Physiological
Measurement

NOTE • **OPEN ACCESS**

Development, validation and test-retest reliability of a load cell-based device for assessment of isometric forearm rotation torque

To cite this article: Miika Köykkä *et al* 2025 *Physiol. Meas.* **46** 08NT01

IOP Publishing Physiol. Meas. 46 (2025) 08NT01 https://doi.org/10.1088/1361-6579/adf488

Physiological Measurement

IPEM Institute of Physics and Engineering in Medicine

OPEN ACCESS

RECEIVED
4 December 2024

REVISED
20 June 2025

ACCEPTED FOR PUBLICATION
25 July 2025

PUBLISHED
8 August 2025

NOTE

Development, validation and test-retest reliability of a load cell-based device for assessment of isometric forearm rotation torque

Miika Köykkä[1,2], **Iida Laatikainen-Raussi**[2], **Sami Vierola**[1], **Neil J Cronin**[2,4], **Benjamin Waller**[3,5,*] and **Tomi Vänttinen**[1]

[1] Finnish Institute of High Performance Sport KIHU, Jyväskylä, Finland
[2] Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland
[3] Physical Activity, Physical Education, Sport and Health Research Centre (PAPESH), Department of Sport Science, Reykjavik University, Reykjavik, Iceland
[4] School of Education and Science, University of Gloucestershire, Gloucestershire, United Kingdom
[5] NIHR Exeter Biomedical Research Centre (BRC), University of Exeter, Exeter, United Kingdom
* Author to whom any correspondence should be addressed.

E-mail: b.waller2@exeter.ac.uk

Keywords: forearm pronation, forearm supination, forearm rotation ratio, strength, rehabilitation

Supplementary material for this article is available online

Abstract

Objectives. This study aimed to develop and validate a load cell-based device for measuring isometric forearm rotation torque and to determine its test-retest reliability. *Approach.* The custom-built device was calibrated using known weights and validated against a high-precision torque transducer. For reliability assessment, 35 physically active participants (20 males, 15 females; age 30 ± 7 years) were tested for isometric forearm pronation and supination strength 5–7 d apart. *Main results.* The custom device demonstrated excellent validity (intraclass correlation coefficient (ICC), absolute agreement = 1.00; $r^2 = 1.00$, $p < 0.001$; mean difference = −1.26–1.44%, $p < 0.001$). Test-retest reliability was excellent for absolute pronation and supination torque (ICC = 0.88–0.97; coefficient of variation percentage (CV%) = 4.1–5.6; minimal detectable change (MDC) at 90% confidence level = 13.1–19.9%), good to excellent for supination:pronation ratios (ICC = 0.60–0.88; CV% = 7.0–8.6; MDC = 0.10–0.13), and fair to good for dominant:non-dominant ratios (ICC = 0.42–0.66; CV% = 6.1–7.6; MDC = 0.07–0.10). Sex significantly influenced absolute torque values, with males demonstrating consistently higher torque, although reliability metrics were similar for both sexes. *Significance.* The device is valid, and the test is reliable. It is suitable for clinical assessments, rehabilitation monitoring, and performance evaluation, facilitating an improved understanding of factors affecting elbow overloading and injuries. Limb ratio metrics should be interpreted with caution due to their lower reliability.

1. Introduction

Forearm pronation and supination are crucial for positioning the hand optimally for various tasks, such as providing rotational movements to objects grasped. The anatomical design of forearm pronator and supinator muscles, whose fibres wrap around the radius, allows the translation of muscle contractions into effective rotational movements and maintaining joint stability during various functional tasks, such as using a screwdriver, hammer, or throwing a ball (Soubeyrand *et al* 2017).

Forearm pronation relies mainly on the pronator teres and pronator quadratus muscles, while supination is largely controlled by the biceps brachii, assisted by the supinator muscle (O'Sullivan and Gallwey 2005, Bader *et al* 2018). During isometric tests, these muscles demonstrate different activity patterns depending on whether pronation or supination is being measured (Bader *et al* 2018). The grip strength required to hold the

dynamometer handle during these tests may influence muscle activity, but it also simulates real-life applications like throwing (Ferragut *et al* 2011), making it a practical addition to the performance evaluation (Gopinath *et al* 2017).

Musculotendinous units crossing the elbow joint provide dynamic stabilisation and protect the ligaments of the joint (Dugas *et al* 2014, Islam *et al* 2020), particularly the ulnar collateral ligament, which is often injured in overarm athletes (Zaremski *et al* 2017, DeFroda *et al* 2018). Thus, the force-generating capacity of these muscles is crucial for reducing stress on elbow ligaments and preventing elbow injuries in overarm throwing (Werner *et al* 1993, Fleisig *et al* 1995, Yanai *et al* 2023), and their strengthening is a key part of the rehabilitation process after an elbow injury (Wilk *et al* 2012). Further, forearm flexor–pronator muscle injuries have been observed to contribute to the failure of nonoperative management of ulnar collateral ligament tears (Ikezu *et al* 2022). Therefore, measuring forearm pronation and supination torque may be valuable for injury prevention, treatment, and rehabilitation, as well as performance optimisation. Specifically, torque measures can help identify muscular weaknesses or imbalances that predispose individuals to injury, guide rehabilitation protocols following injury, and provide baseline measures to monitor athletes or workers involved in repetitive or high-load forearm activities (Wolf *et al* 1987, Ellenbecker *et al* 2006, Morse *et al* 2006).

Literature is inconsistent in the used test positions and devices (Kotte *et al* 2018). Forearm torque-generating capacity has been measured using various custom-built dynamometers (Askew *et al* 1987, O'Sullivan and Gallwey 2005, Mukhopadhyay *et al* 2009, Herrera-Ligero *et al* 2023) and commercial systems such as the CON-TREX isokinetic dynamometer (Medimex, Sainte-Foy-les Lyon, France) (Daumillare *et al* 2020), BTE tools (BTE Technologies, Hanover, MD, USA) (Kramer *et al* 1994, Lucado *et al* 2019), and the Cybex dynamometers (Cybex, division of Lumex, Inc., Ronkonkoma, NY, USA) (Kramer *et al* 1994). While commercial isokinetic dynamometers are accurate and reliable (Bandy and McLaughlin 1993, Shechtman *et al* 2003), they tend to be expensive, large, and require in-depth expertise for reliable testing, which limits their practicality in typical clinical settings (Wong and Moskovitz 2010). Therefore, there is a need for a more practical and accessible alternative that can be custom-built by sports or health science laboratories with limited specialised equipment while still maintaining precision and reliability. This motivated the development of our custom device.

The first aim of this study was to validate a custom forearm isometric torque measurement device calibrated with a theoretical low-cost model. The second aim was to determine its test-retest reliability in physically active individuals without prior experience of the test.

2. Methods

2.1. Construction and calibration of the forearm rotation strength measurement device

The custom-built device consisted of a welded stainless-steel frame with an integrated gripping handle, rotational axis, and lever arm (figure 1). Two load cells (TAS606, SparkFun Electronics, Boulder, CO, USA) were attached at opposite ends of the lever. They could measure loads from 0 to 200 kg, with a maximal hysteresis of 0.1%, repeatability of 0.1% over the entire nominal range, and creep of 0.1%/30 min. Signals from the load cells were amplified via HX711-based amplifiers and 24-bit analogue-to-digital converter units (SEN-13879, SparkFun Electronics), and flowed to an ESP32-based microcontroller (WRL-17743, SparkFun Electronics). An OLED display integrated into the device showed real-time torque readings and indicated rotation direction (clockwise (CW) or counterclockwise (CCW)). During strength tests, the highest torque value was displayed until reset. See supplementary material S1 for detailed mechanical designs, electronic schematics, and source code.

At startup, the load cells were automatically tared three times during the first six seconds to ensure the electronic components had stabilised. Torque values were calculated in real-time at a sampling rate of 80 Hz. Each displayed torque was based on the mean of the four lowest readings from a 16-sample rolling buffer, minimising measurement errors from brief jerks.

Calibration involved applying known torques using weight sets (10–50 kg) suspended from a horizontally attached lever arm (47.5 mm from rotation centre) (figure 1). Each weight set was measured precisely using a weight scale (BWB-800, Tanita Corporation of America, Inc., Illinois, USA). Before hanging the weight set, the device was zeroed. Readings were taken after a 60 s stabilisation period. Linear calibration equations (one for each rotation direction) were derived from the recorded sensor outputs and theoretical torque values:

$$\text{counterclockwise } \tau = 0.0000317354270010086 \times \text{ADC} + 0.133360470185269 \quad (1)$$

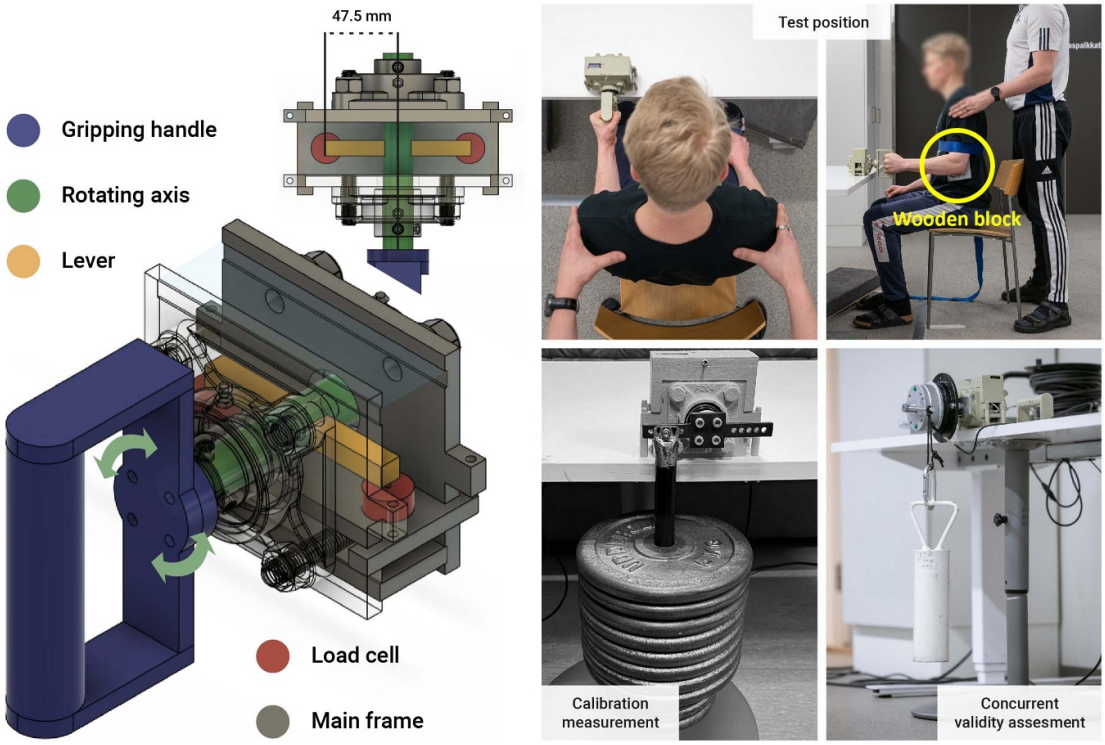

Figure 1. Illustration of the custom-made device, calibration measurement, concurrent validity assessment, and strength test position.

and

$$\text{clockwise } \tau = 0.0000308936870725094 \times \text{ADC} + 0.0857569175449125, \tag{2}$$

where τ corresponds to torque and ADC to analogue-to-digital converter output.

2.2. Assessment of concurrent validity of the forearm rotation strength device

The concurrent validity of the custom device was assessed against a high-precision torque transducer to confirm measurement accuracy post-calibration, ensuring the device's practical precision matched its theoretical calibration. Although isokinetic dynamometers are considered the gold standard, our goal was specifically to assess the practical precision of our calibration method against a reference device that directly measures torque. A FINAS-accredited calibration laboratory (JAMK University of Applied Sciences, School of Technology, Calibration Centre, Finland) conducted the measurements to assess concurrent validity.

The gripping handle of the custom-made device was replaced by a metal plate allowing attachment of the device's rotating axis to the rotating axis of a high-precision torque transducer (Stahlwille 7722, STAHLWILLE Tools LLC, Sturtevant, WI, USA) with a measurement range from 2 to 100 Nm (figure 1), which acted as the reference device. To allow for real-time readings from the custom-made device, the signal processing pipeline on the microcontroller was adjusted to display current values on the OLED display instead of the maximum value, as it does in strength testing mode. The measurement protocol included hanging a set of weights from a lever arm attached to the torque transducer, thereby inducing the same torque to the custom and reference devices. Four different weights were used three times to induce torque in each direction so that the recorded torques ranged from approximately 6–26 Nm, corresponding to the range of forearm pronation and supination torques previously observed by the authors with the custom device among high-level male and female javelin throwers and volleyball players. The readings of each device were recorded 60 s after hanging the weight to ensure that the weight had stopped swinging.

2.3. Test-retest reliability measurements

Thirty-five physically active participants, males ($n = 20$, age $= 29.7 \pm 7.6$ years, height $= 1.81 \pm 0.05$ m, weight $= 84.7 \pm 9.9$ kg) and females ($n = 15$, age $= 29.6 \pm 6.1$ years, height $= 1.69 \pm 0.05$ m, weight $= 66.5 \pm 10.8$ kg) without prior experience of the test protocol participated in the test-retest study. Participants attended two identically structured testing sessions, scheduled 5–7 d apart and at the same time of day. They received comprehensive information about the study design and potential risks prior to

providing written consent. The study was conducted according to the principles of the Declaration of Helsinki and received approval from the University of Jyväskylä Ethical Committee (1635/13.00.04.00/2022). Participants were healthy and injury-free at the time of testing.

Warm-up consisted of five minutes of light aerobic exercise on a stationary rowing ergometer at a rating of 10–12 on Borg's 15-point rating of perceived exertion (6–20) RPE scale (Borg 1998) and submaximal strength exercises (ten repetitions of body-weight push ups, dead rows with dumbbells and shoulder presses with dumbbells). Subsequently, the testing position (figure 1), similar to that used in the study by Axelsson and Kärrholm (2018), was adjusted. Participants sat upright with their upper body unsupported, feet on the ground, the testing arm perpendicular to the device handle, shoulder in 0° flexion and 0° abduction (i.e. vertical), elbow in 90° flexion supported by a wooden block placed on the side, and wrist in neutral position. The wooden block's width was individually adjusted to make sure the shoulder position was correct. A strap was secured around the upper body and the upper arm to ensure that the shoulder position was maintained. In the test, participants performed two submaximal practice trials at 60% and 80% of their perceived maximal strength, respectively, with 15 s rest intervals, and two maximal trials with 60 s rest intervals. If an improvement greater than 5% was observed, additional trials were performed to make sure maximum was reached. The highest value was used for analyses. Verbal cues ('Ready-3-2-1-go') and consistent encouragement were provided for each repetition. At '1', participants first gradually built tension on the force transducer, and at 'go' they exerted maximal force for approximately three seconds. The clinician monitored the specified body positions to make sure no compensatory and unwanted movements occurred during any trial, while the goal was to perform any necessary adjustments and correct false movements already before and during the practice trials. Forearm pronation test was performed first, followed by forearm supination test with the same limb. This procedure was then repeated with the contralateral limb. The participant's posture was aligned for each limb separately by moving the chair on which they were seated.

2.4. Sample size

The sample size for the test-retest protocol was determined based on the null hypothesis *'The ICC will not be higher than 0.85'*, with an expected intra-class correlation coefficient of 0.95, according to results from a previous study (Axelsson and Kärrholm 2018). Using a two-times test-retest approach, the required sample size for 80% power would be 25 participants (Borg *et al* 2022).

2.5. Statistical analysis

All analyses were performed in R version 4.4.3 (R Core Team 2024). The alpha level of 0.05 was applied for all analyses.

Concurrent validity was assessed by comparing the mean absolute and percentage difference (bias) between the devices using paired samples *t*-tests, Bland–Altman plots with 95% limits of agreement, intraclass correlation coefficient model 2 (ICC; calculated for absolute agreement; using psych package version 2.4.6.26 (Revelle 2024)), and Pearson's correlation coefficient. Test-retest reliability was evaluated for both sexes separately and the entire group by Bland–Altman plots with 95% limits of agreement, ICC, standard error of the measurement (SEM), CV%, and MDC. To further assess both within-subject changes across sessions and between-group sex differences, a two-way repeated-measures ANOVA with factors Time (Test, Retest) and Sex (male, female) was performed using the afex package version 1.4-1 (Singmann *et al* 2024), with the main interest on Time effects and Sex:Time interaction effects. Partial eta squared (η_p^2) determined the variance explained by Sex, Time, and their interaction, and were computed with effectsize version 1.0.0 (Ben-Shachar *et al* 2020). Any significant main or interaction effects were followed up by Bonferroni-adjusted pairwise comparisons of estimated marginal means via emmeans version 1.10.5 (Lenth 2024).

ICCs were used to categorise agreement as poor (<0.40), fair (0.40 to <0.60), good (0.60–0.75) or excellent (⩾0.75) (Cicchetti 1994). Pearson's *r* values were squared, with effect sizes interpreted as small ($r^2 \geqslant 0.01$), medium (⩾0.09) or large (⩾0.25) (Cohen 1992). SEM was calculated using the formula: $\text{SEM} = \text{total SD}\sqrt{1 - \text{ICC}}$, where SD represents standard deviation. CV% was determined by first calculating the ratio of the standard deviation to the mean score for each participant's two test scores. These individual CVs were then averaged to obtain the overall CV% (Atkinson and Nevill 1998). MDC at 90% confidence level was calculated using the formula: $\text{MDC} = 1.65 \times \sqrt{2} \times \text{SEM}$.

3. Results

The differences between the custom device and the reference device are shown in table 1. Mean differences were significant in both clockwise and counterclockwise directions, with mean bias values of 0.21 Nm (1.4%) and −0.16 Nm (−1.3%), respectively (both $p < 0.001$). This is also evident from the Bland–Altman plots in

Table 1. Concurrent validity statistics.

	Mean (95% CI) absolute difference	Mean (95% CI) percentage difference	*t* statistic	*p* value	ICC (95% CI)
Clockwise (Nm)	0.21 (0.16–0.27)	1.44% (1.25%–1.62%)	8.60	<0.001	1.00 (1.00–1.00)
Counter clockwise (Nm)	−0.16 (−0.19−0.13)	−1.26% (−1.71%−0.81%)	−10.87	<0.001	1.00 (1.00–1.00)

95% CI = 95% confidence interval; ICC = intraclass correlation coefficient.

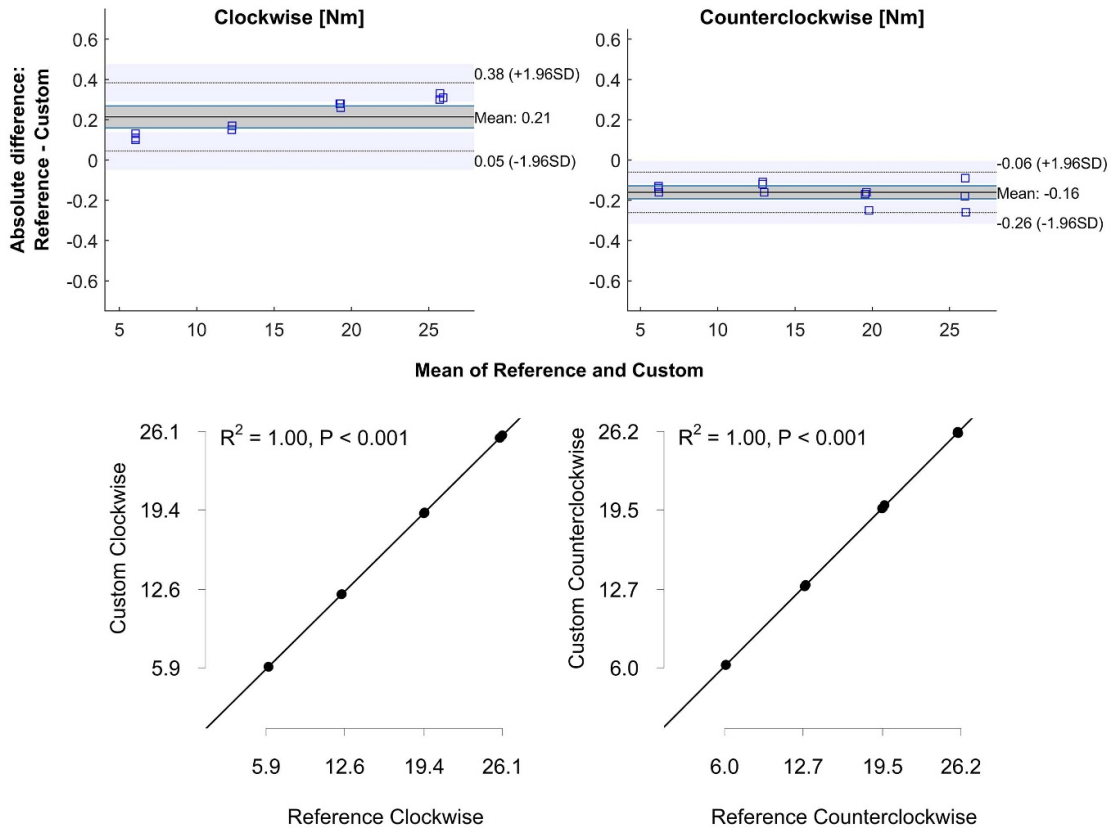

Figure 2. Bland–Altman (top panel) and correlation (bottom panel) plots comparing the reference device with the custom device. In the Bland–Altman plots, the dotted horizontal lines indicate the upper and lower limits of agreement, with the shaded regions representing their 95% confidence intervals. The shaded area around the mean denotes the confidence interval for the mean difference between the devices.

figure 2, where the confidence interval of the mean difference between the methods did not overlap zero. However, the estimates from the two devices showed excellent absolute agreement (table 1), along with correlations representing large effect sizes (figure 2).

In both sexes and the entire group, the isometric forearm rotation torque test indicated excellent test-retest reliability for all absolute torque measures and the supination to pronation ratio for both arms, whereas the dominant to non-dominant arm ratio in supination and pronation exhibited fair to good reliability (figure 3 and table 2).

None of the measures showed significant Time effects ($p = 0.20$–0.96) or Sex:Time interaction effects ($p = 0.17$–0.76). Sex effects were significant for pronation (dominant arm: $F[1, 33] = 32.86$, $\eta_p^2 = 0.50$, $P < 0.001$; non-dominant arm: $F[1, 33] = 28.04$, $\eta_p^2 = 0.46$, $P < 0.001$) and supination (dominant arm: $F[1, 33] = 25.68$, $\eta_p^2 = 0.44$, $P < 0.001$; non-dominant arm: $F[1, 33] = 33.32$, $\eta_p^2 = 0.50$, $P < 0.001$) torque, but not for ratio metrics ($p = 0.09$–1.0). For pronation, males exceeded females by 4.42 Nm on the dominant arm ($t[33] = 5.73$, $p < 0.001$) and by 4.23 Nm on the non-dominant arm ($t[33] = 5.29$, $p < 0.001$) and. For supination, the male–female difference was 5.07 Nm on the dominant side ($t[33] = 5.07$, $p < 0.001$) and 5.47 Nm on the non-dominant side ($t[33] = 5.77$, $p < 0.001$).

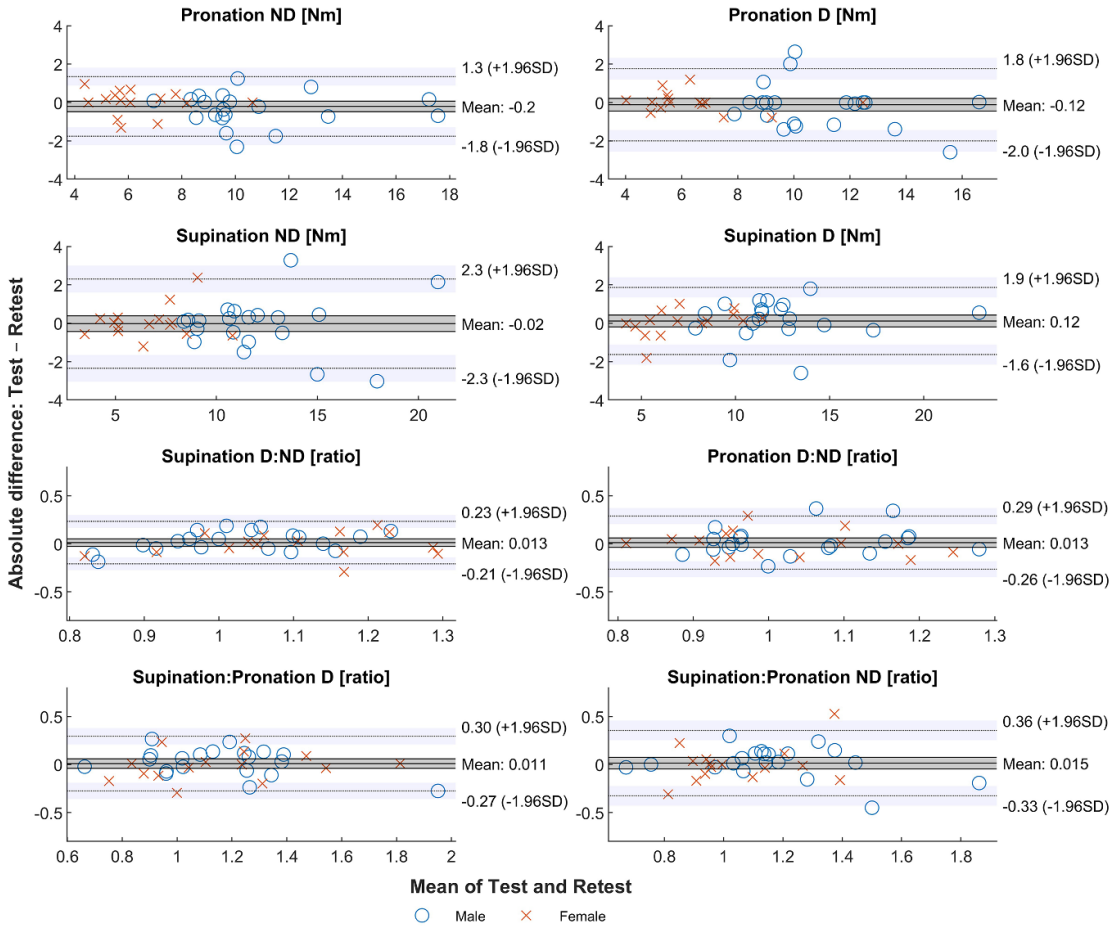

Figure 3. Bland–Altman plots comparing test and retest values based on the highest recorded values. The dotted horizontal lines indicate the upper and lower limits of agreement, with the shaded regions representing their 95% confidence intervals. The shaded area around the mean denotes the confidence interval for the mean difference between test and retest values. D = dominant arm; ND = non-dominant arm.

4. Discussion

In this study, concurrent validity and test-retest reliability of a custom load cell-based device for assessing isometric forearm rotation torque were evaluated. The custom device demonstrated excellent concurrent validity in both rotational directions. Test-retest reliability was excellent for the absolute measures of forearm pronation and supination torque for both the dominant and non-dominant arms, as well as for supination to pronation ratios in both arms. Dominant to non-dominant arm ratios in supination and pronation showed fair and good reliability, respectively.

4.1. Concurrent validity

The validity statistics (table 1 and figure 2) show that the custom device, calibrated with the theoretical model, is precise within the measured range. The minor differences observed might be due to errors such as the length of the rotating lever arm not being exactly 95 mm, imperfect centring of the lever arm through the rotating axis, or inaccuracies during the calibration process, such as errors in measuring the weight sets. For example, the difference between the reference and custom device in the clockwise direction increased a total of ∼0.2 Nm when the torque was increased from ∼6 to ∼26 Nm (figure 2). In practice, this corresponds to an error of 1 mm in lever arm length. Although the differences between the devices were significant, they remain clinically negligible (<1.5%), and only slightly exceeded the ±1% error reported by the manufacturer for the high precision torque transducer.

To the authors' knowledge, no similar concurrent validity studies exist for comparison. This study investigated the device's technical validity for torque measurement by comparing its estimates with a high precision torque transducer. In contrast, previous studies had participants perform the same tests on two different forearm torque measurement devices (Wolf *et al* 1987, Wong and Moskovitz 2010, Axelsson and Kärrholm 2018), and therefore their results are not comparable.

Table 2. Test-retest statistics for each test measure based on the highest recorded values.

		Test mean (SD)	Retest mean (SD)	ICC (95% CI)	SEM	MDC	CV%
Pronation D (Nm)	All	8.92 (3.04)	9.03 (3.32)	0.96 (0.91–0.98)	0.67	1.56 (17.5%)	4.8
	M	10.76 (2.27)	10.98 (2.62)	0.88 (0.74–0.95)	0.82	1.92 (17.8%)	5.3
	F	6.46 (2.04)	6.43 (2.17)	0.97 (0.91–0.99)	0.36	0.84 (13.1%)	4.1
Pronation ND (Nm)	All	8.68 (3.08)	8.88 (3.24)	0.97 (0.94–0.98)	0.57	1.32 (15.2%)	5.1
	M	10.41 (2.78)	10.77 (2.80)	0.95 (0.87–0.98)	0.64	1.49 (14.3%)	4.7
	F	6.36 (1.61)	6.35 (1.68)	0.93 (0.80–0.97)	0.44	1.02 (16.1%)	5.7
Supination D (Nm)	All	10.24 (3.94)	10.12 (3.80)	0.97 (0.95–0.99)	0.62	1.45 (14.2%)	4.9
	M	12.44 (3.37)	12.26 (3.33)	0.95 (0.89–0.98)	0.72	1.67 (13.4%)	4.7
	F	7.30 (2.48)	7.27 (2.18)	0.96 (0.89–0.99)	0.46	1.07 (14.7%)	5.1
Supination ND (Nm)	All	9.77 (3.93)	9.79 (3.91)	0.96 (0.91–0.98)	0.82	1.92 (19.6%)	5.3
	M	12.09 (3.32)	12.16 (3.28)	0.91 (0.80–0.96)	0.96	2.24 (18.6%)	5.1
	F	6.68 (2.14)	6.63 (1.95)	0.92 (0.78–0.97)	0.57	1.33 (19.9%)	5.6
Pronation D:ND (ratio)	All	1.03 (0.14)	1.02 (0.14)	0.47 (0.18–0.69)	0.10	0.23 (22.2%)	7.2
	M	1.05 (0.14)	1.03 (0.12)	0.42 (−0.01–0.72)	0.10	0.24 (22.4%)	6.9
	F	1.01 (0.13)	1.01 (0.15)	0.55 (0.09–0.82)	0.09	0.22 (21.8%)	7.6
Supination D:ND (ratio)	All	1.06 (0.15)	1.05 (0.13)	0.67 (0.44–0.82)	0.08	0.18 (17.2%)	6.2
	M	1.04 (0.14)	1.01 (0.10)	0.65 (0.30–0.84)	0.07	0.17 (16.3%)	6.1
	F	1.10 (0.16)	1.10 (0.14)	0.66 (0.26–0.87)	0.09	0.20 (18.1%)	6.4
Supination: Pronation D (ratio)	All	1.16 (0.28)	1.15 (0.29)	0.87 (0.77–0.93)	0.10	0.24 (20.2%)	7.3
	M	1.17 (0.25)	1.14 (0.30)	0.88 (0.72–0.95)	0.10	0.23 (19.3%)	7.0
	F	1.15 (0.32)	1.16 (0.29)	0.88 (0.68–0.96)	0.10	0.24 (21.1%)	7.7
Supination: Pronation ND (ratio)	All	1.13 (0.25)	1.11 (0.26)	0.76 (0.58–0.87)	0.12	0.28 (25.1%)	7.5
	M	1.18 (0.24)	1.16 (0.30)	0.83 (0.62–0.93)	0.11	0.26 (22.1%)	6.7
	F	1.05 (0.24)	1.05 (0.18)	0.60 (0.16–0.84)	0.13	0.31 (29.1%)	8.6

SD = standard deviation; 95% CI = 95% confidence interval; D = dominant arm; ND = non-dominant arm; ICC = intraclass correlation coefficient; SEM = standard error of the measurement; MDC = minimal detectable change at 90% confidence level; CV% coefficient of variation percentage.

4.2. Test-retest reliability

The custom device demonstrated excellent test-retest reliability for absolute measures of forearm pronation and supination torque in both dominant and non-dominant arms, consistent across the entire group (ICC = 0.96–0.97). Subgroup analyses indicated that males and females separately also showed excellent reliability, with ICC values ranging from 0.88 to 0.97 for males and 0.92–0.97 for females. (table 2) These findings indicate robust measurement consistency suitable for clinical and athletic evaluations and align with previous studies that have also reported ICCs over 0.90 while using a similar testing position (Axelsson and Kärrholm 2018, Gwak *et al* 2021, Herrera-Ligero *et al* 2023). However, Herrera-Ligero *et al* (2023) found slightly lower ICCs, particularly for supination which may be attributed to using only one maximal trial.

MDCs for absolute torque measures ranged between 0.84 and 1.92 Nm (13.1%–19.9%). The relative values are comparable with those reported for other upper extremity isometric strength tests, such as shoulder internal and external rotation (Cools *et al* 2014) and extension (Awatani *et al* 2016), measured using a handheld dynamometer, and the ASH-test (a long lever shoulder strength test) measured using a force platform (Ashworth *et al* 2018). This similarity supports the clinical relevance and applicability of the device for reliably monitoring forearm strength changes over time. However, other studies addressing forearm pronation and supination strength have not reported this metric.

Reliability of the supination to pronation ratio in both arms was excellent (ICC > 0.75). To the authors' knowledge, there are no studies available for comparison regarding forearm strength, but the values are in line with the shoulder external/internal rotation strength ratio obtained using a handheld dynamometer (Conceição *et al* 2018). In the overarm thrower's shoulder, evidence indicates that too low external compared to internal rotator strength reduces training load tolerance (Møller *et al* 2017) and is a significant injury risk factor (Whiteley *et al* 2012, Clarsen *et al* 2014). Therefore, studying the sensitivity of the balance between forearm supinators' and pronators' strength to indicate training load tolerance and injury risks in the future could provide valuable insights to enhance performance in sports involving high-intensity upper extremity

activities, as well as in occupations requiring repetitive and high-load arm movements. However, MDC values should be carefully considered when interpreting changes in a monitoring context.

The considerably lower reliability of the dominant to non-dominant arm ratios in forearm pronation (ICC 0.47) and supination (ICC 0.67) observed in the present study is consistent with previous results (Kramer *et al* 1994). However, the dominant arm has been observed to demonstrate over 60% higher pronation torque than the non-dominant arm in healthy tennis players (Ellenbecker *et al* 2006), a difference much greater than in healthy general population, where it is less than 10% (Kotte *et al* 2018). Despite the lack of reproducibility in dominant to non-dominant ratios, likely due to small variations in the absolute measures, individuals repeatedly exposed to unilateral upper limb tasks requiring high intensity activities may still benefit from ensuring that their dominant arm can produce higher forearm rotation torques than the non-dominant arm. However, it may be reasonable to establish task-specific normative values for dominant to non-dominant ratios of forearm rotation strength in future studies, which have often not been reported.

Sex comparisons revealed consistently reliable torque measurements across both males and females, although males produced higher torques in both pronation and supination tasks (table 2). However, reliability metrics did not differ meaningfully between sexes, confirming that the custom device provides robust measurements for both males and females alike. Nonetheless, establishing sex-specific normative values could further enhance the utility of the device in clinical and athletic contexts.

Previous research indicates that torque output fluctuates depending on body and joint positions, emphasising the importance of maintaining consistent positioning to achieve reproducible results (O'Sullivan and Gallwey 2005, Morse *et al* 2006, Mukhopadhyay *et al* 2009, Herrera-Ligero *et al* 2023). Any changes in posture may therefore have contributed to the observed test-retest differences. Some participants tended to compensate by tilting their upper bodies mediolaterally in the direction of forearm rotation, flexing, extending or deviating the wrist, or by abducting or adducting the shoulder. This was counteracted by the clinician conducting the test, who aimed to teach proper technique during the warm-up trials and excluded trials performed with incorrect technique. However, day-to-day fluctuations in performance are natural due to human factors.

4.3. Clinical implications and future studies

A forearm isometric torque measurement device can be built and calibrated with precision comparable to a high-precision torque transducer using the protocols demonstrated in this study. The test can assess forearm muscle performance and evaluate an athlete's or worker's readiness for repetitive or high-load arm activities. It could potentially form part of a comprehensive assessment protocol for the entire upper extremity. Future studies should investigate the test's sensitivity to cumulative load, its ability to predict upper extremity injury risk, and its responsiveness to rehabilitation interventions following injuries. Additionally, establishing sport-, task- and sex-specific normative values among healthy individuals may help identify those at risk of developing upper extremity injuries.

4.4. Limitations

The test-retest study was conducted in participants who had no prior experience with the test and included only two repeated measurements. Therefore, conclusions about the long-term repeatability of the measurements or possible effects of learning cannot be drawn from this study. However, the test protocol was designed to allow participants to familiarise themselves during the submaximal warm-up trials, as it is common in clinical practice that individuals are unfamiliar with the test during their initial visit. Displaying the torque signal on the screen could enhance learning, as it would allow the participant and the clinician to see how the torque is produced. This, however, would require a different type of display.

In the present study, the effects of long-term use on the validity of the custom device were not investigated. Therefore, it is recommended to re-calibrate the device once a year (Khodabocus and Balgobin 2011), as factors such as bearing friction may change over time. As demonstrated, the theoretical calibration method is valid and can be used to ensure measurement accuracy.

5. Conclusion

The custom load cell-based device is a valid and reliable tool for assessing absolute forearm pronation and supination torque. It is suitable for monitoring absolute torque values, but it may not be reproducible enough for monitoring dominant to non-dominant arm ratios and supination to pronation ratios.

Data availability statement

The data cannot be made publicly available upon publication because they are not available in a format that is sufficiently accessible or reusable by other researchers. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

We thank Kari Koivisto from Vemitek Oy (Peräseinäjoki, Finland) for his assistance in manufacturing the metal parts of the custom-made device, Chuan-Chih Chiang for her assistance in the calibration measurements, Petri Uutela and Juha Roivainen in concurrent validity assessment, and Juuso Sillanpää, Eero Viitikko, Lauri Hietaharju and Tommi Jalomäki for their assistance in planning and piloting the strength measurement protocol. We would also like to thank the participants for their time and effort. The authors have confirmed that any identifiable participants in this study have given their consent for publication.

Contributor roles

Conceptualization: all authors. Data curation: MK. Formal analysis: MK, IL-R, NJC, BW. Funding acquisition: NJC, TV. Investigation: MK, IL-R. Methodology: all authors. Project administration: MK, IL-R. Resources: MK, SV, TV. Software: MK, SV. Supervision: NJC, BW, TV. Validation: all authors. Visualisation: MK. Writing—original draft: MK, IL-R. Writing—review & editing: all authors.

Funding

The study was done as part of the Sport Innovation Hub Jyväskylä project which is co-funded by the European Union (European Regional Development Fund, ERDF; project A80754). The funder had no influence on experiment design or data analysis.

Conflict of interest

None.

Patient consent for publication

Individuals recognisable in the photographs in this article have provided written consent for the publication of their images.

Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Ethical statement

Ethical approval was obtained from the University of Jyväskylä Ethics Committee (1635/13.00.04.00/2022). Participants gave informed consent to participate in the study before taking part. The research was conducted in accordance with the principles embodied in the Declaration of Helsinki and in accordance with local statutory requirements.

ORCID iDs

Miika Köykkä 0000-0003-0451-9099
Iida Laatikainen-Raussi 0009-0008-9258-4010
Neil J Cronin 0000-0002-5332-1188
Benjamin Waller 0000-0002-0738-0670
Tomi Vänttinen 0009-0007-5813-6583

References

Ashworth B, Hogben P, Singh N, Tulloch L and Cohen D D 2018 The Athletic Shoulder (ASH) test: reliability of a novel upper body isometric strength test in elite rugby players *BMJ Open Sport Exercise Med.* **4** e000365

Askew L J, An K N, Morrey B F and S. Chao E Y 1987 Isometric elbow strength in normal individuals *Clin. Orthopaedics Relat. Res.* **222** 261–6
Atkinson G and Nevill A M 1998 Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine *Sports Med.* **26** 217–38
Awatani T, Morikita I, Shinohara J, Mori S, Nariai M, Tatsumi Y, Nagata A and Koshiba H 2016 Intra- and inter-rater reliability of isometric shoulder extensor and internal rotator strength measurements performed using a hand-held dynamometer *J. Phys. Ther. Sci.* **28** 3054–9
Axelsson P and Kärrholm J 2018 New methods to assess forearm torque and lifting strength: reliability and validity *J. Hand Surg.* **43** 865.e1–865.e10
Bader J, Boland M R, Greybe D, Nitz A, Uhl T and Pienkowski D 2018 Muscle activity during maximal isometric forearm rotation using a power grip *J. Biomech.* **68** 24–32
Bandy W D and McLaughlin S 1993 Intramachine and intermachine reliability for selected dynamic muscle performance tests *J. Orthop. Sports Phys. Ther.* **18** 609–13
Ben-Shachar M S, Lüdecke D and Makowski D 2020 Effectsize: estimation of effect size indices and standardized parameters *J. Open Source Softw.* **5** 2815
Borg D N, Bach A J E, O'Brien J L and Sainani K L 2022 Calculating sample size for reliability studies *Pmr* **14** 1018–25
Borg G 1998 *Borg's Perceived Exertion and Pan Scales* (Human Kinetics)
Cicchetti D V 1994 Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology *Psychol. Assess.* **6** 284–90
Clarsen B, Bahr R, Andersson S H, Munk R and Myklebust G 2014 Reduced glenohumeral rotation, external rotation weakness and scapular dyskinesis are risk factors for shoulder injuries among elite male handball players: a prospective cohort study *Br. J. Sports Med.* **48** 1327–33
Cohen J 1992 A power primer *Psychol. Bull.* **112** 155–9
Conceição A, Parraca J, Marinho D, Costa M, Louro H, Silva A and Batalha N 2018 Assessment of isometric strength of the shoulder rotators in swimmers using a handheld dynamometer: a reliability study *Acta Bioeng. Biomech.* **20** 113–9
Cools A M, De Wilde L, Van Tongel A, Ceyssens C, Ryckewaert R and Cambier D C 2014 Measuring shoulder external and internal rotation strength and range of motion: comprehensive intra-rater and inter-rater reliability study of several testing protocols *J. Shoulder Elbow Surg.* **23** 1454–61
Daumillare A, Bosch C, Julia M, Coulet B, Mares O and Chammas M 2020 Volar locking plate fixation of distal radius fractures: isokinetic assessment of the influence of an ulnar styloid process fracture on pronation-supination strength *Hand Surg. Rehabil.* **39** 23–29
DeFroda S F, Goodman A D, Gil J A and Owens B D 2018 Epidemiology of elbow ulnar collateral ligament injuries among baseball players: national collegiate athletic association injury surveillance program, 2009-2010 through 2013-2014 *Am. J. Sports Med.* **46** 2142–7
Dugas J, Chronister J, Cain E L and Andrews J R 2014 Ulnar collateral ligament in the overhead athlete *Sports Med. Arthroscopy Rev.* **22** 169–82
Ellenbecker T S, Roetert E P and Riewald S 2006 Isokinetic profile of wrist and forearm strength in elite female junior tennis players *Br. J. Sports Med.* **40** 411–4
Ferragut C, Vila H, Abraldes J A, Argudo F, Rodriguez N and Alcaraz P E 2011 Relationship among maximal grip, throwing velocity and anthropometric parameters in elite water polo players *J. Sports Med. Phys. Fitness* **51** 26–32
Fleisig G S, Andrews J R, Dillman C J and Escamilla R F 1995 Kinetics of baseball pitching with implications about injury mechanisms *Am. J. Sports Med.* **23** 233–9
Gopinath B, Kifley A, Liew G and Mitchell P 2017 Handgrip strength and its association with functional independence, depressive symptoms and quality of life in older adults *Maturitas* **106** 92–94
Gwak G-T, Hwang U-J, Jung S-H, Kim J-H, Kim M-H and Kwon O-Y 2021 Comparison of maximal isometric forearm supination torque in two elbow positions between subjects with and without limited forearm supination range of motion *Physiother. Theory Pract.* **37** 99–105
Herrera-Ligero C, Sánchez Zuriaga D, Martínez-Iranzo Ú, Corresa S P and De Rosario H 2023 Protocol for forearm pronosupination strength measuring in different postures: reliability and influence of relevant factors *Am. J. Occup. Ther.* **77** 7705205110
Ikezu M *et al* 2022 Sites of flexor-pronator muscle injury and relationship between ulnar collateral ligament injury and flexor-pronator muscle injury in baseball players: a retrospective cohort study *J. Shoulder Elbow Surg.* **31** 1588–94
Islam S U, Glover A, MacFarlane R J, Mehta N and Waseem M 2020 The anatomy and biomechanics of the elbow *Open Orthopaedics J.* **14** 95–99
Khodabocus F and Balgobin K 2011 Implementation and practical benefits of ISO/IEC 17025:2005 in a testing laboratory *Univ. Mauritius Res. J.* **17** 27–60
Kotte S H P, Viveen J, Koenraadt K L M, The B and Eygendaal D 2018 Normative values of isometric elbow strength in healthy adults: a systematic review *Shoulder Elbow* **10** 207–15
Kramer J F, Nusca D, Bisbee L, MacDermid J, Kemp D and Boley S 1994 Forearm pronation and supination: reliability of absolute torques and nondominant/dominant ratios *J. Hand Ther.* **7** 15–20
Lenth R V 2024 Emmeans: estimated marginal means, aka least-squares means (available at: https://CRAN.R-project.org/package=emmeans)
Lucado A, Fraher L, Patel H and Munck G 2019 Comparison of portable handheld versus fixed isokinetic dynamometers in measuring strength of the wrist and forearm *Physiother. Theory Pract.* **35** 677–85
Morse J L, Jung M-C, Bashford G R and Hallbeck M S 2006 Maximal dynamic grip force and wrist torque: the effects of gender, exertion direction, angular velocity, and wrist angle *Appl. Ergon.* **37** 737–42
Mukhopadhyay P, O'Sullivan L W and Gallwey T J 2009 Upper limb discomfort profile due to intermittent isometric pronation torque at different postural combinations of the shoulder-arm system *Ergonomics* **52** 584–600
Møller M, Nielsen R O, Attermann J, Wedderkopp N, Lind M, Sørensen H and Myklebust G 2017 Handball load and shoulder injury rate: a 31-week cohort study of 679 elite youth handball players *Br. J. Sports Med.* **51** 231–7
O'Sullivan L and Gallwey T 2005 Forearm torque strengths and discomfort profiles in pronation and supination *Ergonomics* **48** 703–21
R Core Team 2024 R: a language and environment for statistical computing (available at: www.R-project.org/)
Revelle W 2024 Psych: procedures for psychological, psychometric, and personality research (available at: https://CRAN.R-project.org/package=psych)

Shechtman O, Davenport R, Malcolm M and Nabavi D 2003 Reliability and validity of the BTE-Primus grip tool *J. Hand Ther.* **16** 36–42
Singmann H *et al* 2024 Afex: analysis of factorial experiments (available at: https://CRAN.R-project.org/package=afex)
Soubeyrand M, Assabah B, Bégin M, Laemmel E, Dos Santos A and Crézé M 2017 Pronation and supination of the hand: anatomy and biomechanics *Hand Surg. Rehabil.* **36** 2–11
Werner S L, Fleisig G S, Dillman C J and Andrews J R 1993 Biomechanics of the elbow during baseball pitching *J. Orthop. Sports Phys. Ther.* **17** 274–8
Whiteley R, Oceguera M V, Valencia E B and Mitchell T 2012 Adaptations at the shoulder of the throwing athlete and implications for the clinician *Tech. Shoulder Elbow Surg.* **13** 36–44
Wilk K E, Macrina L C, Cain E L, Dugas J R and Andrews J R 2012 Rehabilitation of the overhead athlete's elbow *Sports Health* **4** 404–14
Wolf L D, Klein L and Cauldwell-Klein E 1987 Comparison of torque strength measurements on two evaluation devices *J. Hand Ther.* **1** 24–27
Wong C K and Moskovitz N 2010 New assessment of forearm strength: reliability and validity *Am. J. Occup. Ther.* **64** 809–13
Yanai T, Onuma K, Crotin R L and Monda D 2023 A novel method intersecting three-dimensional motion capture and medial elbow strength dynamometry to assess elbow injury risk in baseball pitchers *Sci. Rep.* **13** 12253
Zaremski J L, McClelland J, Vincent H K and Horodyski M 2017 Trends in sports-related elbow ulnar collateral ligament injuries *Orthopaedic J. Sports Med.* **5** 232596711773129