, 84, 89, 92, 104] and indicated a pooled strong linear association ($r = 0.91$ [0.84–0.95]). However, individual studies ranged from $r = 0.2$ [52] to $r = 0.99$ [24, 77, 81, 92].

Since analytical heterogeneity was evident [ICC: $Q(438) = 1317.53$, CCC: $Q(63) = 1083.51$, r: $Q(226) = 1550.48$, all $p < 0.001$], subgroup analyses were performed to identify meaningful moderators (Fig. 2).

Moderator Analysis Velocity variable (analyzed with ICCs; $p = 0.630$) and intensity (analyzed with Lin's CCC, $p = 0.480$) did not reach the level of significance.

In contrast, in the ICC-based analysis, squat-based assessments were associated with higher agreement estimates ($\beta = 0.37$, $p < 0.001$, $n = 13$) with excellent ICC values for the T-Force-Tendo LPT comparison (ICC = 0.90) [62], and the RepOne sensor vs Tendo LPT (ICCs = 0.83–0.99). However, Lin's CCC indicated lower concordance

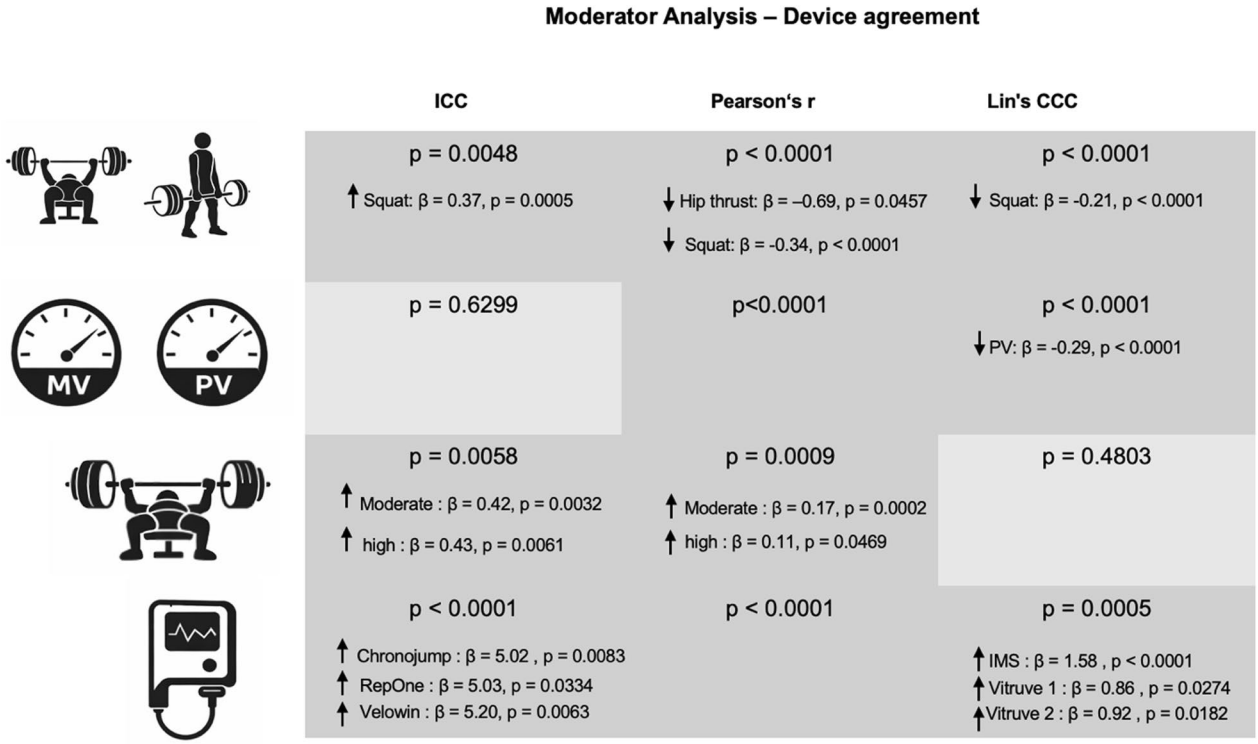

Fig. 2 Moderator analyses for device agreement. Results are shown separately for ICC-based, Pearson's *r*-based and Lin's CCC-based agreement analyses. Rows represent the investigated moderators (exercise type, velocity variable, training intensity, and sensor type). Shaded cells indicate statistically significant moderator effects based on the QM test ($p < 0.05$). Specific moderator levels are displayed only when the analysis reached statistical significance. Directional arrows summarize the direction of significant effects. *ICC* intraclass correlation coefficient, *r* Pearson's correlation coefficient, *Lin's CCC* Lin's concordance correlation coefficient, *p* corresponding significance level of the regression coefficient, β estimated meta-regression coefficient: positive β values indicate higher ICC, *r* or CCC values relative to the reference category, whereas negative β values indicate lower values

for squat movements ($\beta = -0.21$, $p < 0.001$, $n = 3$), and Pearson-based analyses showed a similar reduction in linear association ($\beta = -0.34$, $p < 0.001$, $n = 6$).

Additionally, the variable is a significant moderator in Lin's CCC and Pearson analyses. Peak velocity ($\beta = -0.29$, $p < 0.001$, $n = 3$) showed reduced concordance with CCC = 0.59–0.99 when IMU compared to T-force [64, 67], while mean velocity did not reach statistical significance. The Pearson-based device agreement model reinforced this pattern, with all velocity variables tending toward lower correlations, although none of the individual beta coefficients reached statistical significance.

Training intensity significantly moderated device agreement in both the ICC and Pearson-based analyses. The analysis revealed that agreement was significantly higher at moderate (ICC: $n = 10$; r: $n = 6$) and high (ICC: $n = 4$; r: $n = 2$) intensities compared with low intensities. Good to excellent correlations were found at moderate intensities (ICC = 0.87–0.99, $r = 0.81$–0.99) between RepOne and Tendo [92], indicating the best correlation at 60%1RM with r and ICC = 0.99. At high intensities, correlations remained high between these sensors [92] (ICC = 0.87–0.99), while for low intensities agreement was lower (ICC = 0.79–0.99) [92].

Sensor type significantly moderated device agreement in ICC- and Lin's CCC-based analyses. Certain LPTs [58, 92] demonstrated significantly higher agreement with T-force or Tendo (ICC = 0.82–0.99). In Lin's CCC-based analysis, IMS and Vitruve showed moderate-to-high concordance with T-force IMU (CCCs = 0.58–0.99) [64, 67]. Pearson-based regression revealed Gyko IMU ($\beta = -1.43$, $p = 0.0366$, $n = 1$) with significantly lower agreement with smartcoach LPT; however, this finding is based on a single study. For detailed moderator analysis, see Fig. 2.

Reliability of Velocity Measurements Obtained Using Different Velocity Sensors

Overall Effects For interday reliability, 29 studies reported ICCs [25, 27, 46, 54–56, 58–60, 62, 65, 67, 69, 72, 75, 78, 80, 82–84, 86, 88, 90, 92, 95, 97–99, 101]. The pooled estimate indicated good-to-excellent reliability with ICC = 0.90 [0.85–0.93]; however, individual study estimates ranged widely from poor to excellent (ICC = 0.14 [56]–0.99 [25, 58, 98]. One study reported Lin´s CCC [25] with CCC = 0.99. For interday reliability, Pearson correlations were reported in $N = 3$ studies [65, 69, 98] and were treated as supplementary association metrics rather than primary reliability evidence ($r = 0.91$ [0.38–0.99]) with a range of $r = 0.55$ [65] to 0.99 [98].

Thirteen intraday studies [23, 26, 53, 65, 66, 69, 83, 84, 91, 93, 101–103] indicated good-to-excellent reliability (ICC = 0.91 [0.87–0.95]) with individual studies ranging from 0.29 [23] to 0.99 [103]. One study [65] additionally reported Pearson's correlation coefficient ($r = 0.89$ [0.83–0.93]).

Since analytical heterogeneity was evident for Interday reliability [ICC: $Q(607) = 950.13$, CCC: $Q(5) = 30.61$], [Pearson's r $Q(43) = 232.07$); all $p < 0.001$], as well as for intraday ICCs [$Q(227) = 268.22$, $p = 0.032$], subgroup analyses were performed to identify meaningful moderators.

Moderator Analyses Intraday:

In the intraday analysis, no moderator effect could be found for exercise, velocity variable, or intensity ($p = 0.059$–0.684). Only sensor type significantly moderated intraday reliability (Table 1) when assessed with ICC ($p = 0.001$). The analysis revealed that only the Beast sensor ($\beta = -2.13$, $p = 0.004$, $n = 1$), demonstrated significantly lower reliability, whereas no other sensor showed statistically significant differences.

Interday:

Table 1 Sensor-specific range of velocity sensor reliability across interday and intraday analyses with significant beta values in italics

Sensor	Interday_ICC	Interday_*r*	Intraday_ICC
LPT			
Tendo	0.56–0.99	–	0.65–0.97
GymAware	0.54–0.99	*0.60–0.92*	0.61–0.92
Quantum	0.58–0.97	–	–
Chronojump	0.98–0.99	–	0.72–0.97
HUMAC	0.45–0.99	–	0.91–0.97
Speed4Lifts	0.56–0.75	*0.56–0.75*	0.81–0.97
PowerLift	–	–	0.74–0.87
ADR	0.94–0.99	0.90–0.99	0.75–0.96
Velowin	0.98–0.99	–	0.68–0.83
IMU			
PUSH	0.38–0.99	–	0.46–0.98
Beast	–	–	*0.29–0.64*
Myotest	0.14–0.96	–	–
WIMU	0.91–0.98	–	–
Bar Sensei	0.17–0.45	–	–
PowerTool	0.65–0.79	–	–
EnodePro	0.55–0.99	*0.55–0.81*	0.58–0.94
Jueying	0.79–0.95	–	–
ReoOne	0.59–0.98	–	–
Other			
Vicon (MoCap)	0.69–0.91	–	0.61–0.99
T-force (LVT)	0.7–0.99	–	0.77–0.99

ICC intraclass correlation coefficient, *r* Pearson's correlation coefficient, *LPT* linear position transducer, *IMU* inertial measurement unit, *LVT* linear velocity transducer

Intensity (ICC: $p=0.455$) did not reach the level of significance and could not be conducted for Lin's CCC and Pearson's r analyses.

For interday reliability assessed using Lin's CCC, the velocity variable showed a significant moderating effect ($p=0.0002$), with peak velocity being significantly lower ($\beta=-0.44$, $p<0.001$, $n=1$) for measurement with GymAware, PUSH, and EnodePro [25].

When Pearson's r was considered as a supplementary association metric for interday reliability, exercise category showed a significant moderating effect ($p<0.001$). Lower between-day correlations were observed for hip thrust ($\beta=-2.02$, $p<0.001$, $n=1$), landmine punch throw ($\beta=-1.58$, $p<0.001$, $n=1$), and squat ($\beta=-1.59$, $p<0.001$, $n=1$). For the squat, correlations were moderate for Speed4Lifts ($r=0.75$) and EnodePro ($r=0.81$) [65]. Lower associations were observed for the hip thrust (Speed4Lifts: $r=0.56$; EnodePro: $r=0.55$) [65], and landmine punch throw for GymAware ($r=0.60$–0.92) [69].

Sensor type significantly moderated reliability outcomes in all analyses, except for the intraday analysis based on Pearson's r ($p=0.1153$), with PUSH ($\beta=-0.51$, $p<0.001$, $n=1$), EnodePro ($\beta=-1.77$, $p<0.001$, $n=1$), and GymAware ($\beta=-1.58$, $p<0.001$, $n=1$) being sensors with the poorest reliability. As each estimate was derived from a single study, these findings should be regarded as exploratory. Twenty-one studies were conducted using LPTs ([25, 46, 55, 58–60, 62, 65, 67, 69, 78, 80, 83, 86, 88, 90, 92, 95, 97, 98, 101] and 13 studies [25, 54, 56, 58, 75, 78, 80, 82–84, 92, 99] on IMUs. A linear velocity transducer (LVT) is represented in all studies by the T-Force sensor [27, 58, 67, 72, 84, 95, 99]. Table 1 summarizes all sensors and reports the range of study-level effect sizes for each analysis across studies, exercises, and intensities.

Results Part II: 1RM Prediction Models

Risk of Publication Bias

Funnel plots for all primary meta-analyses can be reviewed in the Supplementary Figures S12–S14.

Search Results

Figure 3 displays the flowchart of the literature search and selection process for the 1RM predictions that returned 502 hits from which 38 studies were included.

Among these, 2 studies included only females [15, 105], 25 focused exclusively on male participants [7, 8, 11, 14, 16, 18–21, 106–121] and the remaining 10 studies included both sexes [118, 122–130] (Supplementary Table S15). One study [131] did not explicitly report the sex of the participants.

The majority of studies ($N=37$) applied linear load–velocity (LV) regression [7, 8, 14, 15, 18–21, 23, 57, 105–116, 118–132], five studies used polynomial regression [11, 20, 111–113], one study used quadratic regression [8], and one study employed multilinear regression [14]. One study applied regularized regression methods [14] and another study [14] implemented artificial neural networks. Across the studies included, a variety of exercises were used for 1RM prediction. Lower-body exercises were used in 25 studies:

1. Squats [7, 8, 15, 18, 19, 21, 109, 112, 122, 124, 127, 129, 132],
2. Deadlift [11, 25, 57, 105, 107, 114, 116, 132],
3. Snatch pull [128]
4. Knee extension [106, 129],
5. Ankle dorsiflexion, seated ankle plantarflexion, standing ankle plantarflexion, hip abduction, hip adduction, hip extension, hip flexion, and knee flexion [129].

Upper-body exercises were used in 22 studies:

1. Bench press throw [131],
2. Power clean [108, 121],
3. Prone bench pull [110, 125],
4. Bench press [14, 15, 18, 19, 111, 113, 115, 117–120, 123, 126, 127],
5. Horizontal pull [18] and lat pull [118],
6. Seated cable row [118]
7. Overhead press [130].

Methodological Quality, Risk of Bias and Quality of Evidence

The studies achieved a mean adapted COSMIN score of 12.24 ± 1.38 points, with a maximum of 15 points [27, 107] and the lowest score of 10 [8, 19, 105, 106, 116, 133]. A detailed table of methodological quality and study characteristics is provided in the supplementary materials (Table S16).

Reliability of Prediction Models

Overall Effects Thirteen studies [7, 19, 21, 107, 108, 111, 115, 120, 121, 124, 125, 132, 134] reported good-to-excellent interday reliability with pooled ICC = 0.90 [0.83–0.94], while individual study ICCs ranged from 0.14 [120] to 0.99 [124, 132] (Fig. 4).

Since analytical heterogeneity was evident [$Q(123)=7136.71$; $p<0.001$], subgroup analyses were performed to identify meaningful moderators.

Moderator Analysis Moderator analyses did not identify any statistically significant influence of prediction method ($p=0.750$) or used velocity ($p=0.205$).

In contrast, exercise type substantially influenced prediction reliability ($p<0.001$), with squat predictions

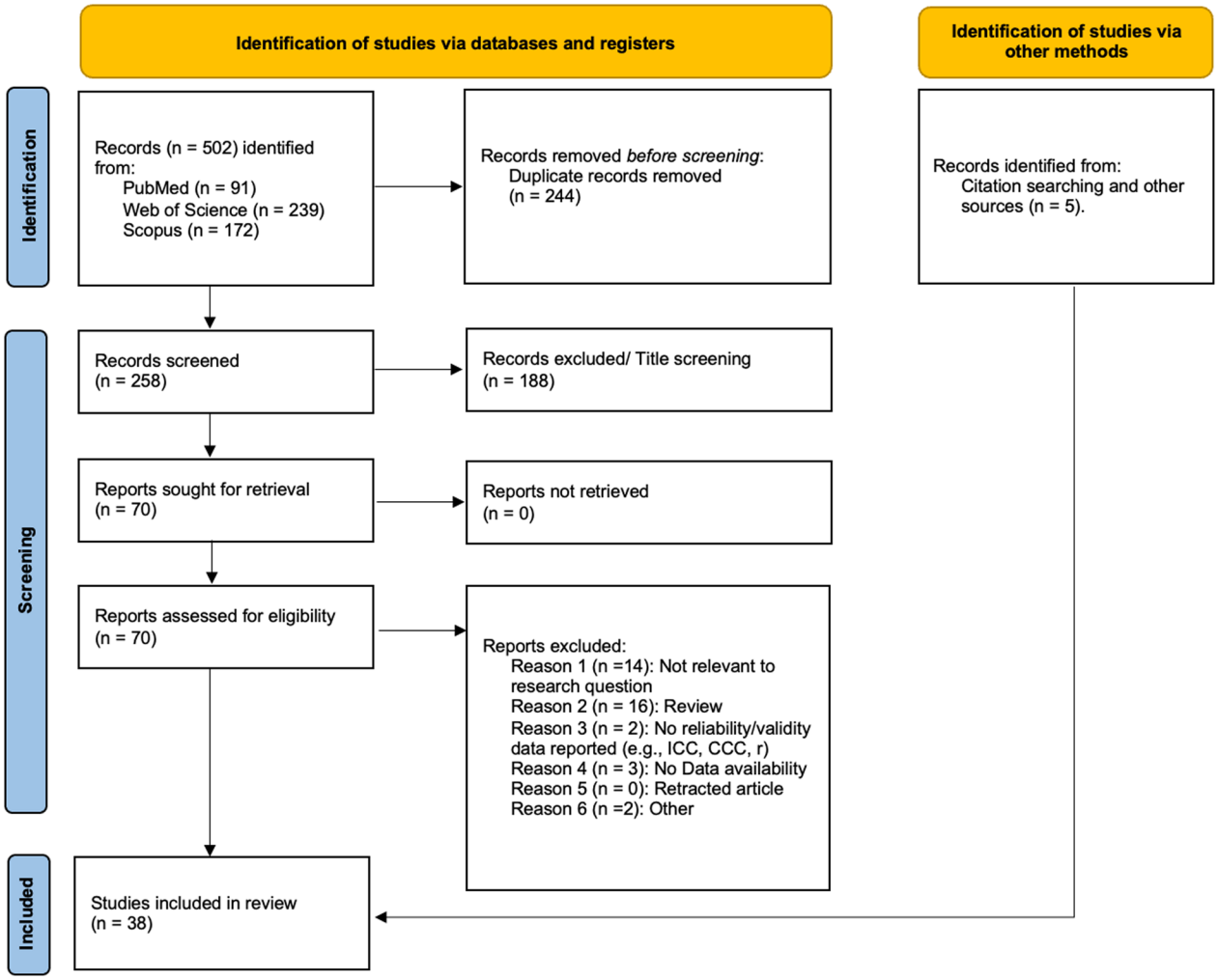

Fig. 3 PRISMA flowchart of the literature search and study selection process for Part II: reliability and validity of velocity-based 1RM prediction models. *ICC* intraclass correlation coefficient, *CCC* Lin's concordance correlation coefficient, *r* Pearson's correlation coefficient

showing lower reliability ($\beta = -1.28$, $p < 0.001$, $n = 5$) when using the LV method to predict 1RM. ICCs ranged between 0.61 [7, 19] and 0.96 [21] for multiple load points and ICC = 0.24–0.99 [132] for the two-point prediction model [129]. In contrast, other exercises did not show statistically significant deviations (all $p = 0.079$–0.768). Bench press–based predictions demonstrated moderate-to-excellent interday reliability for single-point approaches (ICCs = 0.64 [120]–0.91) [115] and ICC = 0.92–0.95 [111] and ICC = 0.89–0.96, respectively, for two-point or polynomial models [111, 120].

Validity of Prediction Models

Overall Effects Three studies [105, 122, 131] used the ICC and yielded moderate-to-excellent relative reliabilities for agreement (ICC = 0.91 [0.72–0.98]) ranging between ICC = 0.68 [131] and 0.98 [105]. Pearson correlations were used in 31 studies [7, 9, 11, 14–16, 18, 19, 25, 107, 109, 111–120, 123–130, 135, 136] with an overall excellent pooled correlation ($r = 0.96$ [0.94–0.97]); however individual studies ranged between 0.44 [109] and 0.99 [8, 14, 113, 117, 118, 126, 127, 133].

Since analytical heterogeneity was evident [ICC: $Q(8) = 914.14$; r: $Q(187) = 1472.50$, $p < 0.001$], subgroup analyses were performed to identify meaningful moderators.

Moderator Analysis In the validity analysis, no moderator effect was evident for the velocity variable (mean vs peak velocity) ($p = 0.486$).

In contrast, exercise type significantly influenced 1RM prediction in both analyses ($p < 0.001$), with squat-based predictions ($\beta = -0.98$, $p < 0.001$, $n = 9$) [7, 8, 15, 18, 19, 109, 112, 127, 129] and deadlift predictions ($\beta = -0.94$, $p < 0.001$, $n = 7$) [11, 16, 20, 107, 114, 116, 126] demonstrating lower validity. Some upper-body pulling exercises

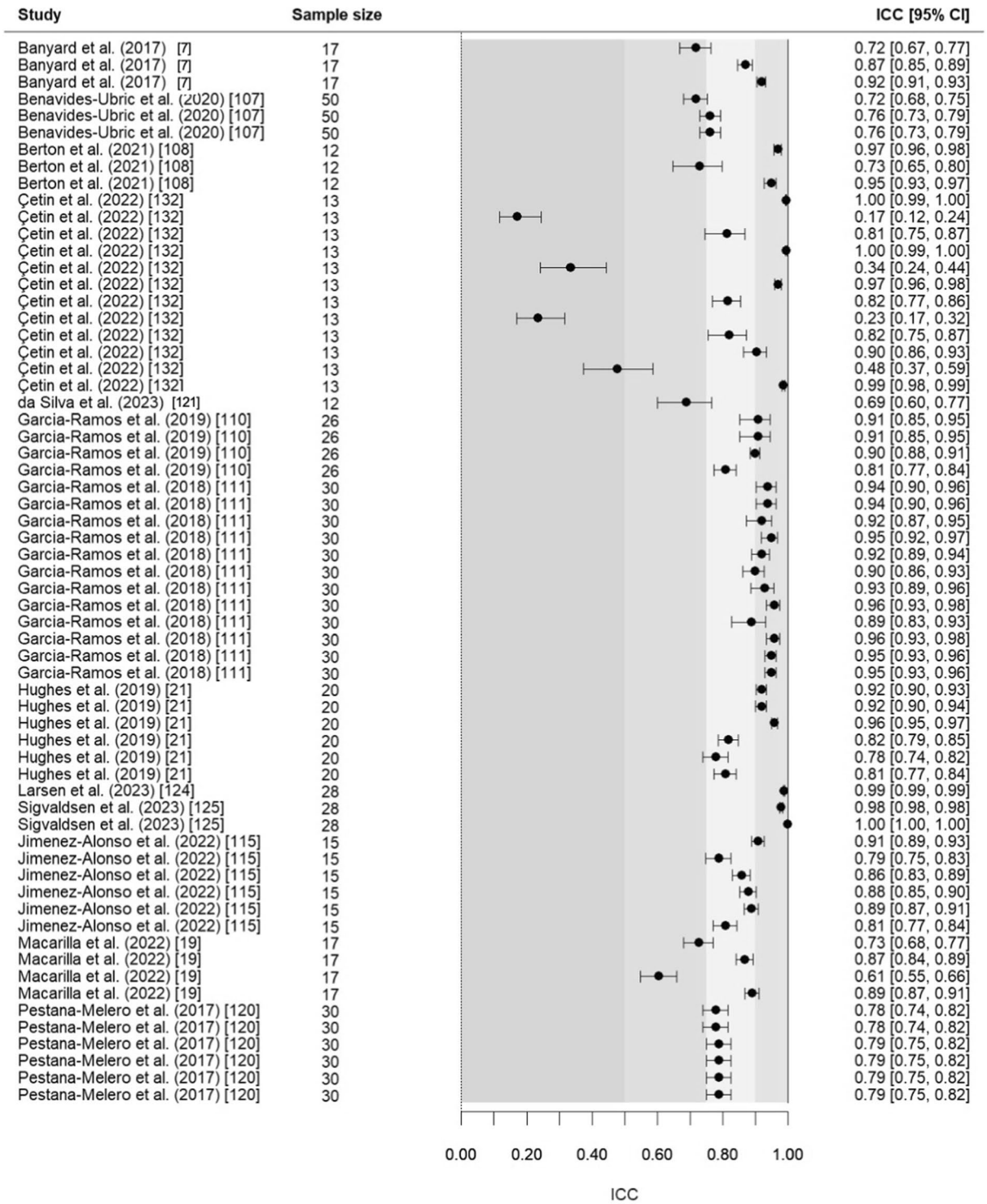

Fig. 4 Forest plot of interday reliability estimates for velocity-based 1RM prediction models. Each point represents an individual study-level ICC estimate, with horizontal lines indicating the 95% confidence interval. Values closer to 1.00 indicate higher relative reliability, whereas lower values indicate poorer test–retest stability. The pooled random-effects estimate summarizes the overall interday reliability across included effects but should be interpreted in light of the wide dispersion of individual study estimates and substantial heterogeneity. *ICC* intraclass correlation coefficient, *CI* confidence interval

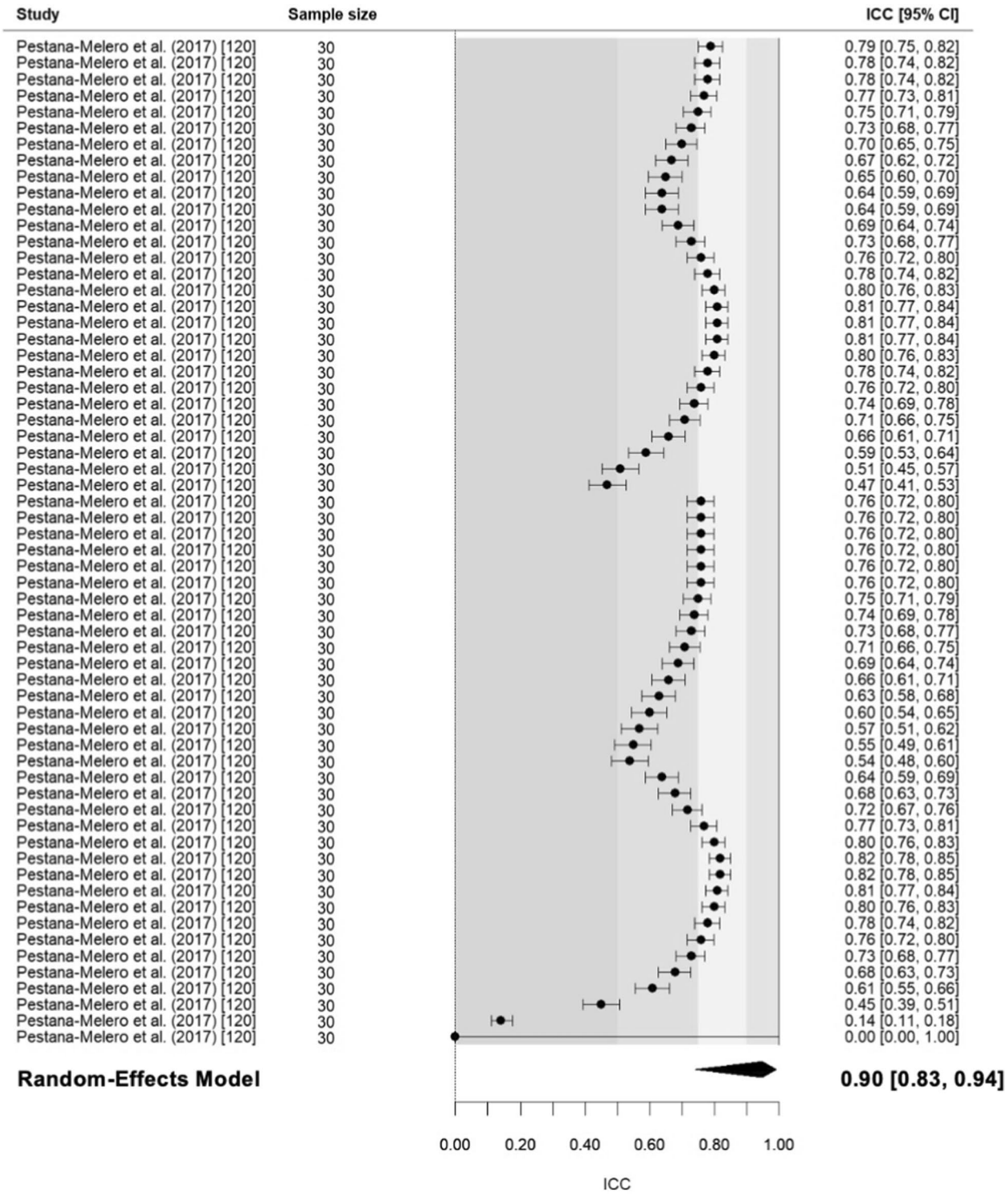

Fig. 4 continued

were conducted in 1 study [118] and exhibited stronger associations, including the lat pull ($\beta = 0.51$, $p = 0.003$, $n = 1$) and cable row ($\beta = 0.47$, $p = 0.005$, $n = 1$). Effect sizes for each exercise (if $n > 2$) are shown in Table 2 with significant β values in italics.

Additionally, the prediction method significantly moderated the validity estimates based on Pearson's

Table 2 Exercise-specific range of validity across 1RM prediction analyses with significant beta values in italics

Exercise	ICC	r
Squat	–	*0.37–0.99*
Deadlift	*0.97–0.98*	*0.60–0.98*
Bench press throw	0.68–0.97	–
Knee extension	–	0.84–0.96
Bench press	–	0.50–0.99

ICC intraclass correlation coefficient, *r* Pearson's correlation coefficient

r ($p<0.001$), with polynomial regression performing significantly worse than other methods ($\beta=-0.3941$; $p<0.001$, $n=5$) [11, 20, 111–113] and showing moderate-to-high correlations of $r=0.85$–0.98 [111] (one-point) and $r=0.78$ [112]–0.98 [113] (multipoint). Most studies ($n=21$) applied multipoint LV regression models, which yielded a wide range of correlations ($r=0.40$ [109]–0.99 [14, 113]). Two-point models ($n=13$) demonstrated moderate-to-excellent correlations ($r=0.50$–0.99 [113]), while single-point ($n=3$) approaches generally showed moderate-to-excellent correlations ($r=0.60$ [116]–0.95 [124]). Quadratic regression models ($r=0.98$–0.99 [8]) were included in fewer studies ($n=4$) but demonstrated consistently excellent validity.

Discussion

This systematic review synthesized evidence from 63 articles to meta-analyze the currently available validity and reliability of commercially available velocity-based measurement systems (Part I) and the accuracy of velocity-based 1RM prediction approaches applied in resistance training exercises (Part II).

In part I, 11 studies indicated high to excellent validity (ICCs = 0.91 [0.87–0.94], $r=0.95$ [0.82–0.98]. More studies ($n=27$) compared velocity tracking with other sensors, showing ICC = 0.92 [0.83–0.97], Lin's CCC = 0.90 [0.52–0.98]. Pearson's r was additionally reported ($r=0.91$ [0.84–0.95]) as a supplementary association metric, which does not assess agreement. Forty-seven articles ($N=47$) evaluated reliability facets, indicating good-to-excellent measurement consistency in 13 intraday reliability studies (ICC = 0.91 [0.87–0.95], $r=0.89$ [0.83–0.93] and 29 studies on interday reliability metrics with ICC = 0.90 [0.85–0.93], LIN's CCC = 0.99 [0.99–0.99] and Pearson's $r=0.91$ [0.38–0.99]. Most studies reported relative reliability, while only 12 studies quantified systematic and/or random measurement errors [23, 25, 46, 53, 62, 65, 66, 91, 93, 102, 103].

Part II comprised 38 studies evaluating the accuracy of the 1RM prediction via different methods (e.g., LV regression). Between-day reliability ($N=13$) showed good-to-excellent ICCs = 0.90 [0.83–0.94], with 8 studies providing detailed measurement error analyses [19, 107, 110, 115, 120, 124, 125, 132]. Validity was quantified via ICCs and Pearson's r, showing ICC = 0.91 [0.72–0.98] ($N=3$) and $r=0.96$ [0.94–0.97] ($N=31$), also focusing on relative reliabilities.

Validity of Velocity-Based Measurement Systems (Part I)

Despite high pooled validity estimates (ICC = 0.91, $r=0.95$), substantial heterogeneity was observed across studies. The wide range of correlation coefficients and analysis heterogeneity within IMU-based systems (ICC = 0.65–0.96, $r=0.40$–0.99) indicates that validity is strongly sensor-dependent rather than uniformly high across technologies. In contrast, LPT-based systems demonstrated more consistent validity estimates with lower variability (ICC = 0.87–0.98, $r=0.77$–0.99). It can be speculated whether the lower validity of IMU-based systems was attributable to indirect acceleration integration, which is more sensitive to signal noise, drift, and calibration assumptions. LPTs calculate speed directly from displacement over time, which likely contributes to their more consistent validity across different studies [24].

Relative reliability metrics should be interpreted cautiously in sensor-validation studies because they reflect rank-order consistency within a sample rather than measurement agreement with the reference device [137]. Validity assessment therefore requires quantifying differences between the new sensor and the reference standard using agreement analyses [138].

Measurement Error Analysis

Relative validity coefficients quantify the strength of association between devices rather than the magnitude of measurement error. Consequently, high ICC or Pearson's r values do not necessarily imply that two sensors can be used interchangeably in practice. For this purpose, agreement analyses and absolute measurement error metrics are required. Few authors included agreement analyses in their validity evaluations [22, 49, 51, 63, 70, 74, 84, 85, 96, 100, 103], and many of those who did provided insufficient or erroneous interpretations. Agreement between sensors should be interpreted by relating the device-to-device differences to the corresponding paired mean velocities [138]. Accordingly, a systematic bias between the gold standard measure and the new device can be interpreted as an over- or underestimation of velocity, that could be attributed to different underlying calculation models. Several studies provided limited agreement information. Two studies presented Bland–Altman plots only qualitatively without numerical agreement indices [51, 70], while Olaya-Cuartero et al. [100] interpreted the finding that "most of the data are between the upper and

lower LoA" as evidence of agreement. However, this is not an adequate interpretation of Bland–Altman analyses, as the LoA are constructed to include approximately 95% of observed differences [30]. Two studies [84, 96] reported numerical LoA incompletely, as these were not described in reference to the respective means, complicating the interpretation of the analysis.

Although several studies concluded that the sensors were valid, high correlation or ICC values alone do not demonstrate acceptable agreement. When interpreted relative to mean velocity, the reported LoA indicated potentially meaningful device-to-device differences, ranging from −0.32 to 0.41 m s^{-1} at 1.03 ± 0.10 m s^{-1} [84] and from 0.022 to 0.410 m s^{-1} at 0.510 ± 0.244 m s^{-1} [96]. Only Askow et al. [49] reported relative LoA (−11.6 to 7.1%) for GymAware versus Qualisys, allowing scale-independent interpretation [80]. LoA appeared narrower for LPTs than for IMUs, although this comparison remains descriptive because protocols and reference systems differed. Tendo showed LoA of −0.141 to 0.099 m·s^{-1} versus 3D motion capture [96] and −0.15 to −0.07 m·s^{-1} versus Optotrak [63], whereas Gheller et al. [103] reported mean bias only for EnodePro versus Vicon. The LoAs quantified a random scattering from unsystematic influences such as inter-individual standardization limitations (e.g., inappropriate sensor placement or unsystematic mathematical model problems). Appropriate reporting and interpretations of agreement analyses are urgently needed for a comprehensive practical recommendation.

Future validation studies should therefore complement relative validity analyses with standardized agreement and measurement error reporting to improve the practical interpretation of sensor performance.

Device Agreement Between Velocity Measurement Systems (Part I)

It is noteworthy that some authors considered LPT sensors such as GymAware and T-Force as the gold standard [52, 61, 74, 81, 94, 104]: *"LPT systems are considered the gold standard for measuring barbell velocity... [.]"* (p.6) [61]. Definitions and terms, as well as awareness of the difference between validity and device objectivity, were blurred and must be reconsidered. In the presented analysis, these studies were in the device agreement section, as they compared different velocity sensors against each other.

Twenty-three studies revealed good-to-excellent correlation metrics (r = 0.91 [0.84–0.95], ICCs = 0.92 [0.83–0.97]. Only 3 studies [64, 67, 80] evaluated agreement via CCC = 0.90 [0.52–0.98]. Increased device agreement variability (Q > 1050, p < 0.001) might be attributable to the lack of gold standard measurements, with both sensors showing individual and specific random and systematic errors. Sensor type, used exercise, intensity or velocity profile influenced results. Comparisons between LPT-based systems and LPT to T-force LVT generally demonstrated moderate-to-excellent agreement (ICC = 0.79–0.99) [92], with sensors such as Chronojump and Velowin showing significantly higher agreement with T-force [58]. IMU to LPT comparisons yielded r = −0.20–0.99 [24, 47, 57, 65, 74–76, 79, 92], ICC = 0.19–0.99 [24, 65, 74, 79, 80, 92], and CCC = 0.36–0.80 [80], while IMU to LVT showed correlations with r = 0.41–0.97 [53, 61, 84, 104], ICC = 0.00–0.98 [58, 64, 87, 94] and CCC = 0.61–0.99 [64, 67]. IMU-to-IMU comparisons yielded ICC = 0.30–0.56 [82], questioning the interchangeability of sensors. However, these findings should be regarded as exploratory rather than interpreted as evidence of superior measurement performance of these sensors, as the comparisons partly stem from a single study. In contrast, comparisons between EnodePro and Tendo reported lower CCCs [76], further highlighting that these effects likely reflect study-specific characteristics rather than generalizable differences in sensor quality. Analyses of squat exercises [24, 75, 76, 89] support this finding: whereas RepOne showed a relatively narrow ICC range (0.83–0.99), EnodePro exhibited a substantially wider range (0.28–0.99), reflecting pronounced inconsistency across measurement conditions. This could suggest that high ICC values observed for the squat may be driven more by sensor-specific characteristics than by uniformly high agreement across systems. In contrast, upper-body exercises such as the bench press demonstrate considerably high correlation values (r = 0.84–0.99), as reported for EnodePro [24] and RepOne [92], which introduces uncertainty regarding the sources of measurement error. Given the large number of potential influencing variables, it remains difficult to determine whether poorer device agreement is primarily attributable to sensor characteristics, exercise-specific factors, or their interaction. Exercise-specific effects are supported by Behrmann et al. [24], who reported a significantly smaller error range for bench press velocity measurements (MAPE = 3.29–8.79%) compared with squat measurements (MAPE = 8.28–29.90%).

Also, intensity or movement velocity moderated results, with agreement being significantly higher at moderate (50–80%1RM) and high intensities (> 80%) compared with low intensities (< 50%). Especially for IMU-based systems compared with T-force or LPTs, the results show a wide range with partly very low for low intensities (r = 0.19–0.99, ICC = 0.37–0.99) [24, 52, 57, 61, 71, 75, 76, 80, 82, 84]. Additionally, Behrmann et al. [24] reported LoA for bench press velocity at 70%1RM (−0.043 to 0.095 m s^{-1}, mean velocity 0.50 ± 0.14 m·s^{-1}),

whereas markedly wider LoA were observed at 30%1RM (−0.109 to 0.246 m s^{-1}, mean velocity 0.94 ± 0.19 m·s^{-1}). Similarly, Warneke et al. [80] reported MAPE between IMU and LPTs ranging from 9 to 11% at low squat intensities and 5.6–7.3% at moderate-to-high intensities.

Measurement Error Analysis

In contrast to correlation-based validity measures, agreement analyses directly quantify whether different devices provide interchangeable measurements by accounting for systematic bias and absolute differences between sensors. Consequently, agreement metrics such as Lin's CCC and Bland–Altman analyses are particularly relevant for practitioners intending to replace or combine measurement systems. However, comparatively few studies reported these analyses, limiting firm conclusions regarding device interchangeability. As for validity, some studies determined agreement via LoA or bias values, but without providing corresponding mean velocities or relative error metrics and without stratifying results by exercise or intensity [52, 58, 61, 65, 71, 72, 74, 76, 79, 81, 84, 87, 89, 92]. Most studies that used agreement analyses for reliability evaluations failed to report systematic bias and numerical LoA [57, 77, 82, 94] with one study even failing to report the mean velocity [74]. Only a few studies evaluated the systematic bias. González-Galán et al. [64] reported mean biases ranging from 2.8 to 13.4% when comparing Vitruve LPT with T-Force. Pérez-Castilla et al. [71] observed small absolute biases of 0.06–0.07 m s^{-1} across mean velocities of 0.48–1.28 m s^{-1}, while Martínez-Cava et al. [87] reported biases between −0.02 and 0.03 m s^{-1} across velocities of 0.7–1.15 m s^{-1}. In studies directly comparing LPT sensors, only two reported BA analyses with sufficient detail: Illera-Domínguez et al. [89] identified a substantial negative mean bias of −0.119 m s^{-1} at relatively low velocities (0.38–0.55 m·s^{-1}), whereas Lawson et al. [92] reported positive biases ranging from 0.039 to 0.110 m s^{-1} across a broader velocity range (0.42–1.52 m s^{-1}). An appropriate random error analysis was only found/rarely reported in 2 studies [80, 93].

Reliability (Part I)

Interday

Overall interday reliability revealed ICCs = 0.90 [0.85–0.93] (N = 29) and r = 0.91 [0.38–0.99] (N = 3) with a range between ICC = 0.14 and 0.99. This wide range of ICCs might be attributable to different methodological constrains. Koo and Li [42] recommended $n \geq 30$ as a rule of thumb to adequately interpret ICCs. While this recommendation is based on robustness of analyses of variance effects with sample sizes exceeding n = 30, larger sample sizes of $n > 150$ were required for stable ICC estimates [30]. Reliability analyses with $n < 50$ were considered "pilot data" [29]. Out of the included studies, only four studies met the criterion from Koo and Li [54, 62, 67, 97], and only one study reached a sample size of $n > 50$ [25], showing that instable reliability metrics could be attributable to comparatively small sample sizes.

A significant moderator for interday reliability was the sensor type ($p < 0.001$), with ICC = 0.44–0.99 for LPTs, ICC = 0.70–0.99 for T-force LVT and ICC = 0.14–0.99 for IMUs. Although the ICC-based meta-regression did not yield statistically significant β coefficients for individual sensors, several individual studies reported very low ICC values for specific IMU-based systems, including the Myotest (ICC = 0.14), Bar Sensei (ICC = 0.17), and the PUSH sensor (ICC = 0.38). The performed regression analysis showed EnodePro ($\beta = -1.77$) and GymAware ($\beta = -1.58$) to be sensors with the poorest reliability when Pearson 'r was calculated. Since regression analyses are impacted by sample size per potential moderator, the comparatively small study number might bias the regression analysis, and individual study results should be considered carefully.

Intraday

Intraday reliability reflects the consistency of velocity measurements obtained within a single testing session and therefore represents the lowest threshold for protocol reliability. Despite good-to-excellent reliability metrics (ICC = 0.91 [0.87–0.95], r = 0.89 [0.83–0.93]), values ranged widely from poor to excellent (ICC = 0.29–0.99) [23, 26, 53, 65, 66, 69, 83, 84, 91, 93, 101–103]. To explain the observed variability, meta-regression identified only sensor type as a significant moderator (p = 0.001), with the Beast IMU sensor ($\beta = -2.13$, p = 0.004) exhibiting significantly lower reliability, with ICC = 0.29–0.64 [23]. Studies employing LPTs (e.g., GymAware, Chronojump) [23, 65, 66, 69, 93, 101, 102] or T-force LVT [9, 23, 66, 84, 93, 102] demonstrated ICC = 0.61–0.99, indicating LPTs or LVTs to be more reliable than IMU sensors.

Measurement Error Analysis

Only eight studies reported CV (1.22–54.9%) [23, 53, 66, 91, 93, 102], TE (1.02–2.46) [103] or MAE (0.03–0.64) [93]. From a practical perspective, coaches and practitioners are often more interested in the expected measurement error than in relative reliability alone, particularly when monitoring small performance changes over time. Therefore, reliability studies should also complement ICC-based analyses with absolute reliability measures and standardized measurement error reporting to facilitate practical interpretation [139]. When interpreting reliability and repeatability quantifications, it should be noted that not sensor reliability, but rather

a combination of sensor reliability limitations and biological variability could cause outlined random errors. A pure sensor evaluation would exclude participants and use (for instance) a machine that moved the sensor with a prescribed velocity. Since all studies evaluated reliability in exercise performed by humans, repeatability errors were caused by the training individual and/or sensor limitations. Therefore, in contrast to systematic sensor over- or underestimations quantified in measurement analyses for validity or device agreement, in test–retest reliability, systematic errors could also be attributed to learning or fatiguing effects of the exercising individual. Reliability estimates from repeated exercise trials may reflect both device error and biological variability and should not be interpreted as device reliability alone [30]. Exercise-specific differences may also reflect habituation demands, particularly in complex multi-joint tasks and untrained participants [140, 141].

1RM Predictions (Part II)

Validity of 1RM Predictions

Based on 38 studies, velocity-based 1RM prediction showed high relative validity estimates with Pearson's $r = 0.96$ [0.94–0.97] ($N = 32$) [7, 9, 11, 14–16, 18, 19, 25, 107, 109, 111–120, 123–130, 134–136]. High heterogeneity [ICC: $Q(8) = 914.14$; *r*: $Q(197) = 1492.73$] could be explained by exercise ($p < 0.001$), variable ($p = 0.046$), and regression model ($p < 0.001$) being significant moderators. Squat and deadlift predictions demonstrated lower validity, with correlation coefficients ranging from $r = 0.37$–0.99 for squats [7, 8, 15, 18, 19, 109, 112, 127, 129] and $r = 0.60$–0.98 for deadlifts [11, 16, 20, 107, 114, 116, 126]. Besides this high variability, Warneke et al. [80] observed mean random over- and underestimations during free-weight squats with MAPE up to 71%, although relative quantifications (ICC/*r*) indicated good agreement. Additionally, the prediction method significantly moderated validity estimates, with polynomial regression models showing comparatively lower validity despite moderate-to-high correlations (one-point: $r = 0.85$–0.98 [111]; multipoint: $r = 0.78$–0.98 [112, 113]). LV regression models yielded a wide range of correlations ($r = 0.40$–0.99) [14, 109, 113], with both two-point and single-point models showing similarly broad validity estimates [112, 116, 124]. Although current practical recommendations suggest that two carefully selected loads may be sufficient [6], heterogeneous loading schemes and limited reporting preclude conclusions regarding the superiority of two-point over multipoint approaches. The apparently narrower validity range of single-point methods should likewise be interpreted cautiously because few studies used this approach. Increasing load points improved prediction accuracy in one study (three points: $r = 0.78$; five points: $r = 0.93$), [7] but this likely reflects broader sampling of the load–velocity spectrum rather than the number of loads alone. Accordingly, load location and distribution may be more important than the absolute number of tested loads [109, 142]. Recent evidence suggests that the location and distribution of tested loads are more important than the absolute number of loads.

Quadratic regression models ($r = 0.98$–0.99) [8] showed high validity but were evaluated in fewer studies. Prediction accuracy also appears to benefit from individualization: individualized regression equations ($r = 0.96$–0.98) outperformed generalized equations ($r = 0.88$–0.94) [15, 20, 107], and individualized or optimized minimum velocity thresholds may reduce bias and absolute prediction error, particularly in lower-body exercises [15, 109, 143]. This may partly explain the lower validity observed for squat and deadlift predictions. However, high group-level validity can obscure substantial individual error. For example, Ruf et al. [16] confirmed consistent group-level overestimations (~ 5–10 kg) in conventional deadlift, with high inter-individual variability (− 14–29 kg).

Reliability of 1RM Predictions

Reliability of 1 RM prediction was evaluated in 13 studies showing good-to-excellent ICC = 0.90 [0.83–0.94]. Despite a wide range of ICC = 0.14–0.99, the exercise type emerged as the only significant moderator ($p < 0.001$), with squat-based predictions demonstrating lower reliability ($\beta = -1.28$, $p < 0.001$) when the LV method was applied. Within the squat exercise, however, prediction stability strongly depended on the number of measured loads (Multipoint: ICC = 0.61–0.96 [7, 19, 21], Two-point method: ICC = 0.24–0.99 [132]. Importantly, even when reliability was reduced for squat-based predictions, Bazuelo-Ruiz et al. [122] and Perez-Castilla et al. [119, 135] reported absolute mean differences between predicted and measured 1RM values of < 5 kg when two-point and multiple-point methods were applied in the bench press. However, recent evidence indicates that back squat prediction accuracy can be substantially improved when optimized MVTs are used [109, 143], suggesting that lower reliability in squat-based predictions may partly reflect suboptimal threshold selection rather than exercise type alone. Nevertheless, biomechanical factors may also contribute. Lower-body movements involve greater displacement and more complex bar paths compared to upper-body exercises, which typically follow a more vertical and linear trajectory and may therefore reduce the predictability and stability of the LV relationship in complex lower-body tasks, resulting in a less predictable relationship [144].

Limitations

Despite the comprehensive scope of this meta-analysis, several limitations must be considered when interpreting these findings. First, substantial between-study heterogeneity was observed across most analyses, reflecting large methodological differences in study design, participant characteristics, exercise selection, sensor technology, and statistical reporting. While moderator analyses were conducted to explore these sources, several moderator levels were represented by only a small number of studies (e.g., Pearson-based sensor type regression). Consequently, some statistically significant moderator effects may reflect study-specific characteristics rather than robust, generalizable patterns. Second, the majority of included studies used small sample sizes ($n < 30$) [42] and male-dominated cohorts, which may reduce the generalizability of findings across populations, training levels, and sex. In addition, incomplete reporting of Bland–Altman analyses and error margins in several studies further constrained the evaluation of systematic and random measurement error. A further limitation is that different ICC formulations may have been pooled, as several studies did not report whether ICCs were based on consistency or absolute agreement models. Since these ICC types are not directly interchangeable, this may have contributed to the observed heterogeneity. These limitations, although originating from the primary studies, may have biased the analysis and affected the interpretation of the meta-analytical results, calling for further evaluation.

Conclusion

Overall, commercially available velocity sensors demonstrated good-to-excellent pooled relative validity and reliability, although these findings were mainly supported by LPT-based systems. CCC-based evidence for agreement was limited but generally favorable where available. However, substantial heterogeneity across studies highlights that measurement quality is strongly influenced by contextual factors such as exercise complexity, sensor technology, and movement intensity. LPT-based systems appeared more consistent than IMU-based systems, while prediction accuracy was generally more robust for upper-body exercises and less consistent for complex lower-body movements such as the squat and deadlift, although device-specific performance should be considered. Also, velocity-based 1RM prediction models showed high average validity and reliability, supporting their potential use as a practical alternative to maximal strength testing, but their practical use depends on acceptable absolute prediction errors. Although recent evidence suggests that optimized MVT selection can substantially improve prediction accuracy in these exercises, prediction accuracy varied considerably between exercises, with complex lower-body movements such as the squat and deadlift demonstrating reduced stability and validity. These findings suggest that velocity-based prediction should be interpreted cautiously and practitioners should preferentially rely on sensor- and exercise-specific evidence rather than pooled estimates alone when implementing velocity-based monitoring or 1RM prediction. However, some subgroup-specific recommendations remain preliminary because they are based on a limited number of studies. Consequently, conclusions regarding the practical precision of velocity-based monitoring and prediction remain limited.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40798-026-01102-0.

Supplementary material 1.

Author Contributions

NL, AB, AF, MS and NC conducted the literature search, study screening, and data extraction. AB and NL performed the methodological quality assessment, with NC resolving disagreements. SS conducted the statistical analyses. NC and KW drafted the manuscript. NC and KW conceived the study and supervised the project. All authors contributed to the interpretation of the data, critically revised the manuscript for important intellectual content, and approved the final version.

Funding

Open Access funding enabled and organized by Projekt DEAL. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The datasets and code used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics Approval and Consent to Participate

Ethical approval was not required for this study as it involved analysis of previously published data.

Competing Interests

The authors declare that they have no relevant financial or non-financial interests to disclose. Furthermore, the authors have no affiliations with or involvement in any organization or entity that could be perceived as influencing the work reported in this article.

Author details

[1]Institute of Human Movement Science and Exercise Physiology, University Osnabrück, Osnabrück, Germany. [2]Institute of Human Movement Science and Exercise Physiology, Friedrich Schiller University Jena, Jena, Germany. [3]Institute of Human Movement Science and Exercise Physiology, University of Hamburg, Hamburg, Germany. [4]Institute of Sustainability Psychology, Leuphana University Lüneburg, Lüneburg, Germany. [5]Experimental Orthopaedics, University Hospital Jena, Campus Eisenberg, Waldkliniken Eisenberg, Friedrich-Schiller-Universität Jena, Jena, Germany. [6]Orthopaedics, University Hospital Jena, Campus Eisenberg, Waldkliniken Eisenberg, Friedrich-Schiller-University Jena, Jena, Germany.

Received: 26 March 2026 Accepted: 20 August 2026
Published online: 03 September 2026

References

1. American College of Sports Medicine. American College of Sports Medicine position stand. Progression Models in Resistance Training for Healthy Adults. Med Sci Sports Exerc. 2009;41:687–708. https://doi.org/10.1249/MSS.0b013e3181915670.
2. Schoenfeld BJ, Grgic J, Ogborn D, Krieger JW. Strength and hypertrophy adaptations between low- vs. high-load resistance training: a systematic review and meta-analysis. J Strength Cond Res. 2017;31:3508–23. https://doi.org/10.1519/JSC.0000000000002200.
3. Greig L, Stephens Hemingway BH, Aspe RR, Cooper K, Comfort P, Swinton PA. Autoregulation in resistance training: addressing the inconsistencies. Sports Med. 2020;50:1873–87. https://doi.org/10.1007/s40279-020-01330-8.
4. Helms ER, Byrnes RK, Cooke DM, Haischer MH, Carzoli JP, Johnson TK, et al. RPE vs. percentage 1RM loading in periodized programs matched for sets and repetitions. Front Physiol. 2018;9:247. https://doi.org/10.3389/fphys.2018.00247.
5. González-Badillo JJ, Sánchez-Medina L. Movement velocity as a measure of loading intensity in resistance training. Int J Sports Med. 2010;31:347–52. https://doi.org/10.1055/s-0030-1248333.
6. Ramos AG. Resistance training intensity prescription methods based on lifting velocity monitoring. Int J Sports Med. 2024;45:257–66. https://doi.org/10.1055/a-2158-3848.
7. Banyard HG, Nosaka K, Haff GG. Reliability and validity of the load–velocity relationship to predict the 1RM back squat. J Strength Cond Res. 2017;31:1897–904. https://doi.org/10.1519/JSC.0000000000001657.
8. Thompson SW, Rogerson D, Ruddock A, Greig L, Dorrell HF, Barnes A, et al. A novel approach to 1RM prediction using the load-velocity profile: a comparison of models 2021. Sports (Basel). 2021;9:88. https://doi.org/10.3390/sports.
9. Balsalobre-Fernández C, Torres-Ronda L. The implementation of velocity-based training paradigm for team sports: framework, technologies, practical recommendations and challenges. Sports (Basel). 2021;9:47. https://doi.org/10.3390/sports9040047.
10. Held S, Speer K, Rappelt L, Wicker P, Donath L. The effectiveness of traditional vs. velocity-based strength training on explosive and maximal strength performance: a network meta-analysis. Front Physiol. 2022. https://doi.org/10.3389/fphys.2022.926972.
11. Lopes dos Santos M, Mann JB, Berton R, Alvar BA, Lockie RG, Dawes JJ. Using the load-velocity profile for predicting the 1RM of the hexagonal barbell deadlift exercise. J Strength Cond Res. 2023;37:220–3. https://doi.org/10.1519/JSC.0000000000004216.
12. Marston KJ, Forrest MRL, Teo SYM, Mansfield SK, Peiffer JJ, Scott BR. Load-velocity relationships and predicted maximal strength: a systematic review of the validity and reliability of current methods. PLoS ONE. 2022;17:e0267937. https://doi.org/10.1371/journal.pone.0267937.
13. Pareja-Blanco F, Loturco I. Velocity-based training for monitoring training load and assessing training effects. In: Muñoz-López A, Taiar R, Sañudo B, editors. Resistance training methods: from theory to practice. Cham: Springer; 2022. p. 153–79. https://doi.org/10.1007/978-3-030-81989-7_9.
14. Balsalobre-Fernández C, Kipp K. Use of machine-learning and load-velocity profiling to estimate 1-repetition maximums for two variations of the bench-press exercise. Sports (Basel). 2021. https://doi.org/10.3390/sports9030039.
15. Caven EJG, Bryan TJE, Dingley AF, Drury B, Garcia-Ramos A, Perez-Castilla A, et al. Group versus individualized minimum velocity thresholds in the prediction of maximal strength in trained female athletes. Int J Environ Res Public Health. 2020;17:7811. https://doi.org/10.3390/ijerph17217811.
16. Ruf L, Chéry C, Taylor K-L. Validity and reliability of the load-velocity relationship to predict the one-repetition maximum in deadlift. J Strength Cond Res. 2018;32:681–9. https://doi.org/10.1519/JSC.0000000000002369.
17. Guppy SN, Kendall KL, Haff GG. Velocity-based training—a critical review. Strength Cond J. 2024;46:295–307.
18. Fernandes JFT, Dingley AF, Garcia-Ramos A, Perez-Castilla A, Tufano JJ, Twist C. Prediction of one repetition maximum using reference minimum velocity threshold values in young and middle-aged resistance-trained males. Behav Sci. 2021;11:71. https://doi.org/10.3390/bs11050071.
19. Macarilla C, Sautter N, Robinson Z, Juber M, Hickmott L, Cerminaro R, et al. Accuracy of predicting one-repetition maximum from submaximal velocity in the barbell back squat and bench press. J Hum Kinet. 2022;82:201–12. https://doi.org/10.2478/hukin-2022-0046.
20. Jukic I, García-Ramos A, Malecek J, Omcirk D, Tufano JJ. Validity of load-velocity relationship to predict 1 repetition maximum during deadlifts performed with and without lifting straps: the accuracy of six prediction models. J Strength Cond Res. 2022;36:902–10. https://doi.org/10.1519/JSC.0000000000003596.
21. Hughes LJ, Banyard HG, Dempsey AR, Scott BR. Using a load-velocity relationship to predict one repetition maximum in free-weight exercise: a comparison of the different methods. J Strength Cond Res. 2019;33:2409–19. https://doi.org/10.1519/JSC.0000000000002550.
22. Lorenzetti S, Lamparter T, Lüthy F. Validity and reliability of simple measurement device to assess the velocity of the barbell during squats. BMC Res Notes. 2017;10:707. https://doi.org/10.1186/s13104-017-3012-z.
23. Pérez-Castilla A, Piepoli A, Delgado-García G, Garrido-Blanca G, García-Ramos A. Reliability and concurrent validity of seven commercially available devices for the assessment of movement velocity at different intensities during the bench press. J Strength Cond Res. 2019;33:1258–65. https://doi.org/10.1519/jsc.0000000000003118.
24. Behrmann N, Hillebrecht M, Afonso J, Warneke K. Is the EnodePro® a valid tool to determine the bar velocity in the bench press and barbell back squat? A comparative analysis. Sensors. 2025;25:549. https://doi.org/10.3390/s25020549.
25. Jukic I, King A, Sousa CA, Prnjak K, McGuigan MR. Implementing a velocity-based approach to resistance training: the reproducibility and sensitivity of different velocity monitoring technologies. Sci Rep. 2023;13:7152. https://doi.org/10.1038/s41598-023-34416-0.
26. Lake J, Augustus S, Austin K, Comfort P, McMahon J, Mundy P, et al. The reliability and validity of the bar-mounted PUSH Band ™ 2.0 during bench press with moderate and heavy loads. J Sports Sci. 2019;37:2685–90. https://doi.org/10.1080/02640414.2019.1656703.
27. García-Ramos A, Haff GG, Padial P, Feriche B. Reliability of power and velocity variables collected during the traditional and ballistic bench press exercise. Sports Biomech. 2018;17:117–30. https://doi.org/10.1080/14763141.2016.1260767.
28. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998;26:217–38.
29. Hopkins WG. Measures of reliability in sports medicine and science. Sports Med. 2000;30:1–15.
30. Warneke K, Gronwald T, Wallot S, Magno A, Hillebrecht M, Wirth K. Discussion on the validity of commonly used reliability indices in sports medicine and exercise science: a critical review with data simulations. Eur J Appl Physiol. 2025;125:1511–26. https://doi.org/10.1007/s00421-025-05720-6.
31. Weakley J, Morrison M, García-Ramos A, Johnston R, James L, Cole MH. The validity and reliability of commercially available resistance training monitoring devices: a systematic review. Sports Med. 2021;51:443–502. https://doi.org/10.1007/s40279-020-01382-w.
32. Page MJ, Mckenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2020;372:n71. https://doi.org/10.1136/bmj.n71.
33. Wager E, Kleinert S. MASA responsible research publication: international standards for authors. Contribution. Sec Med Sci. 2014. https://doi.org/10.1515/prilozi-2015-000.
34. Eriksen MB, Frandsen TF. The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: a systematic review. J Med Libr Assoc. 2018. https://doi.org/10.5195/jmla.2018.345.
35. Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN risk of bias checklist for systematic reviews of

patient-reported outcome measures. Qual Life Res. 2018;27:1171–9. https://doi.org/10.1007/s11136-017-1765-4.
36. Jeffries AC, Wallace L, Coutts AJ, McLaren SJ, McCall A, Impellizzeri FM. Athlete-reported outcome measures for monitoring training responses: a systematic review of risk of bias and measurement property quality according to the COSMIN guidelines. Int J Sports Physiol Perform. 2020;15:1203–15. https://doi.org/10.1123/ijspp.2020-0386.
37. Fisher Z, Tipton E. robumeta: an R-package for robust variance estimation in meta-analysis. arXiv. 2015. arXiv:1503.02220. https://doi.org/10.48550/arXiv.1503.02220.
38. Noble C, Sala G, Peter M, Lingwood J, Rowland CF, Gobet F, et al. The impact of shared book reading on children's language skills: a meta-analysis. Educ Res Rev. 2019;28:100290. https://doi.org/10.1016/j.edurev.2019.100290.
39. Donner A. A review of inference procedures for the intraclass correlation coefficient in the one-way random effects model. Int Stat Rev. 1986;54:67. https://doi.org/10.2307/1403259.
40. Noble S, Scheinost D, Constable RT. A decade of test-retest reliability of functional connectivity: a systematic review and meta-analysis. Neuroimage. 2019;203:116157. https://doi.org/10.1016/j.neuroimage.2019.116157.
41. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to meta-analysis. Chichester: Wiley; 2009. https://doi.org/10.1002/9780470743386.
42. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15:155–63. https://doi.org/10.1016/j.jcm.2016.02.012.
43. Cochran WG. The combination of estimates from different experiments. Biometrics. 1954;10:101. https://doi.org/10.2307/3001666.
44. Mavridis D, Salanti G. How to assess publication bias: funnel plot, trim-and-fill method and selection models. Evid Based Ment Health. 2014;17:30–30. https://doi.org/10.1136/eb-2013-101699.
45. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315:629–34. https://doi.org/10.1136/bmj.315.7109.629.
46. Orange ST, Metcalfe JW, Marshall P, Vince RV, Madden LA, Liefeith A. Test-retest reliability of a commercial linear position transducer (GymAware PowerTool) to measure velocity and power in the back squat and bench press. J Strength Cond Res. 2020;34:728–37. https://doi.org/10.1519/JSC.0000000000002715.
47. Jovanovic M, Jukic I. Within-unit reliability and between-units agreement of the commercially available linear position transducer and barbell-mounted inertial sensor to measure movement velocity. J Strength Cond Res. 2020. https://doi.org/10.1519/jsc.0000000000003776.
48. Weakley J, Mann B, Banyard H, McLaren S, Scott T, Garcia-Ramos A. Velocity-based training: from theory to application. Strength Cond J. 2021;43:31–49. https://doi.org/10.1519/SSC.0000000000000560.
49. Askow AT, Stone JD, Arndts DJ, King AC, Goto S, Hannon JP, et al. Validity and reliability of a commercially-available velocity and power testing device. Sports (Basel). 2018;6:170. https://doi.org/10.3390/sports6040170.
50. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al editors. Cochrane handbook for systematic reviews of interventions. Chichester: Wiley; 2019. https://doi.org/10.1002/9781119536604.
51. Abbott JC, Wagle JP, Sato K, Painter K, Light TJ, Stone MH. Validation of inertial sensor to measure barbell kinematics across a spectrum of loading conditions. Sports (Basel). 2020. https://doi.org/10.3390/sports8070093.
52. Arede J, Figueira B, Gonzalo-Skok O, Leite N. Validity and reliability of Gyko Sport for the measurement of barbell velocity on the bench-press exercise. J Sports Med Phys Fit. 2019. https://doi.org/10.23736/S0022-4707.19.09770-6.
53. Balsalobre-Fernández C, Kuzdub M, Poveda-Ortiz P, Campo-Vecino JD. Validity and reliability of the PUSH wearable device to measure movement velocity during the back squat exercise. J Strength Cond Res. 2016;30:1968–74. https://doi.org/10.1519/jsc.0000000000001284.
54. Bampouras T, Relph N, Orme D, Esformes J. Validity and reliability of the Myotest Pro wireless accelerometer in squat jumps. Isokinet Exerc Sci. 2013;21:101. https://doi.org/10.3233/IES-130484.
55. Boehringer S, Whyte DG. Validity and test-retest reliability of the 1080 quantum system for bench press exercise. J Strength Cond Res. 2019;33:3242–51. https://doi.org/10.1519/JSC.0000000000003184.
56. Caruso JF, Olson NM, Taylor ST, McLagan JR, Shepherd CM, Borgsmiller JA, et al. Front squat data reproducibility collected with a triple-axis accelerometer. J Strength Cond Res. 2012;26:40–6. https://doi.org/10.1519/JSC.0b013e31821d5ed7.
57. Chéry C, Ruf L. Reliability of the load-velocity relationship and validity of the PUSH to measure velocity in the deadlift. J Strength Cond Res. 2019;33:2370–80. https://doi.org/10.1519/JSC.0000000000002663.
58. Courel-Ibáñez J, Martínez-Cava A, Morán-Navarro R, Escribano-Peñas P, Chavarren-Cabrero J, González-Badillo JJ, et al. Reproducibility and repeatability of five different technologies for bar velocity measurement in resistance training. Ann Biomed Eng. 2019;47:1523–38. https://doi.org/10.1007/s10439-019-02265-6.
59. Elsworthy N, Callaghan DE, Scanlan AT, Kertesz AHM, Kean CO, Dascombe BJ, et al. Validity and reliability of using load-velocity relationship profiles to establish back squat 1 m·s-1 load. J Strength Cond Res. 2021;35:340–6. https://doi.org/10.1519/jsc.0000000000003871.
60. Ferro A, Floría P, Villacieros J, Muñoz-López A. Maximum velocity during loaded countermovement jumps obtained with an accelerometer, linear encoder and force platform: a comparison of technologies. J Biomech. 2019;95:109281. https://doi.org/10.1016/j.jbiomech.2019.07.025.
61. García-Pinillos F, Latorre-Román PA, Valdivieso-Ruano F, Balsalobre-Fernández C, Párraga-Montilla JA. Validity and reliability of the WIMU ® system to measure barbell velocity during the half-squat exercise. Proc Inst Mech Eng P J Sport Eng Technol. 2019;233:408–15. https://doi.org/10.1177/1754337119831890.
62. Garnacho-Castaño MV, López-Lastra S, Maté-Muñoz JL. Reliability and validity assessment of a linear position transducer. J Sports Sci Med. 2015;14:128–36.
63. Goldsmith JA, Trepeck C, Halle JL, Mendez KM, Klemp A, Cooke DM, et al. Validity of the open barbell and tendo weightlifting analyzer systems versus the optotrak certus 3D motion-capture system for barbell velocity. Int J Sports Physiol Perform. 2019;14:540–3. https://doi.org/10.1123/ijspp.2018-0684.
64. González-Galán J, Herrera-Bermudo JC, González-Badillo JJ, Rodríguez-Rosell D. Validity and concordance of a linear position transducer (Vitruve) for measuring movement velocity during resistance training. Sensors (Basel). 2024;24:6444. https://doi.org/10.3390/s24196444.
65. Held S, Rappelt L, Deutsch JP, Donath L. Valid and reliable barbell velocity estimation using an inertial measurement unit. Int J Environ Res Public Health. 2021. https://doi.org/10.3390/ijerph18179170.
66. Lopez-Torres O, Fernandez-Elias VE, Li J, Gomez-Ruano MA, Guadalupe-Grau A. Validity and reliability of a new low-cost linear position transducer to measure mean propulsive velocity: the ADR device. Proc Inst Mech Eng P J Sport Eng Technol. 2025;239:126–34. https://doi.org/10.1177/17543371221104345.
67. Maroto-Izquierdo S, Nosaka K, Alarcón-Gómez J, Martín-Rivera F. Validity and reliability of inertial measurement system for linear movement velocity in flywheel squat. Sensors (Basel). 2023;23:2193. https://doi.org/10.3390/s23042193.
68. Martín-San Agustín R, Sánchez-Barbadora M, García-Vidal JA. Validity of an inertial system for measuring velocity, force, and power during hamstring exercises performed on a flywheel resistance training device. PeerJ. 2020;8:e10169. https://doi.org/10.7717/peerj.10169.
69. Oleksy Ł, Kuchciak M, Bril G, Mika A, Przydział M, Pazdan-Śliż I, et al. Intra-rater and test-retest reliability of barbell force, velocity, and power during the landmine punch throw test assessed by the GymAware linear transducer. Appl Sci (Basel). 2023;13:10875. https://doi.org/10.3390/app131910875.
70. Pelka EZ, Gadola C, McLaughlin D, Slattery E, Claytor RP. Comparison of the PUSH Band 2.0 and vicon motion capture to measure concentric movement velocity during the barbell back squat and bench press. Sports. 2022;11:6. https://doi.org/10.3390/sports11010006.
71. Pérez-Castilla A, Boullosa D, García-Ramos A. Reliability and validity of the iLOAD application for monitoring the mean set velocity during the back squat and bench press exercises performed against different loads. J Strength Cond Res. 2021;35:S57-65. https://doi.org/10.1519/JSC.0000000000003739.

72. Pérez-Castilla A, Jiménez-Reyes P, Haff GG, García-Ramos A. Assessment of the loaded squat jump and countermovement jump exercises with a linear velocity transducer: which velocity variable provides the highest reliability? Sports Biomech. 2021;20:247–60. https://doi.org/10.1080/14763141.2018.1540651.
73. Pueo B, Lopez JJ, Mossi JM, Colomer A, Jimenez-Olmedo JM. Video-based system for automatic measurement of barbell velocity in back squat. Sensors (Basel). 2021;21:925. https://doi.org/10.3390/s21030925.
74. Zhang S, Yang C, Zhu J, Tian J, Fu M, Chen C, et al. A method for obtaining barbell velocity and displacement and motion counting based on IMU. Mob Netw Appl. 2024;29:623–33. https://doi.org/10.1007/s11036-024-02293-0.
75. Qu H, Qian D, Xu S, Shen Y. Validity and test-retest reliability of a resistance training device for Smith machine back squat exercise. iScience. 2024;27:108582. https://doi.org/10.1016/j.isci.2023.108582.
76. Ruiz-Alias SA, Şentürk D, Akyildiz Z, Cetin O, Kaya S, Pérez-Castilla A, et al. Validity and reliability of velocity and power measures provided by the Vitruve linear position transducer. PLoS ONE. 2024;19(10):e0312348. https://doi.org/10.1371/journal.pone.0312348.
77. Sato K, Smith SL, Sands WA. Validation of an accelerometer for measuring sport performance. J Strength Cond Res. 2009;23:341–7. https://doi.org/10.1519/JSC.0b013e3181876a01.
78. Suchomel TJ, Techmanski BS, Kissick CR, Comfort P. Reliability, validity, and comparison of barbell velocity measurement devices during the jump shrug and hang high pull. J Funct Morphol Kinesiol. 2023. https://doi.org/10.3390/jfmk8010035.
79. Van Den Tillaar R, Ball N. Validity and reliability of kinematics measured with PUSH band vs. linear encoder in bench press and push-ups. Sports. 2019;7:207. https://doi.org/10.3390/sports7090207.
80. Warneke K, Skratek J, Wagner C-M, Wirth K, Keiner M. Random measurement and prediction errors limit the practical relevance of two velocity sensors to estimate the 1RM back squat. Front Physiol. 2024. https://doi.org/10.3389/fphys.2024.1435103.
81. Justo-Álvarez A, García-López J, Sabido R, García-Valverde A. Validity of a new portable sensor to measure velocity-based resistance training. Methods Protoc. 2025;8:9. https://doi.org/10.3390/mps8010009.
82. Beckham GK, Layne DK, Kim SB, Martin EA, Perez BG, Adams KJ. Reliability and criterion validity of the Assess2Perform Bar Sensei. Sports. 2019;7:230. https://doi.org/10.3390/sports7110230.
83. Chavda S, Sandau I, Bishop C, Xu J, Turner AN, Lake JP. Validity and reliability of a commercially available inertial sensor for measuring barbell mechanics during weightlifting. Appl Sci. 2024;14:7397. https://doi.org/10.3390/app14167397.
84. Feuerbacher JF, Jacobs MW, Dragutinovic B, Goldmann J-P, Cheng S, Schumann M. Validity and test-retest reliability of the Vmaxpro sensor for evaluation of movement velocity in the deep squat. J Strength Cond Res. 2023;37:35–40. https://doi.org/10.1519/JSC.0000000000004207.
85. Fritschi R, Seiler J, Gross M. Validity and effects of placement of velocity-based training devices. Sports. 2021;9:123. https://doi.org/10.3390/sports9090123.
86. Gant RW, Pinzone AG, Tagesen EC, Rivera J, Davis BR, Pelka EZ, et al. Validity of linear position transducer to measure velocity during barbell back squats. Med Sci Sports Exerc. 2022;54:97–8. https://doi.org/10.1249/01.mss.0000876264.53146.f5.
87. Martínez-Cava A, Hernández-Belmonte A, Courel-Ibáñez J, Morán-Navarro R, José González-Badillo JJ, Pallarés JG. Reliability of technologies to measure the barbell velocity: Implications for monitoring resistance training. PLoS ONE. 2020;15:e0232465. https://doi.org/10.1371/journal.pone.0232465.
88. Grgic J, Scapec B, Pedisic Z, Mikulic P. Test-retest reliability of velocity and power in the deadlift and squat exercises assessed by the GymAware PowerTool system. Front Physiol. 2020. https://doi.org/10.3389/fphys.2020.561682.
89. Illera-Domínguez V, Fernández-Valdés B, Gisbert-Orozco J, Ramirez-Lopez C, Nuell S, González J, et al. Validity of a low-cost friction encoder for measuring velocity, force and power in flywheel exercise devices. Biol Sport. 2023;40:805–11. https://doi.org/10.5114/biolsport.2023.119991.
90. Janicijevic D, García-Ramos A, Lamas-Cepero JL, García-Pinillos F, Marcos-Blanco A, Rojas FJ, et al. Comparison of the two most commonly used gold-standard velocity monitoring devices (GymAware and T-Force) to assess lifting velocity during the free-weight barbell back squat exercise. Proc Inst Mech Eng P J Sport Eng Technol. 2023;237:205–12. https://doi.org/10.1177/17543371211029614.
91. Lake JP, Augustus S, Austin K, Mundy P, Mcmahon JJ, Comfort P, et al. The Validity of the Push Band 2.0 during Vertical Jump Performance. Sports. 2018;6(4):140 https://doi.org/10.3390/sports6040140.
92. Lawson DJ, Olmos AA, Stahl CA, Lopes Dos Santos M, Goodin JR, Dawes JJ. Validity, reliability, and sensitivity of a commercially available barbell velocity measuring device when performing selected exercises. Int J Exerc Sci. 2024. https://doi.org/10.70252/NVPI2815.
93. Wagner CM, Keiner M, Puschkasch-Möck S, Wirth K, Schiemann S, Warneke K. Can the velocity profile in the bench press and the bench pull sufficiently estimate the one repetition maximum in youth elite cross-country ski and biathlon athletes? BMC Sports Sci Med Rehabil. 2025;17:102. https://doi.org/10.1186/s13102-025-01137-y.
94. van den Tillaar R, Falch HN, Larsen S. Comparison of velocity and estimated one repetition maximum measured with different measuring tools in bench presses and squats. Sensors (Basel). 2024;24:7422. https://doi.org/10.3390/s24237422.
95. Stock MS, Beck TW, DeFreitas JM, Dillon MA. Test–retest reliability of barbell velocity during the free-weight bench-press exercise. J Strength Cond Res. 2011;25:171–7. https://doi.org/10.1519/JSC.0b013e318201bdf9.
96. McGrath GA, Flanagan EP, O'Donovan P, Collins DJ, Kenny IC. Velocity based training: validity of monitoring devices to assess mean concentric velocity in the bench press exercise. J Aust Strength Cond. 2018;26:23–30.
97. Miller RM, Freitas EDS, Heishman AD, Koziol KJ, Galletti BAR, Kaur J, et al. Test-retest reliability between free weight and machine-based movement velocities. J Strength Cond Res. 2020;34:440–4. https://doi.org/10.1519/JSC.0000000000002817.
98. Moreno-Villanueva A, Rico-González M, Pérez-Caballero CE, Rodríguez-Valero G, Pino-Ortega J. Level of agreement and reliability of ADR encoder to monitor mean propulsive velocity during the bench press exercise. Proc Inst Mech Eng P J Sport Eng Technol. 2024;238:264–75. https://doi.org/10.1177/17543371221100395.
99. Muyor JM, Granero-Gil P, Pino-Ortega J. Reliability and validity of a new accelerometer (Wimu®) system for measuring velocity during resistance exercises. Proc Inst Mech Eng P J Sport Eng Technol. 2018;232:218–24. https://doi.org/10.1177/1754337117731700.
100. Olaya-Cuartero J, Villalon-Gasch L, Penichet-Tomás A, Jimenez-Olmedo JM. Validity and Reliability of the VmaxPro IMU for back squat exercise in multipower machine. J Phys Educ Sport. 2022;22:2920–6. https://doi.org/10.7752/jpes.2022.11369.
101. Pinzone AG, Gant RW, Rivera J, Pelka EZ, Tagesen EC, Lebron MA, et al. Validity and reliability of a linear position transducer to measure barbell velocity, duration, and displacement during the bench press. Int J Exerc Sci. 2024. https://doi.org/10.70252/MFJK8861.
102. Sousa T, Pereira A, Luís Marques D, Pereira Neiva H, Almeida Marinho D, Cardoso Marques M. Validity and reliability of the T-Force and Chronojump systems to measure movement velocity on resistance machines with older adults. Sci Rep. 2024;14:23727. https://doi.org/10.1038/s41598-024-74159-0.
103. Gheller RG, Kons RL, Külkamp W, Dal Pupo J, Detanico D. Validity and reliability of *ergonauta* encoder to assess countermovement jump performance. Proc Inst Mech Eng P J Sport Eng Technol. 2025;239:604–10. https://doi.org/10.1177/17543371231171192.
104. Pérez-Castilla A, García-Ramos A, Gijón-Nieto LM, Marcos-Blanco A, García-Pinillos F. Reliability and concurrent validity of the PUSH BandTM 2.0 to measure barbell velocity during the free-weight and Smith machine squat exercises. Proc Inst Mech Eng P J Sport Eng Technol. 2022;236:351–9. https://doi.org/10.1177/17543371211024018.
105. Gantois P, Gilo D, Denise A, Muñoz-López A, Nakamura FY, de Souza FF. Load-velocity relationship to estimate the one-repetition maximum in the hexagonal bar deadlift exercise in women. Women Sport Phys Act J. 2022;30:144–50. https://doi.org/10.1123/wspaj.2022-0015.
106. Balsalobre-Fernández C, Cardiel-García M, Jiménez SL. Bilateral and unilateral load-velocity profiling in a machine-based, single-joint, lower body exercise. PLoS ONE. 2019;14:e0222632. https://doi.org/10.1371/journal.pone.0222632.

107. Benavides-Ubric A, Díez-Fernández DM, Rodríguez-Pérez MA, Ortega-Becerra M, Pareja-Blanco F. Analysis of the load-velocity relationship in deadlift exercise. J Sports Sci Med. 2020;19:452–9.
108. Berton R, Soriano M, da Silva DD, dos Santos ML, Teixeira G, Reiser F, et al. Concurrent validity and reliability of the load-velocity relationship to predict the one-repetition maximum during three weightlifting derivatives. Kinesiology. 2021;53:215–25. https://doi.org/10.26582/k.53.2.4.
109. Fitas A, Gomes M, Santos P, Gonçalves AD, Pezarat-Correia P, Mendonca GV. Optimal minimum-velocity threshold to predict one-repetition maximum for the back squat. Int J Sports Med. 2024;45:923–9. https://doi.org/10.1055/a-2335-4143.
110. García-Ramos A, Barboza-González P, Ulloa-Díaz D, Rodriguez-Perea A, Martinez-Garcia D, Guede-Rojas F, et al. Reliability and validity of different methods of estimating the one-repetition maximum during the free-weight prone bench pull exercise. J Sports Sci. 2019;37:2205–12. https://doi.org/10.1080/02640414.2019.1626071.
111. García-Ramos A, Haff GG, Pestaña-Melero FL, Pérez-Castilla A, Rojas FJ, Balsalobre-Fernández C, et al. Feasibility of the 2-point method for determining the 1-repetition maximum in the bench press exercise. Int J Sports Physiol Perform. 2018;13:474–81. https://doi.org/10.1123/ijspp.2017-0374.
112. Kilgallon J, Cushion E, Joffe S, Tallent J. Reliability and validity of velocity measures and regression methods to predict maximal strength ability in the back-squat using a novel linear position transducer. Proc Inst Mech Eng P J Sport Eng Technol. 2025;239:335–48. https://doi.org/10.1177/17543371221093189.
113. Janicijevic D, Jukic I, Weakley J, García-Ramos A. Bench press 1-repetition maximum estimation through the individualized load–velocity relationship: comparison of different regression models and minimal velocity thresholds. Int J Sports Physiol Perform. 2021;16:1074–81. https://doi.org/10.1123/ijspp.2020-0312.
114. Janicijevic D, Senturk D, Akyildiz Z, Gu Y, García-Ramos A. Hexagonal barbell deadlift one-repetition maximum estimation using velocity recordings. Int J Sports Med. 2024;45:844–50. https://doi.org/10.1055/a-2330-3001.
115. Jiménez-Alonso A, García-Ramos A, Cepero M, Miras-Moreno S, Rojas FJ, Pérez-Castilla A. Velocity performance feedback during the free-weight bench press testing procedure: an effective strategy to increase the reliability and one repetition maximum accuracy prediction. J Strength Cond Res. 2022;36:1077–83. https://doi.org/10.1519/JSC.0000000000003609.
116. Lake J, Naworynsky D, Duncan F, Jackson M. Comparison of different minimal velocity thresholds to establish deadlift one repetition maximum. Sports. 2017. https://doi.org/10.3390/sports5030070.
117. Pérez-Castilla A, Fernandes JFT, García-Ramos A. Validity of the bench press one-repetition maximum test predicted through individualized load-velocity relationship using different repetition criteria and minimal velocity thresholds. Isokinet Exerc Sci. 2021;29:369–77. https://doi.org/10.3233/IES-202247.
118. Pérez-Castilla A, Suzovic D, Domanovic A, Fernandes JFT, García-Ramos A. Validity of Different Velocity-Based Methods and Repetitions-to-Failure Equations for Predicting the 1 Repetition Maximum During 2 Upper-Body Pulling Exercises. J Strength Cond Res. 2021;35:1800–8. https://doi.org/10.1519/JSC.0000000000003076.
119. Pérez-Castilla A, Piepoli A, Garrido-Blanca G, Delgado-García G, Balsalobre-Fernández C, García-Ramos A. Precision of 7 commercially available devices for predicting bench-press 1-repetition maximum from the individual load–velocity relationship. Int J Sports Physiol Perform. 2019;14:1442–6. https://doi.org/10.1123/ijspp.2018-0801.
120. Pestaña-Melero FL, Haff GG, Rojas FJ, Pérez-Castilla A, García-Ramos A. Reliability of the load–velocity relationship obtained through linear and polynomial regression models to predict the 1-repetition maximum load. J Appl Biomech. 2018;34:184–90. https://doi.org/10.1123/jab.2017-0266.
121. Berton R, Soares Teixeira G, Barros MH, Reiser F, Tricoli V, da Silva DD. Using the bar-velocity to predict the one-repetition maximum in the power clean from the knee. Kinesiology. 2023;55:254–61. https://doi.org/10.26582/k.55.2.9.
122. Bazuelo-Ruiz B, Padial P, García-Ramos A, Morales-Artacho AJ, Miranda MT, Feriche B. Predicting maximal dynamic strength from the load-velocity relationship in squat exercise. J Strength Cond Res. 2015;29:1999–2005. https://doi.org/10.1519/JSC.0000000000000821.
123. Bosquet L, Porta-Benache J, Blais J. Validity of a commercial linear encoder to estimate bench press 1 RM from the force-velocity relationship. J Sports Sci Med. 2010;9:459–63.
124. Larsen F, Loturco I, Sigvaldsen E, Strand MF, Kalhovde JM, Haugen T. Reliability and validity of different lower-limb strength tests to determine 1RM in the Keiser A300 leg press. J Strength Cond Res. 2023;37:1963–8. https://doi.org/10.1519/JSC.0000000000004516.
125. Sigvaldsen E, Loturco I, Larsen F, Bruusgaard J, Kalhovde JM, Haugen T. Validity and reliability of upper body push and pull tests to determine one-repetition maximum. PLoS ONE. 2023;18:e0288649. https://doi.org/10.1371/journal.pone.0288649.
126. Oberhofer K, Erni R, Sayers M, Huber D, Lüthy F, Lorenzetti S. Validation of a smartwatch-based workout analysis application in exercise recognition, repetition count and prediction of 1RM in the strength training-specific setting 2021. Sports (Basel). 2021;9:118. https://doi.org/10.3390/sports9090118.
127. Picerno P, Iannetta D, Comotto S, Donati M, Pecoraro F, Zok M, et al. 1RM prediction: a novel methodology based on the force–velocity and load–velocity relationships. Eur J Appl Physiol. 2016;116:2035–43. https://doi.org/10.1007/s00421-016-3457-0.
128. Sandau I, Chaabene H, Granacher U. Predictive validity of the snatch pull force-velocity profile to determine the snatch one repetition-maximum in male and female elite weightlifters. J Funct Morphol Kinesiol. 2021;6:35. https://doi.org/10.3390/jfmk6020035.
129. Simonsen MB, Jolas E, Pedersen SR, Jensen JG, Faarup M, Jakobsen RT, et al. Analyzing one-repetition-max predictions: load-velocity relationship vs. repetition to failure equation in ten lower extremity exercises. Sport Sci Health. 2025;21:343–53. https://doi.org/10.1007/s11332-024-01264-y.
130. Soriano MA, Jiménez-Ormeño E, Haff GG, Comfort P, Giráldez-Costas V, Ruiz-Moreno C, et al. A comparison of three load-velocity based methods to estimate maximum overhead press performance in weightlifters. Int J Sports Sci Coach. 2023;18:1585–93. https://doi.org/10.1177/17479541221115854.
131. Sayers MGL, Schlaeppi M, Hitz M, Lorenzetti S. The impact of test loads on the accuracy of 1RM prediction using the load-velocity relationship. BMC Sports Sci Med Rehabil. 2018;10:9. https://doi.org/10.1186/s13102-018-0099-z.
132. Çetin O, Akyildiz Z, Demirtaş B, Sungur Y, Clemente FM, Cazan F, et al. Reliability and validity of the multi-point method and the 2-point method's variations of estimating the one-repetition maximum for deadlift and back squat exercises. PeerJ. 2022;10:e13013. https://doi.org/10.7717/peerj.13013.
133. Sandau I, Chaabene H, Granacher U. Validity and reliability of a snatch pull test to model the force-velocity relationship in male elite weightlifters. J Strength Cond Res. 2022;36:2808–15. https://doi.org/10.1519/jsc.0000000000003949.
134. García-Ramos A, Janicijevic D, González-Hernández JM, Keogh JWL, Weakley J. Reliability of the velocity achieved during the last repetition of sets to failure and its association with the velocity of the 1-repetition maximum. PeerJ. 2020;8:e8760. https://doi.org/10.7717/peerj.8760.
135. Pérez-Castilla A, Jerez-Mayorga D, Martínez-García D, Rodríguez-Perea Á, Chirosa-Ríos LJ, García-Ramos A. Comparison of the bench press one-repetition maximum obtained by different procedures: direct assessment vs. lifts-to-failure equations vs. two-point method. Int J Sports Sci Coach. 2020;15:337–46. https://doi.org/10.1177/1747954120911312.
136. Thompson SW, Rogerson D, Dorrell HF, Ruddock A, Barnes A. The reliability and validity of current technologies for measuring barbell velocity in the free-weight back squat and power clean. Sports (Basel). 2020;8:94. https://doi.org/10.3390/sports8070094.
137. Bartlett JW, Frost C. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. Ultrasound Obstet Gynecol. 2008;31:466–75. https://doi.org/10.1002/uog.5256.
138. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307–10. https://doi.org/10.1016/S0140-6736(86)90837-8.
139. Grgic J, Lazinica B, Schoenfeld BJ, Pedisic Z. Test–retest reliability of the one-repetition maximum (1RM) strength assessment: a systematic

review. Sports Med Open. 2020;6:1–16. https://doi.org/10.1186/s40798-020-00260-z.
140. Dias RMR, Cyrino ES, Salvador EP, Caldeira LFS, Nakamura FY, Papst RR, et al. Influence of familiarization process on muscular strength assessment in 1-RM tests. Rev Bras Med Esporte. 2005;11:39–42. https://doi.org/10.1590/S1517-86922005000100004.
141. Ploutz-Snyder LL, Giamis EL. Orientation and familiarization to 1RM strength testing in old and young women. J Strength Cond Res. 2001;15:519–23. https://doi.org/10.1519/00124278-200111000-00020.
142. García-Ramos A. The 2-point method: theoretical basis, methodological considerations, experimental support, and its application under field conditions. Int J Sports Physiol Perform. 2023;18:1092–100. https://doi.org/10.1123/ijspp.2023-0127.
143. Chen Z, Xiao F, Mao Y, Zhang X, García-Ramos A. An efficient and accurate approach for estimating the free-weight back squat 1-repetition maximum based on the 2-point method and optimal minimal velocity threshold. J Strength Cond Res. 2025;39:e530–7. https://doi.org/10.1519/JSC.0000000000005040.
144. Bobbert MF. Why is the force-velocity relationship in leg press tasks quasi-linear rather than hyperbolic. J Appl Physiol. 2012;112:1975–83. https://doi.org/10.1152/japplphysiol.00787.2011.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.