Within- and between-day test-retest reliability and agreement of isometric and isokinetic, multi-joint, upper- and lower-body strength testing

Michael Nash[1], James Steele[1,2], James P. Fisher[1*]

[1]Department of Sport and Health, Solent University, East Park Terrace, Southampton, UK
[2]Steele Research Limited, UK

*Correspondence: james.fisher@outlook.com

Cite as: Michael Nash et al. (2024). Within- and between-day test-retest reliability and agreement of isometric and isokinetic, multi-joint, upper- and lower-body strength testing. *SportRxiv.* DOI 10.51224/SRXIV.462

Supplementary Materials: https://osf.io/

Abstract

The aim of this study was to assess the within- and between-day test-retest agreement and reliability of upper- (chest press; CP and seated row; SR) and lower-body (leg press; LP) multi-joint isometric and isokinetic dynamometry using Exerbotics devices. Fourteen recreationally active adults (male = 11, female = 3) completed three testing sessions over a three-week period. On each day, participants performed two isometric testing trials followed by one isokinetic testing trial. Analyses were conducted within a Bayesian estimation framework using multivariate models summarising posterior distributions by their mean point estimate and 95% quantile intervals (QI) for both limits of agreement and variance decomposition ratios (comparable to intraclass correlation coefficients). For isometric testing, the within-day limits of agreement with the mean for the CP were ± 28.82 [95%QI: 23.12, 35.88], for the LP were ± 38.76 [95%QI: 30.87, 49.22], and for the SR were ± 16.65 [95%QI: 13.3, 21.18], and between-day limits of agreement with the mean for the CP were ± 34.06 [95%QI: 28.46, 41.4], for the LP were ± 50.46 [95%QI: 41.57, 62.43], and for the SR were ± 23.2 [95%QI: 19.04, 28.98]. Variance decomposition ratios were 0.939 [95%QI: 0.895, 0.964] and 0.937 [95%QI: 0.893, 0.963] for CP, 0.968 [95%QI: 0.944, 0.981] and 0.967 [95%QI: 0.942, 0.981] for LP, and 0.970 [95%QI: 0.947, 0.983] and 0.969 [95%QI: 0.946, 0.982] for the SR, for within- and between-day, respectively. For isokinetic testing, between day limits of agreement for concentric muscle actions for the CP were ± 25.73 [95%QI: 17.94, 37.67], for the LP were ± 74.24 [95%QI: 55.11, 100.1], and for the SR were ± 17.87 [95%QI: 12.08, 26.91], and for eccentric muscle actions for the CP were ± 37.42 [95%QI: 25.82, 55.54], for the LP were ± 119.25 [95%QI: 85.73, 168.24], and for the SR were ± 28.55 [95%QI: 19.88, 41.77]. Variance decomposition ratios were 0.954 [95%QI: 0.89, 0.984] and 0.931 [95%QI: 0.828, 0.977] for CP, 0.912 [95%QI: 0.797, 0.968] and 0.855 [95%QI: 0.652, 0.95] for LP, and 0.971 [95%QI: 0.927, 0.991] and 0.937 [95%QI: 0.851, 0.978] for the SR, for concentric and eccentric muscle actions respectively. In summary, these data suggest good agreement and good to excellent reliability for multi-joint isometric and isokinetic force measurement.

Keywords: *force; chest press; leg press; seated row*

Introduction

There are three principal methods for assessing muscular strength: isoinertial, isometric and isokinetic. Isoinertial strength assessments consist of lifting a constant external load over a given distance, relying upon trial-and-error with increasing load until performance of a complete single repetition (1-repetition maximum [RM]) (Kroemer et al., 1990) or other (e.g., 10RM) is impossible. In contrast, isometric and isokinetic strength assessments (also known as dynamometry) measure the force or torque applied against an immovable object and fixed angle (Wilson & Murphy, 1996) or at a computer-controlled velocity (Nuzzo et al., 2019), respectively. Accurately assessing muscular strength is important in numerous contexts and for multiple purposes, and in order to do this, testing methods are assessed for validity, reliability, and agreement (Hopkins, 2000).

Isokinetic testing is often viewed as the gold standard for assessing muscle strength (Dirnberger et al., 2013; Parraca et al., 2022). Additionally, isokinetic exercise is an effective training modality for increasing muscular strength more generally (Ratamess et al., 2016) and for injury rehabilitation and prehabilitation (Coudeyre et al., 2016). However, isokinetic strength testing appears to be predominantly performed at the single-joint level, with a relative dearth of application and attention given towards multi-joint testing (Dvir & Müller, 2020). This is despite multi-joint strength assessments being more reflective of the nature of muscular functions in everyday life and sports (Paoli et al., 2017). Of the studies which have assessed the reliability of multi-joint isokinetic dynamometry, data supports this to be a moderately-to-highly reliable means of assessing both concentric and eccentric muscular force. For example, considering test-retest reliability of the Exerbotics squat device over two days using seventeen trained males, Stock and Luera (2014) reported ICC values of 0.74 and 0.70 for concentric and eccentric peak force, respectively. Reliability using the same device with a similar population group improves to 0.95 and 0.90 for concentric and eccentric, respectively, when testing is completed over three days, suggesting a learning effect (Dirnberger et al., 2013). Indeed, Dirnberger et al. (2013) also observed significant increases in reliability values from day one and two (ICC=0.82-0.94) to days two and three (ICC=0.91-0.97) for the Isomed 2000 leg press (LP) isokinetic dynamometer. Such variation is potentially due to a combination of the greater complexity in muscle co-ordination and recruitment in isokinetic compared to isometric testing (Callaghan et al., 2000), the unfamiliar nature of isokinetic movement testing more generally (Hopkins et al., 2001; Schärer et al., 2019), and the impact of measurement scheduling (Kroll, 1970). These findings led the authors to posit that a familiarisation session would improve reliability for multi-joint isokinetic movements. Two additional training studies reference the reliability of multi-joint isokinetic devices. These studies utilised the chest press (CP) (Hoffman et al., 2011), and CP and seated row (SR) (Ratamess et al., 2016) by Exerbotics. The latter references the former for test-retest reliability of the device, with the former stating "Test-retest reliability of the dynamometer has been established in our laboratory as r=0.99" (pp. 2237). However, no further information on the test-retest reliability assessment is provided. Further, Hoffman and colleagues' (2011) intervention was limited to the CP, thus, it is unclear if the reliability established therein was limited to the CP or also included the SR. Finally, and perhaps most importantly, the body of literature has typically investigated reliability – a measure of whether persons can be distinguished from each other, whereas, given that most utilisation of strength measurement is for the purpose of determining whether a change has occurred from test-to-test, a focus should be placed on agreement (i.e., how close the results of repeated measurements are) (Vet et al., 2006). With the above in mind, the aim of this study was to determine the reliability and agreement of isometric and isokinetic upper- and lower-body multi-joint dynamometry using Exerbotics CP, LP and SR exercises.

METHODS

Experimental Approach to the Problem

A within-group, repeated measures study design was employed, with participants reporting to the laboratory over three separate days. On each visit, participants performed two isometric testing trials followed by one isokinetic testing trial. Testing was conducted for three exercises: LP, CP, and SR (Exerbotics, LLC, Tulsa, OK, USA). This methodological design permitted assessment of within- and between-day agreement for peak isometric force for CP, LP, and SR, and between-day agreement of isokinetic concentric and eccentric force for the same exercises as, based on prior research suggesting a learning effect for isokinetic exercise we treated the first visit as familiarisation. Isokinetic values are reported as the average force throughout the full range of motion for each phase (concentric and eccentric) of a repetition for the repetition with the highest result.

Participants

Following ethical approval for this study by Solent University Health, Exercise, and Sports Science Research Ethics and Innovation Committee (reference number: nashm1HESS2023), recreationally active adults (male or female) aged between 18 and 45 years of any resistance training experience level (including none), were recruited. Inclusion criteria required participants to be absent of any injury and/or cardiorespiratory medical condition (including hypertensive or prehypertensive blood pressure) preventing or contraindicating maximal strength testing. Exclusion criteria included anyone not meeting the inclusion criteria as well as anyone who was pregnant or using uncontrolled performance enhancing drugs. Through convenience snowball sampling, 17 participants were recruited. One participant withdrew due to illness unrelated to the study, one withdrew due to personal reasons unrelated to the study and one other participant's data were excluded due to non-adherence to on-going eligibility criteria. All other participants completed all testing, resulting in data for 14 participants (mean ± SD; age, 31.07 ± 9.24 years; height, 178.06 ± 9.02 cm; mass, 79.38 ± 10.22 kg; resistance training experience, 9.93 ± 10.97 years) being available for final analysis. All participants completed a physical activity readiness questionnaire and signed an informed consent form.

Procedures

Machine set-up

This study consisted of three separate, consecutive, weekly visits to the laboratory. Testing days were scheduled on the same day of the week and the same time of day (±1 hour) to control for the impact of diurnal variations on maximum strength performance (Chtourou et al., 2012; Knaier et al., 2022). On testing day one, participants' range of motion and joint angles were established and recorded. For isometric testing, as per manufacturer's recommendations, joint angles were set at 90° flexion of the elbow for the CP and SR, and 90° flexion of the knee for the LP. For isokinetic testing, as per manufacturer's recommendations and to reflect usage in commercial settings, range of motion was set as follows; CP: 90° to 140° flexion of the elbow; SR: 140° to 90° flexion of the elbow; and LP: 90° to 120° flexion of the knee. Participants were allowed to self-select seat position, foot position, and hand position, which were also recorded to ensure consistency between days and trials. On each day of testing, participants performed a standardised warm-up on a cycle ergometer (Ergomedic 874E; Monark, Uppsala, Sweden) for five minutes up to 60% age-predicted heart rate maximum (220-age).

Isometric Testing

Following the warm-up, participants performed two trials consisting of three isometric contractions per exercise. Isometric contractions one and two of each set acted as specific warm-up performed at an estimated 50% and 75% maximal effort, respectively. The third repetition of each trial was performed at maximal effort and used for analysis herein. This was repeated for each exercise, and each testing condition. For maximal tests, participants were instructed to push as hard and as fast as they could against the movement arm and keep pushing for 3 seconds to ensure the maximal value was recorded. Isometric testing was always performed prior to isokinetic testing and always in the order of CP, LP, SR exercises (with 1-minute rest between exercises, and 2-minutes rest between trials), to ensure each testing modality and each exercise was performed in the same state of within-day fatigue and, thus, differences in between-day states of fatigue did not confound results.

Isokinetic Testing

Following the completion of any isometric testing and a subsequent 3-minute rest interval, isokinetic testing was performed. On day one, participants underwent a familiarisation session with the devices, which consisted of eight repetitions per exercise. Participants were instructed to perform the repetitions at an estimated 30-50% of their maximal effort. On days two and three, participants performed one trial per exercise, with each trial consisting of three repetitions: one repetition at an estimated 30-50% maximal effort immediately followed by two maximal effort repetitions. Participants rested for 2-minutes before testing on the next exercise. The maximal effort repetition with the highest mean concentric and eccentric peak torque for each exercise was recorded for analysis. Isokinetic testing was performed at a repetition cadence of four-second concentric, half-second pause, four-second eccentric muscle actions. This equated to 12.5°/sec at the elbow angle for CP and SR, and 7.5°/sec for at the knee for LP, as per manufacturers recommendations.

Monitors were positioned such that participants received no visual feedback. Standardised verbal encouragement @(Engel et al., 2019) of "push" during the isometric and concentric of the isokinetic CP and LP tests, "pull" during the isometric and concentric of the isokinetic SR test, and "resist" during the eccentric of all isokinetic tests, was provided. Participants were instructed to maintain their dietary, hydration and caffeine habits, and to refrain from strenuous physical exercise other than activities of daily living for at least 48 hours prior to each session.

Statistical Analysis

All code utilised for data preparation and analyses are available in either the Open Science Framework page for this project https://osf.io/zrxjp/ or the corresponding GitHub repository https://github.com/jamessteeleii/isokinetic_isometric_agreement_reliability. We cite all software and packages used in the analysis pipeline using the `grateful` package (Rodriguez-Sanchez et al., 2023) which can be seen here: https://osf.io/pgx6v. This project was not pre-registered, but had an exploratory estimation goal. All analyses have been conducted within a Bayesian posterior estimation framework and all posterior estimates and their precision, along with conclusions based upon them, will be interpreted continuously and probabilistically, considering priors, data quality, and all within the context of each outcome and the assumptions of the model employed as the estimator (Kruschke & Liddell, 2018). Given that most utilisation of strength measurement is for the purpose of determining whether a change has occurred from test-to-test, we focused on the *agreement* of measurements as opposed to the typical *reliability* statistics which instead reflect the ability of measurements to distinguish between individuals (Berchtold, 2016; Kottner & Streiner, 2011; Vet et al., 2006). However, we report on the reliability in the form of variance decomposition ratios for our Bayesian models calculated directly from the posterior predictve distributions which are comparable to intraclass correlation coefficients (ICC) in order to compare to prior research. Here we also employed multivariate mixed effects methods for examining agreement by variance components enabling us to model all three exercises, chest press, leg press, and row, simultaneously extending previous approaches (Schluter, 2009) thus offering greater precision, robustness, and efficiency of estimates. Two sets of models, each detailed below, were used to examine both the between- and within-day agreement for isometric outcomes, and for between-day agreement for isokinetic outcomes. In each model we employed informative yet weakly regularising priors which are detailed below. All models were fit with four Markov Chain Mote Carlo chains using 2000 warmup and 6000 sampling iterations. Trace plots were produced along with $\hat{R}$ values to examine whether chains had converged, and posterior predictive checks for each model were also examined to understand the model implied distributions. Note, all values are in Newtons of force.

Isometric Outcomes

Given that for isometric outcomes we had both three repeated days of testing, and two repeated trials within each day, we opted to adapt the methods described by Jones et al. (2011) and Christensen et al. (2020) to derive the limits of agreement with the mean. Typically where there are two measurements to compare in terms of agreement the traditional Bland-Altman Limits of Agreement approach can be employed (Bland & Altman, 1986). When there are multiple measurements (whether multiple methods, observers, or tests, or whether there are replicates within these) it is more difficult to apply these typical models. Instead, we can model the outcomes directly and derive the agreement with the *mean* value for the participant over the repeated measurements made. In the case where we can assume there is no bias for a particular measurement (in our case no particular bias across days for example), then we can assume that the mean reflects a good estimate of the true value and the 95% limits of agreement then reflect the range over which we would expect measurements to fall about the true value 95% of the time.

In the present case we sought to partition the variance components such that we could determine separately the between-participant variance (i.e., α_{0i} below), the between-day variance (i.e., α_{1ij} below), and the residual variance which here reflects the within-day variance (i.e., $\alpha_{\mu_{ij}}$ below). We estimate these variance components through a multivariate mixed effects model of the joint three exercise outcomes observed (chest press, leg press, and row; see Equation 1). The model included a population (i.e., fixed) effect for day which was Helmert coded; This meant that for the three days we have two coefficients in the model for each outcome with the first, β_1Day below, reflecting the difference between the mean of day one and the mean of day two and three, and the second, β_2Day below, reflected the difference between mean of day two and the mean of day three. This allowed us to examine whether there was any systematic bias, and in this case the Helmert coding was specifically used because we anticipated that any bias would manifest in terms of a "familiarisation" effect whereby participants improved systematically with repeated measurement. The model also included random (i.e., group level terms) intercepts for participants (i.e., α_{0i} below), and random intercepts for day nested within participant (i.e., α_{1ij} below). Each of these were modelled as correlated between outcomes reflecting the models assumption that typically participants that are stronger are stronger across each exercise tested (i.e., strength is correlated between exercises), and also that variation across days was likely to also be related reflecting that lower/higher values on a given exercise on one day would likely be related to lower/higher values on a different exercise on that day. Lastly, the residual errors were also modelled as correlated. The model for isometric measurements can thus be represented as follows in Equation 1:

$$
\begin{aligned}
\begin{bmatrix} \text{Chest Press}_{ijk} \\ \text{Leg Press}_{ijk} \\ \text{Row}_{ijk} \end{bmatrix} &\sim MVN\left(\begin{bmatrix} \mu_{ij}^{\text{Chest Press}} \\ \mu_{ij}^{\text{Leg Press}} \\ \mu_{ij}^{\text{Row}} \end{bmatrix}, \Sigma_{obs}\right) \\
\mu_{ij}^{\text{Chest Press}} &= \beta_0^{\text{Chest Press}} + \beta_1^{\text{Chest Press}}\text{Day}_1 + \beta_2^{\text{Chest Press}}\text{Day}_2 + \alpha_{0i}^{\text{Chest Press}} + \alpha_{1ij}^{\text{Chest Press}} \\
\mu_{ij}^{\text{Leg Press}} &= \beta_0^{\text{Leg Press}} + \beta_1^{\text{Leg Press}}\text{Day}_1 + \beta_2^{\text{Leg Press}}\text{Day}_2 + \alpha_{0i}^{\text{Leg Press}} + \alpha_{1ij}^{\text{Leg Press}} \\
\mu_{ij}^{\text{Row}} &= \beta_0^{\text{Row}} + \beta_1^{\text{Row}}\text{Day}_1 + \beta_2^{\text{Row}}\text{Day}_2 + \alpha_{0i}^{\text{Row}} + \alpha_{1ij}^{\text{Row}} \\
\begin{bmatrix} \alpha_{0i}^{\text{Chest Press}} \\ \alpha_{0i}^{\text{Leg Press}} \\ \alpha_{0i}^{\text{Row}} \end{bmatrix} &\sim MVN\left(0, \Sigma_{Participant}\right) \\
\begin{bmatrix} \alpha_{1ij}^{\text{Chest Press}} \\ \alpha_{1ij}^{\text{Leg Press}} \\ \alpha_{1ij}^{\text{Row}} \end{bmatrix} &\sim MVN\left(0, \Sigma_{Participant:Day}\right) \\
\Sigma_{Obs} &= \begin{pmatrix} \sigma_{\mu_{ij}^{\text{Chest Press}}} & 0 & 0 \\ 0 & \sigma_{\mu_{ij}^{\text{Leg Press}}} & 0 \\ 0 & 0 & \sigma_{\mu_{ij}^{\text{Row}}} \end{pmatrix} \mathrm{R} \begin{pmatrix} \sigma_{\mu_{ij}^{\text{Chest Press}}} & 0 & 0 \\ 0 & \sigma_{\mu_{ij}^{\text{Leg Press}}} & 0 \\ 0 & 0 & \sigma_{\mu_{ij}^{\text{Row}}} \end{pmatrix} \\
\Sigma_{Participant} &= \begin{pmatrix} \sigma_{\alpha_{0i}^{\text{Chest Press}}} & 0 & 0 \\ 0 & \sigma_{\alpha_{0i}^{\text{Leg Press}}} & 0 \\ 0 & 0 & \sigma_{\alpha_{0i}^{\text{Row}}} \end{pmatrix} \mathrm{R} \begin{pmatrix} \sigma_{\alpha_{0i}^{\text{Chest Press}}} & 0 & 0 \\ 0 & \sigma_{\alpha_{0i}^{\text{Leg Press}}} & 0 \\ 0 & 0 & \sigma_{\alpha_{0i}^{\text{Row}}} \end{pmatrix} \\
\Sigma_{Participant:Day} &= \begin{pmatrix} \sigma_{\alpha_{1ij}^{\text{Chest Press}}} & 0 & 0 \\ 0 & \sigma_{\alpha_{1ij}^{\text{Leg Press}}} & 0 \\ 0 & 0 & \sigma_{\alpha_{1ij}^{\text{Row}}} \end{pmatrix} \mathrm{R} \begin{pmatrix} \sigma_{\alpha_{1ij}^{\text{Chest Press}}} & 0 & 0 \\ 0 & \sigma_{\alpha_{1ij}^{\text{Leg Press}}} & 0 \\ 0 & 0 & \sigma_{\alpha_{1ij}^{\text{Row}}} \end{pmatrix}
\end{aligned}
\tag{1}
$$

where each exercise outcome is represented by a superscript for observations and model parameters (i.e., $\text{Chest Press}, \text{Leg Press}, or$ Row), and for a given exercise the subscripts reflect the kth measurement (k = 1, ... ,K), from the jth day (j = 1, ... ,J) for the ith participant (i = 1, ... ,I). Population and group (i.e., fixed and random) parameters are described above. The covariance matrices for observations, random intercepts for participant, and random intercepts for day within participant are given by Σ_{Obs}, $\Sigma_{Participant}$, $\Sigma_{Participant:Day}$.

As mentioned above we adopted informative yet weakly regularising priors. Default priors in the `brms` R package used to fit the model are weakly regularising on all intercept terms (i.e., β_0) and are set such that they are centred and scaled using a $student\ t$ distribution with $df = 3$ and represent the expected response value when all predictors are at their means, all group level terms are set with a $student\ t$ distribution with $df = 3$, a $\mu = 0$, and scaled to the expected response values, and all correlation matrices R are set with an $LKJcorr\left(1\right)$

distribution. The remaining population effect coefficients are by default set with an improper flat prior on the reals (i.e., $uniform(lb = -\infty, ub = \infty)$) and thus we opted to set our own informative weakly regularising priors based on the raw data descriptives. Typically a 10% coefficient of variation is deemed acceptable for strength measures test-retest variation (Nuzzo et al., 2019) and so we opted for a slightly more skeptical prior with a location set at 20% of the sample arithmetic mean of all observations for each exercise outcome. Further, we assumed a simple propagation of error approach for two independent measurements (i.e., ignoring covariance and thus reflecting a lack of knowledge about the exact nature of it) again utilising the sample variance of all observations for each exercise outcome. Thus, the priors for the model in Equation 1 were:

$$\begin{aligned}
&\begin{matrix}\beta_1^{\text{Chest Press}}\text{Day}_1\\ \beta_2^{\text{Chest Press}}\text{Day}_2\end{matrix} \sim student\ t\left(df=3, \mu=0.2\left(\frac{1}{n}\sum_{ijk=ijk}^{n}\text{Chest Press}_{ijk}\right), \sigma=2\sqrt{\frac{1}{N-1}\sum_{ijk=1}^{N}(\text{Chest Press}_{ijk}-\overline{\text{Chest Press}})^2}\right)\\
&\begin{matrix}\beta_1^{\text{Leg Press}}\text{Day}_1\\ \beta_2^{\text{Leg Press}}\text{Day}_2\end{matrix} \sim student\ t\left(df=3, \mu=0.2\left(\frac{1}{n}\sum_{ijk=ijk}^{n}\text{Leg Press}_{ijk}\right), \sigma=2\sqrt{\frac{1}{N-1}\sum_{ijk=1}^{N}(\text{Leg Press}_{ijk}-\overline{\text{Leg Press}})^2}\right)\\
&\begin{matrix}\beta_1^{\text{Row}}\text{Day}_1\\ \beta_2^{\text{Row}}\text{Day}_2\end{matrix} \sim student\ t\left(df=3, \mu=0.2\left(\frac{1}{n}\sum_{ijk=ijk}^{n}\text{Row}_{ijk}\right), \sigma=2\sqrt{\frac{1}{N-1}\sum_{ijk=1}^{N}(\text{Row}_{ijk}-\overline{\text{Row}})^2}\right)
\end{aligned} \tag{2}$$

The between- and within-day 95% limits of agreement with the mean where calculated from the posterior draws of the relevant variance components. These were calculated adapting the approach of Christensen et al. (2020) adjusting for the degrees of freedom based upon the number of days, and number of replicates within days. The between-day limits of agreement with the mean were calculated using the between day variance component for each exercise outcome, α_{1ij}, as:

$$\pm 1.96\sqrt{\frac{J-1}{J}\alpha_{1ij}^2} \tag{3}$$

And the within-day limits of agreement with the mean utilising the remaining within day residual variance component for each exercise outcome, $\alpha_{\mu_{ij}}$, as:

$$\pm 1.96\sqrt{\frac{JK-1}{JK}\alpha_{\mu_{ij}}^2} \tag{4}$$

For each posterior draw the limits of agreement with the mean were calculated and then the mean and 95% quantile intervals (QI) determined. These were then presented graphically with a limits of agreement with the mean plot following the methods described by Jones et al. (2011) and Christensen et al. (2020) where the participant mean for each exercise outcome, $\overline{y}_{i..}$, was plot on the x-axis and the difference between each observation for each exercise outcome with the mean, $y_{ijk} - \overline{y}_{i..}$, plot on the y-axis with the corresponding limits of agreement with the mean for both between- and within-day plot about these. We also present the posterior distribution for the bias reflected by the β_1Day and β_2Day coefficients along with their corresponding mean and 95% quantile interval. The variance decomposition ratios were calculated for both between- and within-day by calculating the ratio between the variance for draws from the posterior predictive distribution not conditioned on random (i.e., group level terms) and the variance for draws conditioned on the appropriate random effects. The mean and 95% quantile interval for these were then calculated.

Isokinetic Outcomes

For the isokinetic outcomes we only had two repeated days of testing, and for each day a single "best" repetition measured for each exercise outcome and for both concentric and eccentric phases. We opted to model the concentric and eccentric phases separately as we suspected, whilst they may be correlated, the between-day agreement might differ for either. As such, given we only had two measurements between days for each exercise outcome and muscle action a traditional Bland-Altman Limits of Agreement approach could be employed (Bland & Altman, 1986). We did however still adopt a multivariate approach allowing for the residual errors were also modelled as correlated which in this case, given our model described below had no other predictors (i.e., they include an intercept only for each outcome) these are the correlations between the between-day differences in each outcome. Where δ_i for each exercise outcome in each muscle action is the difference between days for the ith participant (i = 1, … ,I) i.e., $\delta_i = y_{i2} - y_{i1}$, the model is:

$$\begin{bmatrix} \delta_i^{\text{Chest Press (con)}} \\ \delta_i^{\text{Leg Press (con)}} \\ \delta_i^{\text{Row (con)}} \\ \delta_i^{\text{Chest Press (ecc)}} \\ \delta_i^{\text{Leg Press (ecc)}} \\ \delta_i^{\text{Row (ecc)}} \end{bmatrix} \sim MVN \left(\begin{bmatrix} \mu_i^{\text{Chest Press (con)}} \\ \mu_i^{\text{Leg Press (con)}} \\ \mu_i^{\text{Row (con)}} \\ \mu_i^{\text{Chest Press (ecc)}} \\ \mu_i^{\text{Leg Press (ecc)}} \\ \mu_i^{\text{Row (ecc)}} \end{bmatrix}, \Sigma_{obs}^* \right) \tag{5}$$

where Σ_{Obs}^* is the residual covariance matrix (omitted due to size).

Weakly regularising default priors were used again on all intercept terms (i.e., μ_i) for each outcome and set such that they were centred and scaled using a $student\ t$ distribution with $df = 3$ representing the expected response value when all predictors are at their means which in this case meant the raw means. The residual correlation matrix R was set with an $LKJcorr\,(1)$ distribution.

In this model the 95% limits of agreement are calculated simply as $\pm 1.96 \times \sigma_{\mu i}$ for each outcome. Again this was calculated for each outcome for each posterior draw and the mean and 95% quantile intervals calculated. The bias for each outcome then are the intercept terms μ_i and the corresponding means and 95% quantile intervals for these were also determined. These were plot together in a traditional Bland-Altman limits of agreement plot where the participant mean for each exercise outcome, $\bar{y}_i$, was plot on the x-axis and the difference between days for each exercise outcome, δ_i, plot on the y-axis with the corresponding limits of agreement and mean bias plot about these.

For the variance decomposition ratios for isokinetic outcomes a separate multivariate model was fit for y_{ij} for each exercise outcome and muscle action as follows:

$$\begin{bmatrix} \text{Chest Press (con)}_{ij} \\ \text{Leg Press (con)}_{ij} \\ \text{Row (con)}_{ij} \\ \text{Chest Press (ecc)}_{ij} \\ \text{Leg Press (ecc)}_{ij} \\ \text{Row (ecc)}_{ij} \end{bmatrix} \sim MVN\left(\begin{bmatrix} \mu_i^{\text{Chest Press (con)}} \\ \mu_i^{\text{Leg Press (con)}} \\ \mu_i^{\text{Row (con)}} \\ \mu_i^{\text{Chest Press (ecc)}} \\ \mu_i^{\text{Leg Press (ecc)}} \\ \mu_i^{\text{Row (ecc)}} \end{bmatrix}, \Sigma_{obs}^* \right)$$
$$\begin{bmatrix} \mu_i^{\text{Chest Press (con)}} \\ \mu_i^{\text{Leg Press (con)}} \\ \mu_i^{\text{Row (con)}} \\ \mu_i^{\text{Chest Press (ecc)}} \\ \mu_i^{\text{Leg Press (ecc)}} \\ \mu_i^{\text{Row (ecc)}} \end{bmatrix} \sim MVN\left(0, \Sigma_{Participant}\right) \tag{6}$$

where Σ^*_{Obs} is the residual covariance matrix and $\Sigma^*_{Participant}$ the random intercept covariance matrix (both also omitted due to size).

Weakly regularising default priors were used again on all intercept terms (i.e., μ_i) for each outcome and set such that they were centred and scaled using a $student\ t$ distribution with $df = 3$ representing the expected response value when all predictors are at their means which in this case meant the raw means. The residual correlation matrix R was set with an $LKJcorr\,(1)$ distribution.

Variance decomposition ratios were then calculated for between-day by calculating the ratio between the variance for draws from the posterior predictive distribution not conditioned on random (i.e., group level terms) and the variance for draws conditioned on the appropriate random effects. The mean and 95% quantile interval for these were then calculated.

Results

Isometric Outcomes

The isometric model showed chain convergence with $\hat{R}$ values all < 1.01 and posterior predictive checks were good. Model diagnostics can be seen in the supplementary materials here: https://osf.io/ypt59.

Figure 1 shows the mean bias and limits of agreement with the mean for both between- and within-day for each exercise. There was no clear evidence of a "familiarisation" biasing effect between days given that the sign of the contrasts both between- and within-exercises was variable and the posterior distributions typically all ranged from both small positive to negative effects. As might be expected, the between-day limits of agreement with the mean were greater than the within-day agreement. The between-day limits of agreement with the mean for the chest press were $\pm$ 34.06 [95%QI: 28.46, 41.4], for the leg press were $\pm$ 50.46

[95%QI: 41.57, 62.43], and for the row were $\pm$ 23.2 [95%QI: 19.04, 28.98]. The within-day limits of agreement with the mean for the chest press were $\pm$ 28.82 [95%QI: 23.12, 35.88], for the leg press were $\pm$ 38.76 [95%QI: 30.87, 49.22], and for the row were $\pm$ 16.65 [95%QI: 13.3, 21.18].

The variance decomposition ratio (comparable to the ICC) for the chest press between-day was 0.937 [95%QI: 0.893, 0.963] and within-day was 0.939 [95%QI: 0.895, 0.964]. For the leg press the variance decomposition ratio between-day was 0.967 [95%QI: 0.942, 0.981] and within-day was 0.968 [95%QI: 0.944, 0.981]. For the row the variance decomposition ratio between-day was 0.969 [95%QI: 0.946, 0.982] and within-day was 0.97 [95%QI: 0.947, 0.983].

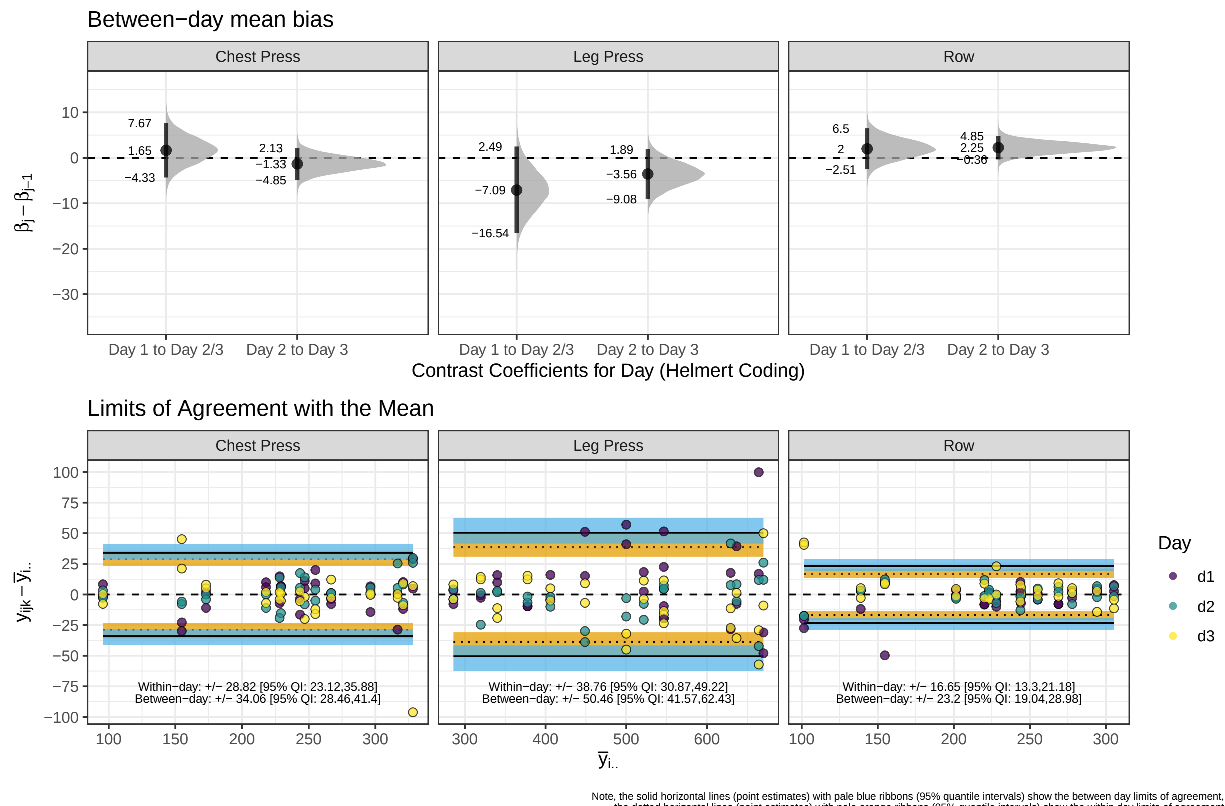

Figure 1: Bias and limits of agreement with the mean for isometric outcomes for both between- and within-day.

Isokinetic Outcomes

The isokinetic model also showed chain convergence with $\hat{R}$ values all < 1.01 and posterior predictive checks were good. Model diagnostics can be seen in the supplementary materials here: https://osf.io/wzh9m.

Figure 2 shows the mean bias and limits of agreement between-day for each exercise and muscle action. There was no clear evidence of a “familiarisation” biasing effect between days for most exercises and muscle actions, perhaps with the exception of the eccentric leg press which showed somewhat of an improvement from day one to day two: 3.87 [95%QI: -14.02, 21.98]. The limits of agreement for concentric muscle actions for the chest press were $\pm$ 25.73 [95%QI: 17.94, 37.67], for the leg press were $\pm$ 74.24 [95%QI: 55.11, 100.1], and for the row were $\pm$ 17.87 [95%QI: 12.08, 26.91]. The limits of agreement for eccentric muscle actions for the chest press were $\pm$ 37.42 [95%QI: 25.82, 55.54], for the leg press were $\pm$ 119.25 [95%QI: 85.73, 168.24], and for the row were $\pm$ 28.55 [95%QI: 19.88, 41.77].

The variance decomposition ratio (comparable to the ICC) for concentric muscle actions for the chest press between-day was 0.954 [95%QI: 0.89, 0.984], for the leg press was 0.912 [95%QI: 0.797, 0.968], and for the row was 0.971 [95%QI: 0.927, 0.991]. For eccentric muscle actions the variance decomposition ratio for the chest press was 0.931 [95%QI: 0.828, 0.977], for the leg press was 0.855 [95%QI: 0.652, 0.95], and for the row was 0.937 [95%QI: 0.851, 0.978].

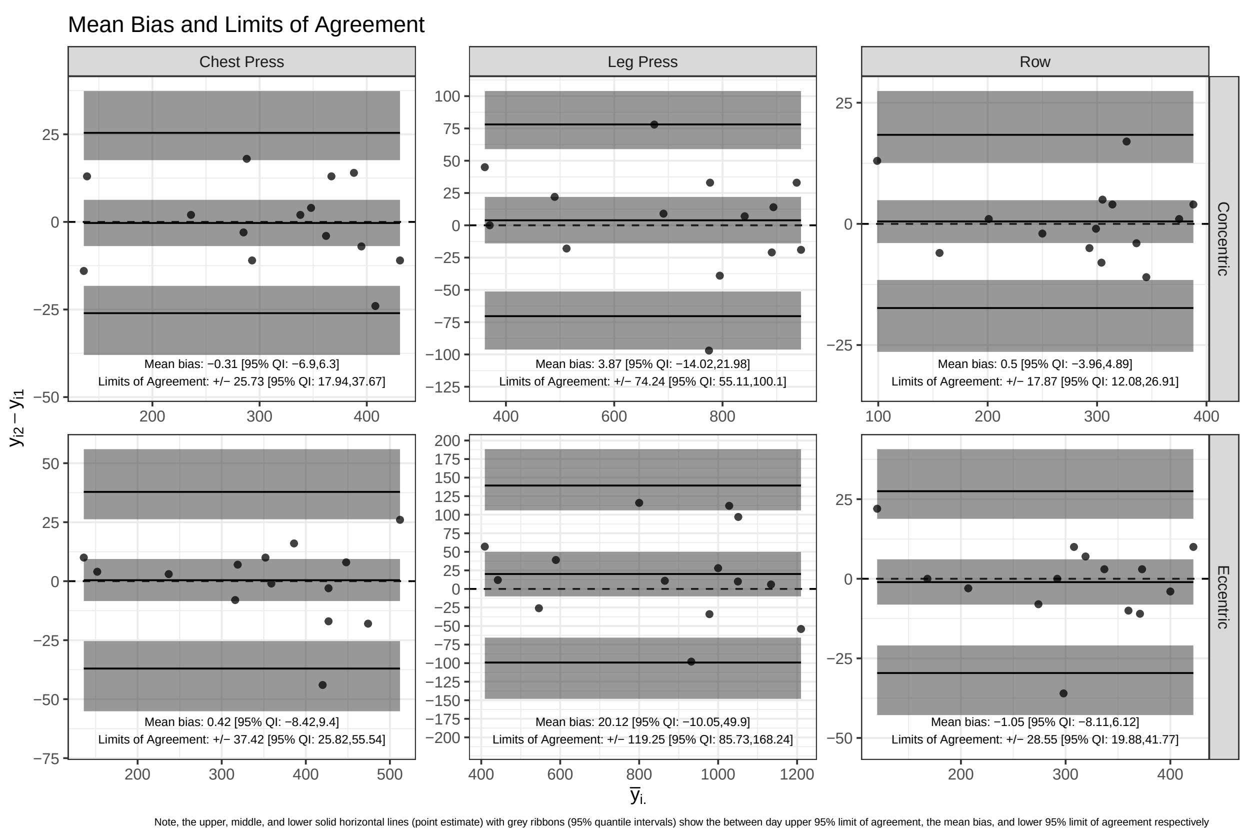

Figure 2: Between day bias and limits of agreement for isokinetic outcomes.

DISCUSSION

The aim of this study was to determine the reliability and agreement of isometric and isokinetic upper- and lower-body multi-joint dynamometry using Exerbotics CP, LP and SR devices. To the author's knowledge, this is the first study to conduct a dedicated investigation into the agreement of multi-joint isometric dynamometry using these particular Exerbotics devices, and to examine bilateral isokinetic dynamometry using the CP and SR exercises irrespective of manufacturer. This study, therefore, produced novel findings that provide valuable contributions to the literature.

Firstly, considering isometric testing, our data suggests no clear evidence of a "familiarisation" biasing effect between days. However, and as might be expected, the between-day limits of agreement were greater than the within-day agreement (between-day = CP ± 34.06 [95%QI: 28.46, 41.4], LP ± 50.46 [95%QI: 41.57, 62.43], and SR ± 23.2 [95%QI: 19.04, 28.98]; within-day CP ± 28.82 [95%QI: 23.12, 35.88], LP ± 38.76 [95%QI: 30.87, 49.22], and SR ± 16.65 [95%QI: 13.3, 21.18]). Although, between-day agreement is likely more important for determining strength changes following an intervention, rather than assessment of acute fatigue following a given task. In this sense we could consider the between-day agreement relative to the mean across all tests for a given exercise where we assume no training effect here; CP=327N, agreement ~10%; LP=668N, agreement ~7%; SR=306N, agreement ~8%. In context, a recent meta-analysis reported an overall estimate strength increase of ~22% (~19% to ~25%) for 2683 participants across 111 studies (Steele et al., 2023). Thus, between day agreement seems acceptable enough for this testing method to have the sensitivity to determine true strength changes. As such we propose these tests show good between-day agreement. Finally, our variance decomposition ratio values can be interpreted similar to ICC values. That is to say that the CP showed good to excellent reliability within- (0.939 [95%QI: 0.895, 0.964]) and between-day (0.937 [95%QI: 0.893, 0.963]), and the LP and SR showed excellent reliability within- (LP=0.968 [95%QI: 0.944, 0.981], SR=0.97 [95%QI: 0.947, 0.983]), and between-day (LP=0.967 [95%QI: 0.942, 0.981], SR=0.969 [95%QI: 0.946, 0.982]) interpreted by values proposed by Koo and Li (2016).

Considering isokinetic testing, once again, there was no clear evidence of a "familiarisation" biasing effect between days for most exercises and muscle actions, perhaps with the exception of the eccentric LP which showed a slight improvement from day one to day two of ~4%. This is in line with verbal feedback from multiple participants who reported how unusual it felt to press against a motorised footplate moving the lower body into flexion. Indeed, authors have reported that variation in strength testing can be exacerbated by the unfamiliar nature of isokinetic movements (Hopkins et al., 2001). The between-day limits of agreement were better for concentric muscle actions compared to eccentric muscle actions (concentric mean estimates = CP ± 25.73 [95%QI: 17.94, 37.67], LP ± 74.24 [95%QI: 55.11, 100.1], and SR ± 17.87 [95%QI: 12.08, 26.91]; eccentric mean estimates = CP ± 37.42 [95%QI: 25.82, 55.54], LP ± 119.25 [95%QI: 85.73, 168.24], and SR ± 28.55 [95%QI: 19.88, 41.77]). This might be expected as a person might be more familiar with applying to force to move an object rather than applying force to resist movement of an object. We can also present between-day agreement in terms of relative to the mean across all tests for a given exercise and muscle action where we assume no training effect here; CP= 214Nm, agreement ~12% concentric, 240Nm, agreement ~15% eccentric; LP= 482Nm, agreement ~15% concentric, 583Nm, agreement ~20% eccentric; SR= 193Nm, agreement ~10% concentric, 206Nm, agreement ~14% eccentric. Whilst agreement is relatively speaking poorer for eccentric movements, for most exercises and muscle actions it might be deemed acceptable in that a true strength change might be detectable using these devices and this test mode. Furthermore, while eccentric force values were greater than concentric force values (mean; = 240Nm and 214Nm, difference ~12%; LP= 583Nm and 482Nm, difference ~21%; SR= 206Nm and 193Nm, difference ~6%, for eccentric and concentric muscle actions, respectively) these were considerably less than the difference of ~40% recently reported in a large meta-analysis, though well within the prediction interval across studies (-3% to 103%) (Nuzzo et al., 2023). However, the authors reported findings from a meta-regression supporting that eccentric: concentric strength ratio was impacted by movement velocity, with an increase in velocity producing an increase in the ratio (Nuzzo et al., 2023). Within the present study, movement velocities were very low; 12.5°/sec at the elbow for CP and SR, and 7.5°/sec for at the knee for LP, as per manufacturers recommendations.

While comparison to previous research is difficult because of a dearth of literature considering multi-joint isokinetic strength testing and since previous studies have failed to calculate or report agreement statistics specifically, we have reported the variance decomposition ratio from our Bayesian models which is comparable to ICC values. Previous literature using lower body multi-joint isokinetic testing reported mean ICC values of 0.95 (90%CI=0.87-0.98) for concentric and 0.90 (90%CI=0.76-0.97) for eccentric (Bridgeman et al., 2016), and 0.804 and 0.736 for concentric and eccentric, respectively (Stock & Luera, 2014). Our own data shows similar reliability (LP mean estimates; concentric=0.912 [95%QI: 0.797, 0.968] considered good to excellent (Koo & Li, 2016), eccentric=0.855 [95%QI: 0.652, 0.95]) interpreted as moderate to excellent (Koo & Li, 2016). However, the previous studies used a squat device compared to a LP device in the present study – which while superficially similar apply different forces through the lower and upper back. Furthermore, the manufacturers (Exerbotics) recommend a different ROM for the two devices; squat = knee angle of 90-170° (Bridgeman et al., 2016; Stock & Luera, 2014), and LP = knee angle of 90-120° herein. We also reported variance decomposition ratios for upper body exercises (CP mean estimates; concentric=0.954 [95%QI: 0.89, 0.984], eccentric=0.931 [95%QI: 0.828, 0.977], SR mean estimates; concentric=0.971 [95%QI: 0.927, 0.991], eccentric=0.937 [95%QI: 0.851, 0.978]) which are all deemed good to excellent (Koo & Li, 2016), though there are no known data in the literature to compare these values against. Similar to previous research (Bridgeman et al., 2016; Stock & Luera, 2014) our own data for concentric muscle actions appears more reliable than for eccentric muscle actions (i.e., larger values and narrower ranges).

Finally, we should recognise potential limitations of our research. While our sample size is commensurate with previous studies discussing test-retest of isokinetic devices (e.g., n=10-17) (Bridgeman et al., 2016; Hoffman et al., 2011; Ratamess et al., 2016; Stock & Luera, 2014) we acknowledge that a larger sample would be of benefit in terms of the precision of estimates for both agreement and reliability. In addition, while we have interpreted our data as showing good agreement and ranging from moderate to excellent reliability, we would still encourage exercise physiologists and trainers to consider these data when interpreting whether real strength changes have occurred. Furthermore, we would like to remind the reader that these data are not assessment of the margin of error or reliability of the devices themselves, but rather these data represent agreement of participants performing repeated tests (i.e., that human variance plays the largest role in error).

CONCLUSION

The findings of this study indicate that both isometric and isokinetic upper- and lower-body multi-joint dynamometry using Exerbotics devices shows good agreement and typically good to excellent reliability in measuring muscular force. Future research should consider the specificity of dynamic training and whether muscle force increases are similar in both isokinetic and isometric testing after an intervention using a specific training modality (e.g., to confirm whether transfer of strength occurs between contraction types).

Contributions

All authors conceived the study. JPF and MN carried out the study. JS completed data analyses. All authors drafted the manuscript and approved the final version for submission.

Funding Information

No funding was obtained or used for this research project.

Data and Supplementary Material Accessibility

All data is available at the corresponding open science framework page: https://osf.io/d2kst.

References

Berchtold, A. (2016). Test–retest: Agreement or reliability? *Methodological Innovations*, *9*, 2059799116672875. https://doi.org/10.1177/2059799116672875

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet (London, England)*, *1*(8476), 307–310.

Bridgeman, L. A., McGuigan, M. R., Gill, N. D., & Dulson, D. K. (2016). Test-Retest Reliability of a Novel Isokinetic Squat Device With Strength-Trained Athletes. *Journal of Strength and Conditioning Research*, *30*(11), 3261–3265. https://doi.org/10.1519/JSC.0000000000001410

Callaghan, M. J., McCarthy, C. J., Al-Omar, A., & Oldham, J. A. (2000). The reproducibility of multi-joint isokinetic and isometric assessments in a healthy and patient population. *Clinical Biomechanics*, *15*(9), 678–683. https://doi.org/10.1016/S0268-0033(00)00032-2

Christensen, H. S., Borgbjerg, J., Børty, L., & Bøgsted, M. (2020). On Jones et al.'s method for extending Bland-Altman plots to limits of agreement with the mean for multiple observers. *BMC Medical Research Methodology*, *20*(1), 304. https://doi.org/10.1186/s12874-020-01182-w

Chtourou, H., Driss, T., Souissi, S., Gam, A., Chaouachi, A., & Souissi, N. (2012). The effect of strength training at the same time of the day on the diurnal fluctuations of muscular anaerobic performances. *Journal of Strength and Conditioning Research*, *26*(1), 217–225. https://doi.org/10.1519/JSC.0b013e31821d5e8d

Coudeyre, E., Jegu, A. G., Giustanini, M., Marrel, J. P., Edouard, P., & Pereira, B. (2016). Isokinetic muscle strengthening for knee osteoarthritis: A systematic review of randomized controlled trials with meta-analysis. *Annals of Physical and Rehabilitation Medicine*, *59*(3), 207–215. https://doi.org/10.1016/j.rehab.2016.01.013

Dirnberger, J., Huber, C., Hoop, D., Kösters, A., & Müller, E. (2013). Reproducibility of concentric and eccentric isokinetic multi-joint leg extension measurements using the IsoMed 2000-system. *Isokinetics and Exercise Science*, *21*(3), 195–202. https://doi.org/10.3233/IES-130511

Dvir, Z., & Müller, S. (2020). Multiple-Joint Isokinetic Dynamometry: A Critical Review. *Journal of Strength and Conditioning Research*, *34*(2), 587–601. https://doi.org/10.1519/JSC.0000000000002982

Engel, F. A., Faude, O., Kölling, S., Kellmann, M., & Donath, L. (2019). Verbal Encouragement and Between-Day Reliability During High-Intensity Functional Strength and Endurance Performance Testing. *Frontiers in Physiology*, *10*, 460. https://doi.org/10.3389/fphys.2019.00460

Hoffman, J. R., Ratamess, N. A., Kang, J., Gonzalez, A. M., Beller, N. A., & Craig, S. A. S. (2011). Effect of 15 days of betaine ingestion on concentric and eccentric force outputs during isokinetic exercise. *Journal of Strength and Conditioning Research*, *25*(8), 2235–2241. https://doi.org/10.1519/JSC.0b013e3182162530

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

Hopkins, W. G., Schabort, E. J., & Hawley, J. A. (2001). Reliability of power in physical performance tests. *Sports Medicine (Auckland, N.Z.)*, *31*(3), 211–234. https://doi.org/10.2165/00007256-200131030-00005

Jones, M., Dobson, A., & O'Brian, S. (2011). A graphical method for assessing agreement with the mean between multiple observers using continuous measures. *International Journal of Epidemiology*, *40*(5), 1308–1313. https://doi.org/10.1093/ije/dyr109

Knaier, R., Qian, J., Roth, R., Infanger, D., Notter, T., Wang, W., Cajochen, C., & Scheer, F. A. J. L. (2022). Diurnal Variation in Maximum Endurance and Maximum Strength Performance: A Systematic Review and Meta-analysis. *Medicine and Science in Sports and Exercise*, *54*(1), 169–180. https://doi.org/10.1249/MSS.0000000000002773

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kottner, J., & Streiner, D. L. (2011). The difference between reliability and agreement. *Journal of Clinical Epidemiology*, *64*(6), 701–702. https://doi.org/10.1016/j.jclinepi.2010.12.001

Kroemer, K. H. E., Marras, W. S., McGlothlin, J. D., McIntyre, D. R., & Nordin, M. (1990). On the measurement of human strength. *International Journal of Industrial Ergonomics*, *6*(3), 199–210. https://doi.org/10.1016/0169-8141(90)90034-Y

Kroll, W. (1970). Test Reliability and Errors of Measurement at Several Levels of Absolute Isometric Strength. *Research Quarterly. American Association for Health, Physical Education and Recreation*, *41*(2), 155–163. https://doi.org/10.1080/10671188.1970.10614966

Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective [Journal Article]. *Psychonomic Bulletin & Review*, *25*, 178–206.

Nuzzo, J. L., Pinto, M. D., Nosaka, K., & Steele, J. (2023). The Eccentric:Concentric Strength Ratio of Human Skeletal Muscle In Vivo: Meta-analysis of the Influences of Sex, Age, Joint Action, and Velocity. *Sports Medicine (Auckland, N.Z.)*, *53*(6), 1125–1136. https://doi.org/10.1007/s40279-023-01851-y

Nuzzo, J. L., Taylor, J. L., & Gandevia, S. C. (2019). CORP: Measurement of upper and lower limb muscle strength and voluntary activation. *Journal of Applied Physiology*, *126*(3), 513–543. https://doi.org/10.1152/japplphysiol.00569.2018

Paoli, A., Gentil, P., Moro, T., Marcolin, G., & Bianco, A. (2017). Resistance Training with Single vs. Multi-joint Exercises at Equal Total Load Volume: Effects on Body Composition, Cardiorespiratory Fitness, and Muscle Strength. *Frontiers in Physiology*, *8*, 1105. https://doi.org/10.3389/fphys.2017.01105

Parraca, J. A., Adsuar, J. C., Domínguez-Muñoz, F. J., Barrios-Fernandez, S., & Tomas-Carus, P. (2022). Test-Retest Reliability of Isokinetic Strength Measurements in Lower Limbs in Elderly. *Biology*, *11*(6), 802. https://doi.org/10.3390/biology11060802

Ratamess, N. A., Beller, N. A., Gonzalez, A. M., Spatz, G. E., Hoffman, J. R., Ross, R. E., Faigenbaum, A. D., & Kang, J. (2016). The Effects of Multiple-Joint Isokinetic Resistance Training on Maximal Isokinetic and Dynamic Muscle Strength and Local Muscular Endurance. *Journal of Sports Science & Medicine*, *15*(1), 34–40.

Rodriguez-Sanchez, F., cre, cph, Jackson, C. P., Hutchins, S. D., & Clawson, J. M. (2023). *Grateful: Facilitate Citation of R Packages*.

Schärer, C., Tacchelli, L., Göpfert, B., Gross, M., Lüthy, F., Taube, W., & Hübner, K. (2019). Specific Eccentric–Isokinetic Cluster Training Improves Static Strength Elements on Rings for Elite Gymnasts. *International Journal of Environmental Research and Public Health*, *16*(22), 4571. https://doi.org/10.3390/ijerph16224571

Schluter, P. J. (2009). A multivariate hierarchical Bayesian approach to measuring agreement in repeated measurement method comparison studies. *BMC Medical Research Methodology*, *9*(1), 6. https://doi.org/10.1186/1471-2288-9-6

Steele, J., Fisher, J. P., Smith, D., Schoenfeld, B. J., Yang, Y., & Nakagawa, S. (2023). Meta-analysis of variation in sport and exercise science: Examples of application within resistance training research. *Journal of Sports Sciences*, *41*(17), 1617–1634. https://doi.org/10.1080/02640414.2023.2286748

Stock, M. S., & Luera, M. J. (2014). Consistency of peak and mean concentric and eccentric force using a novel squat testing device. *Journal of Applied Biomechanics*, *30*(2), 322–325. https://doi.org/10.1123/jab.2013-0191

Vet, H. C. W. de, Terwee, C. B., Knol, D. L., & Bouter, L. M. (2006). When to use agreement versus reliability measures. *Journal of Clinical Epidemiology*, *59*(10), 1033–1039. https://doi.org/10.1016/j.jclinepi.2005.10.015

Wilson, G. J., & Murphy, A. J. (1996). The use of isometric tests of muscular function in athletic assessment. *Sports Medicine (Auckland, N.Z.)*, *22*(1), 19–37. https://doi.org/10.2165/00007256-199622010-00003