

Original Research

Reliability and Validity of a Crane Scale for Isometric Knee and Shoulder Strength Assessment

Jon Dignazio¹, Hannah Marchant², Brady Hutchinson³, Adam Popchak⁴, Andrew L Sprague^{4a}

¹ Department of Athletic Training, University of Pittsburgh, ² Sports Physical Therapy Residency Program, Cleveland Clinic, ³ OrthoVirginia, ⁴ Department of Physical Therapy, University of Pittsburgh

Keywords: dynamometer, muscle, rehabilitation, torque, strain gauge, force

<https://doi.org/10.26603/001c.132166>

International Journal of Sports Physical Therapy

Vol. 20, Issue 4, 2025

Background

Due to cost barriers and/or space constraints, electromechanical and hand-held dynamometry are underutilized in rehabilitation settings. As a result, clinicians often rely on semi-quantitative methods that may fail to detect deficits in upper or lower extremity strength. The purpose of this study was to examine the between-trials test-retest reliability and concurrent criterion validity of an accessible crane scale for measuring isometric knee and shoulder strength.

Study Design

Observational cohort, Cross-sectional study

Methods

Twenty healthy, recreationally active adults underwent isometric knee and shoulder strength testing using a crane scale, electromechanical dynamometer, and hand-held dynamometer during a single session. Knee extension and flexion, and shoulder internal rotation, external rotation, and abduction were tested. Three maximal effort trials were performed per device and motion, with the order of device testing randomized. Between-trials test-retest reliability of the crane scale was assessed using intraclass correlation coefficients (ICCs). Concurrent criterion validity was examined using ICCs, absolute and relative error, Bland-Altman plots, and simple linear regression.

Results

Test-retest reliability of the crane scale was excellent for all motions (ICCs ≥ 0.90). Concurrent validity was poor to moderate with the electromechanical dynamometer (ICCs = 0.36–0.72) but excellent with the hand-held dynamometer (ICCs ≥ 0.90). There was a negative bias for the crane scale compared to alternative devices and evidence of a proportional bias for knee extension and shoulder abduction.

Conclusions

A crane scale provides values similar to hand-held dynamometry but is not comparable to results from electromechanical dynamometry. The lack of concurrent validity between the crane scale and electromechanical dynamometry may be partially attributable to differences in stabilization and participant positioning.

^a Corresponding Author:
Andrew Sprague, PT, MS, DPT, PhD
Department of Physical Therapy
100 Technology Drive, Suite 210
Pittsburgh, PA 15219
asprague@pitt.edu

Level of Evidence

3b, reliability and validity study

Restoration of muscle strength is a primary goal after orthopedic knee and shoulder injuries. Deficits in quadriceps and hamstring peak force place patients at a greater risk of re-injury following anterior cruciate ligament reconstruction and hamstring strains, respectively.^{1,2} Similarly, decreased shoulder internal and external rotation strength is associated with recurrent anterior shoulder instability.³ Despite the importance of restoring muscle force production after injury, many clinicians still rely on highly subjective methods to assess muscle strength.

In recent surveys of physical therapists' return-to-sport decision-making process, 56% reported using manual muscle testing (MMT) as their sole means of measuring knee strength after anterior cruciate ligament reconstruction and 83% relied on MMT for the upper extremity.^{4,5} MMT is a semi-quantitative method using joint motion against gravity and manual resistance to grade muscle force on a 6-point ordinal scale (0 = No observed muscle activity, 5 = "Normal" muscle function).⁶ While quick and easy to perform, MMT has varying inter-rater reliability and limited sensitivity to detect strength deficits once a minimal threshold of force production is achieved.⁶⁻⁹

Electromechanical dynamometry (BIO) and hand-held dynamometry (HHD) are more objective alternatives to MMT. BIO is currently the gold standard in measuring single-joint peak force production.¹⁰ While demonstrating excellent reliability and sensitivity to detect between-limb peak force production deficits, BIO is not routinely used in clinical settings due to its high cost, space requirements, and lengthy testing procedures. Hand-held dynamometry, a portable and more affordable dynamometer, has demonstrated acceptable reliability and validity in both the knee and shoulder when compared to BIO.¹⁰⁻¹² Despite the accessibility of HHD, continued reliance on MMT suggests that cost may be one of the barriers to widespread adoption.^{4,5}

Commercially available crane scales (CS), or strain gauges, have been explored as a method to assess single-joint muscle force production, with potentially greater objectivity than MMT and substantially less cost than electromechanical dynamometers or HHD devices. However, the reliability and validity of these devices have not been established for measuring single-joint muscle force production. Therefore, the purpose of this study is to determine if a commercially available CS is a valid alternative for isometric single-joint muscle peak force testing of the knee and shoulder when compared to HHD and BIO. The two primary aims of the study were to determine the between-trials test-retest reliability and concurrent criterion validity of an accessible crane scale for measuring isometric knee and shoulder strength. It was hypothesized that the CS would demonstrate excellent between-trials test-retest reliability (ICCs >0.9) and clinically acceptable concurrent criterion validity when compared to BIO and HHD (ICCs > 0.75).

METHODS

PARTICIPANTS

The University of Pittsburgh Institutional Review Board approved the study, and all participants provided written informed consent. A convenience sample of healthy, physically active individuals, 18 to 45 years old, was recruited through email solicitation, posted flyers, and word-of-mouth from the University of Pittsburgh community from May 1st through December 31st, 2022. Individuals were excluded if they currently had a musculoskeletal injury to the upper or lower extremities or a health condition that limited their ability to complete strenuous exercise, per the American College of Sports Medicine Preparticipation Health Screening Recommendations.¹³ An a priori power analysis was not performed.

STUDY DESIGN

This was a single-session, between-trials test-retest reliability and concurrent criterion validity study. The device of interest was a Klau Mini Weighing Crane Scale (KlauDirect), rated to measure up to 300 kg of force. The device was chosen for investigation due to its wide availability from online retailers, ability to hold peak force readings, and range of force measurement. The gold standard comparator for concurrent criterion validity was a Biodex System 4 in isometric mode (Biodex Medical Systems, Inc., Shirley, New York, USA). Additionally, a Lafayette Hand-Held Dynamometer (Model 01663, Lafayette Instrument, Lafayette, IN, USA) was used as an additional comparator since these devices are commonly available in clinical settings.

Age, height, weight, sex, and dominant limb were collected before strength testing. The participants' preferred lower extremity to kick a soccer ball was established as the dominant lower limb. Their preferred upper extremity for throwing tasks and activities of daily living was established as the dominant upper limb.

Isometric muscle strength was assessed using each device for six motions: knee extension (EXT), knee flexion (FLEX), shoulder internal rotation (IR), shoulder external rotation (ER), and shoulder abduction (ABD). To limit any influence of limb dominance and fatigue, one limb was tested for each participant and the limb to be tested and order of device testing was randomized using a block randomization scheme (www.randomizer.org) with a block size of 12. The study team was not blinded to the randomization scheme before participant enrollment. For each device, motions were evaluated in a standardized order as follows: EXT, FLEX, IR, ER, then ABD. Testing was completed by three student physical therapists in the final year of their education, who completed approximately two hours of training before initiation of the study. Testing for each device was completed by the same examiner across all participants. Given the nature of the study design, testers were

not blinded to testing results for each device and testing position.

ISOMETRIC STRENGTH TESTING

A standardized warm-up and testing process was followed for all devices and motions. Participants were instructed to complete two warmup repetitions at 50 and 75 percent of their self-perceived maximal intensity for each testing position. Approximately 30 seconds were provided between warmup repetitions. The moment arm for CS and HHD testing was recorded using a flexible tape measure. Three maximal intensity trials were then completed in each testing position with 60 seconds between trials. Once all motions were evaluated for the respective device, approximately three to five minutes were provided to switch between devices. Participants received strong verbal encouragement during all testing trials to ensure maximal intensity was achieved. Testing positions and the degree of stabilization for the CS and HHD were selected to closely mimic how these tools are being utilized in a clinical setting.

KNEE EXTENSION

Knee extension testing was performed with the participant seated at 90° of hip and knee flexion. If necessary, a cushioned pad was added to the seat to ensure that the participant's foot did not contact the ground. For the CS assessment, the device was attached to the lower leg using a padded ankle cuff positioned two finger breadths above the lateral malleolus. The CS was oriented behind the lower leg, perpendicular to the shank, and attached to an immovable object using an inelastic strap ([Figure 1A](#)). For the HHD assessment, the device was secured to the front of the lower leg, approximately two finger breadths above the lateral malleolus using an inelastic loop, with the opposite end of the loop wrapped around an immovable object.¹⁴ The moment arm for CS and HHD testing was recorded as the distance in centimeters from the lateral femoral condyle to the center of the ankle cuff along the lateral lower leg. The shank cuff for the BIO assessment was attached two finger breadths above the lateral malleolus, and inelastic straps were used to stabilize the thigh, waist, and trunk.¹⁴ Participants were instructed to hold onto the chair or stabilization handles to limit extraneous movement.

KNEE FLEXION

Participant positioning for knee flexion was identical to knee extension for each testing device. For the CS assessment, the CS was oriented in front of the lower leg, perpendicular to the shank, and attached to an immovable object using an inelastic strap ([Figure 1B](#)). The examiner stood behind the participant and stabilized the chair to maximize stabilization during testing trials. For the HHD assessment, the device was secured to the back of the lower leg, approximately 2 finger breadths above the lateral malleolus using an inelastic loop, with the opposite end of the loop wrapped around an immovable object.¹⁵ The moment arm for CS and

HHD testing was recorded as the distance in centimeters from the lateral femoral condyle to the center of the ankle cuff along the lateral lower leg. The shank cuff for the BIO assessment was attached 2 finger breadths above the lateral malleolus.¹⁵

SHOULDER EXTERNAL ROTATION

Shoulder external rotation for CS and BIO measurements was performed with the participant seated at 90° of hip and knee flexion, approximately 0° of shoulder abduction, neutral shoulder rotation, 90° of elbow flexion, and neutral forearm position. For the CS assessment, the device was attached to the distal forearm using a wrist cuff positioned just proximal to the wrist joint. The CS was oriented medial to the upper extremity, perpendicular to the forearm, and attached to an immovable object using an inelastic strap. A small towel was placed between the elbow and the trunk to maintain the arm at 0° abduction. For the HHD assessment, the participant was positioned supine with the shoulder abducted and elbow flexed to 90°. The device was placed on the dorsal forearm, just proximal to the wrist joint, with stabilization of the HHD provided by the tester.¹¹ The moment arm for CS and HHD assessment was recorded as the distance in centimeters from the lateral epicondyle of the humerus to the center of the wrist cuff or the HHD pad along the dorsal forearm. Participants were placed into a padded elbow attachment for the BIO assessment while holding a stabilized handle in the aforementioned testing position.¹¹

SHOULDER INTERNAL ROTATION

Participant positions for shoulder internal rotation were identical to shoulder external rotation for each testing device. For the CS assessment, the device was attached to the distal forearm using a wrist cuff positioned just proximal to the wrist joint. The CS was oriented lateral to the upper extremity, perpendicular to the forearm, and attached to an immovable object using an inelastic strap ([Figure 1D](#)). For the HHD assessment, the participant positioning was identical to external rotation but the device was placed on the volar forearm, just proximal to the wrist joint, with stabilization provided by the tester.¹² The moment arm for CS and HHD assessment was recorded as the distance from the lateral epicondyle of the humerus to the center of the wrist strap or dynamometer along the dorsal forearm. Participants were placed into a padded elbow attachment for the BIO assessment while holding a stabilized handle in the aforementioned testing position.¹¹

SHOULDER ABDUCTION

CS and BIO measurements of shoulder abduction were completed with the participant seated at 90° of hip and knee flexion, 90° of shoulder elevation in the scapular plane, and neutral shoulder rotation. For the CS assessment, the device was attached to the distal forearm using a wrist cuff positioned just proximal to the wrist joint. The CS was oriented inferior to the upper extremity, perpendicu-



Figure 1. Participant positioning for crane scale assessment of A) knee extension, B) knee flexion, C) shoulder abduction, and D) shoulder internal rotation.

lar to the forearm, and attached to an immovable object using an inelastic strap (Figure 1C). For the HHD assessment, the device was secured to the dorsal forearm, just proximal to the wrist joint, with stabilization provided by the tester while the participant laid supine with the shoulder abducted to 90°.16 The moment arm for CS and HHD assessment was recorded as the distance from the acromioclavicular joint to the center of the wrist cuff. Participants held a stabilized handle for the BIO assessment while seated in the aforementioned testing position.17

STATISTICAL ANALYSES

Between-trials test-retest reliability was examined using intraclass correlation coefficients (ICC) using a single-rating, absolute agreement, two-way mixed effects model.18 ICCs were classified as poor (<0.5), moderate (0.5 to <0.75), good (0.75 to <0.90), or excellent (.90 or above).18 The standard error of measurement (SEM) was calculated based on the obtained ICCs: $SEM = s_1 \times (\sqrt{1 - ICC})$, where s_1 = standard deviation of sample measurements.19 Addition-

ally, the SEM was used to calculate the minimal detectable change (MDC): $MDC_{95\%} = 1.96 \times \sqrt{2} \times SEM$.20

Before validity analyses, CS and HHD results were converted to Newton meters. Since the direction of force was oriented approximately 90 degrees relative to the moment arm, the equation was simplified to $\tau = MA \times F$, where τ = torque, MA = moment arm, and F = force.

Validity was examined using ICCs, Pearson's correlation coefficients, absolute and relative error calculations, Bland-Altman plots, and simple linear regression. Normality was confirmed for torque values of each device and motion using Shapiro-Wilk's tests and visual inspection of quantile-quantile plots. ICCs were computed using an average rating, absolute agreement, two-way mixed effects model, classified according to the cut-offs described above.18 Pearson's correlation coefficient (r) was used to examine the relationships between torque obtained by the CS to BIO and HHD. Correlation coefficients were classified as negligible (<0.10), weak (0.10 to 0.39), moderate (0.40 to 0.69), strong (0.70 to 0.89), or very strong (0.90 to 1.00).21 While ICCs are the preferred method for examining the concurrent criterion validity of continuous measures,18 Pearson's

Table 1. ICCs with 95% confidence intervals, individual trial mean values, standard error of the measurement, and minimal detectable change for crane scale measurements.

	ICC (95% CI)	Trial 1 mean (SD) [N]	Trial 2 mean (SD) [N]	Trial 3 mean (SD) [N]	SEM (N)	MDC _{95%} (N)
Knee flexion	0.96 (0.92, 0.98)	195.2 (76.4)	197.4 (75.5)	201.8 (78.1)	15.1	41.9
Knee extension	0.95 (0.90, 0.98)	385.0 (107.6)	396.4 (109.3)	401.2 (103.1)	23.3	64.6
Shoulder ER	0.95 (0.88, 0.98)	110.7 (37.8)	116.9 (42.6)	120.3 (44.9)	9.3	25.8
Shoulder IR	0.97 (0.93, 0.99)	177.5 (64.0)	177.1 (66.9)	181.0 (75.2)	12.5	34.6
Shoulder abduction	0.95 (0.90, 0.98)	90.0 (33.0)	88.8 (30.2)	91.2 (32.7)	7.2	20.0

ER = external rotation; IR = internal rotation; ICC = Intra-class correlation coefficient, CI = confidence interval; SD = standard deviation; N = newtons; SEM = standard error of the measure, MDC_{95%} = Minimal detectable change at the 95% confidence level.

correlation coefficient was calculated to facilitate comparison with previous studies. Bland-Altman plots were constructed with 95% limits of agreement (LoA) and bias estimates.²² Additionally, simple linear regression was utilized, with the average of measurements from the CS and each alternative device regressed on the difference between devices. The resulting linear regression line was overlaid on the Bland-Altman plots and visually inspected for evidence of a proportional bias.²³ A substantial deviation of the regression line slope from zero would imply a non-constant bias across the potential range. Bland-Altman plots with regression lines were generated in Prism 10 v. 10.2.2. (Graphpad Software, LLC, Boston, MA) and all other analyses were performed in SPSS Statistics v. 28 (IBM Corp., Armonk, NY).

RESULTS

A total of 20 healthy participants were recruited. Participants were 50% female and had a mean (SD) age of 25.4 (3.8) years, height of 172.1 (9.2) cm, weight of 74.2 (13.3) kg, and body mass index of 24.9 (2.9) m/kg². All 20 participants completed CS testing for all motions. However, due to equipment malfunctions, 19 participants completed BIO shoulder abduction testing, 12 completed HHD knee extension testing, and 11 completed HHD testing for all other motions.

TEST-RETEST RELIABILITY

ICCs were excellent for between-trials test-retest reliability (Table 1). The SEM ranged from 7.2 to 23.3 N and the MDC_{95%} ranged from 20.0 to 64.6 N (Table 1).

CONCURRENT VALIDITY – BIODEX (BIO)

ICCs for validity ranged from poor to moderate, with substantial absolute and relative error (Table 2). There was a strong correlation between devices for knee flexion and moderate for other motions (Table 2). Absolute and relative error was greater for knee motions relative to the shoulder

motions. On average, the CS underestimated the values obtained by BIO, with a negative bias for all motions at the knee and shoulder (Figure 2). There appeared to be evidence for a proportional bias with knee extension and shoulder abduction, with bias increasing as the average torque produced increased (Figures 2B and 2E).

CONCURRENT VALIDITY – HAND-HELD DYNAMOMETRY (HHD)

ICCs for validity were excellent (Table 3). Very strong correlations were found between devices for knee flexion, shoulder IR, and shoulder abduction, and strong correlations for knee extension and shoulder external rotation. Relative error was 9.7% or greater for all motions evaluated but less than the error observed for validity with BIO (Table 3). On average, the CS underestimated values obtained by HHD, with a negative bias for all tested motions (Figure 3). Similar to the comparison with BIO, there was evidence for a proportional bias with knee extension and shoulder abduction, with bias increasing as the average torque produced increased (Figures 3B and 3E).

See Supplementary Tables 1-3 for results of linear regression.

DISCUSSION

This study investigated a CS as an alternative to BIO and HHD for measuring isometric knee and shoulder strength. The CS device demonstrated excellent between-trial test-retest reliability for assessing force during knee and shoulder strength testing. Concurrent criterion validity was moderate to poor compared to gold standard BIO but was excellent compared to HHD. In general, the CS underestimated torque values obtained by both BIO and HHD and there was evidence of a proportional bias for knee extension and shoulder abduction.

The hypothesis regarding between-trials test-retest reliability was supported, as ICCs were excellent (ICCs >0.9) for all motions. Therefore, the CS provides consistent values within a single testing session. Since all data were collected

Table 2. ICCs and Pearson’s r with 95% confidence intervals, mean torque values, and absolute and relative error for comparisons between crane scale and Biodex measurements.

	ICC (95% CI)	Pearson’s r (95% CI)	Crane scale mean (SD) [Nm]	Biodex mean (SD) [Nm]	Absolute error (SD) [Nm]	Relative error (SD) [%]
Knee flexion	0.74 (-0.05, 0.92)	0.75 (0.46, 0.90)	62.9 (32.8)	87.8 (33.5)	28.0 (19.5)	31.7 (19.7)
Knee extension	0.36 (-0.19, 0.74)	0.59 (0.19, 0.82)	124.4 (49.2)	222.2 (59.5)	97.8 (50.4)	43.7 (18.2)
Shoulder ER	0.53 (-0.04, 0.80)	0.41 (-0.04, 0.72)	25.7 (10.5)	32.3 (12.1)	8.9 (10.7)	25.4 (22.0)
Shoulder IR	0.67 (0.21, 0.87)	0.55 (0.14, 0.80)	39.7 (17.6)	47.6 (20.7)	12.9 (15.2)	26.3 (32.4)
Shoulder abduction (n = 19)	0.72 (0.30, 0.89)	0.63 (0.24, 0.84)	46.7 (17.7)	54.4 (26.2)	14.1 (16.5)	25.8 (39.5)

n = 20 unless otherwise noted; ER = external rotation; IR = internal rotation; ICC = Intra-class correlation coefficient, CI = confidence interval; SD = standard deviation; Nm = new-ton-meter.

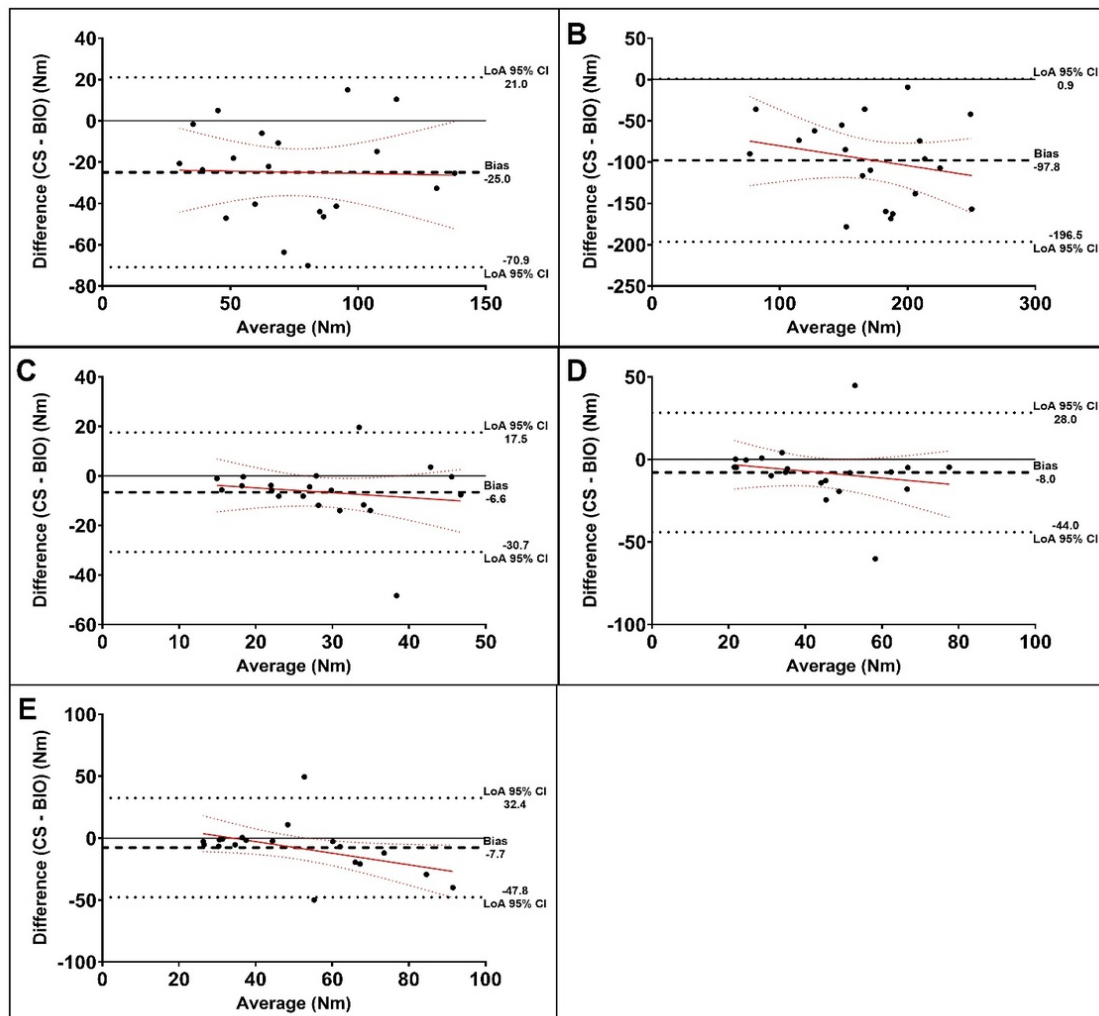


Figure 2. Simple linear regression (red solid line) with 95% confidence intervals (red dashed line) of the relationship between the difference and mean of torque values obtained with the crane scale (CS) and Biodex (BIO) on Bland-Altman plots for A) knee flexion, B) knee extension, C) shoulder external rotation, D) shoulder internal rotation, and E) shoulder abduction. LOA = Limits of Agreement.

Table 3. ICCs and Pearson’s r with 95% confidence intervals, mean torque values, and absolute and relative error for comparisons between crane scale and hand-held dynamometry measurements.

	ICC (95% CI)	Pearson’s r (95% CI)	Crane scale mean (SD) [Nm]	HHD mean (SD) [Nm]	Absolute error (SD) [Nm]	Relative error (SD) [%]
Knee flexion	0.96 (0.86, 0.99)	0.93 (0.76, 0.98)	75.8 (37.8)	82.5 (40.4)	11.9 (10.0)	14.0 (12.3)
Knee extension (n = 12)	0.91 (0.70, 0.97)	0.86 (0.57, 0.96)	139.6 (51.1)	147.8 (66.7)	24.8 (24.5)	18.0 (5.7)
Shoulder ER	0.90 (0.57, 0.97)	0.85 (0.51, 0.96)	29.5 (12.4)	33.9 (12.6)	6.7 (4.2)	22.0 (12.0)
Shoulder IR	0.97 (0.88, 0.99)	0.94 (0.79, 0.99)	49.0 (21.5)	51.0 (20.8)	5.0 (5.0)	9.7 (8.3)
Shoulder abduction	0.92 (0.55, 0.98)	0.95 (0.83, 0.99)	50.1 (18.5)	58.3 (25.8)	10.3 (7.4)	18.7 (12.8)

n = 11 unless otherwise noted; ER = external rotation; IR = internal rotation; ICC = Intra-class correlation coefficient, CI = confidence interval; SD = standard deviation; Nm = new-ton-meter.

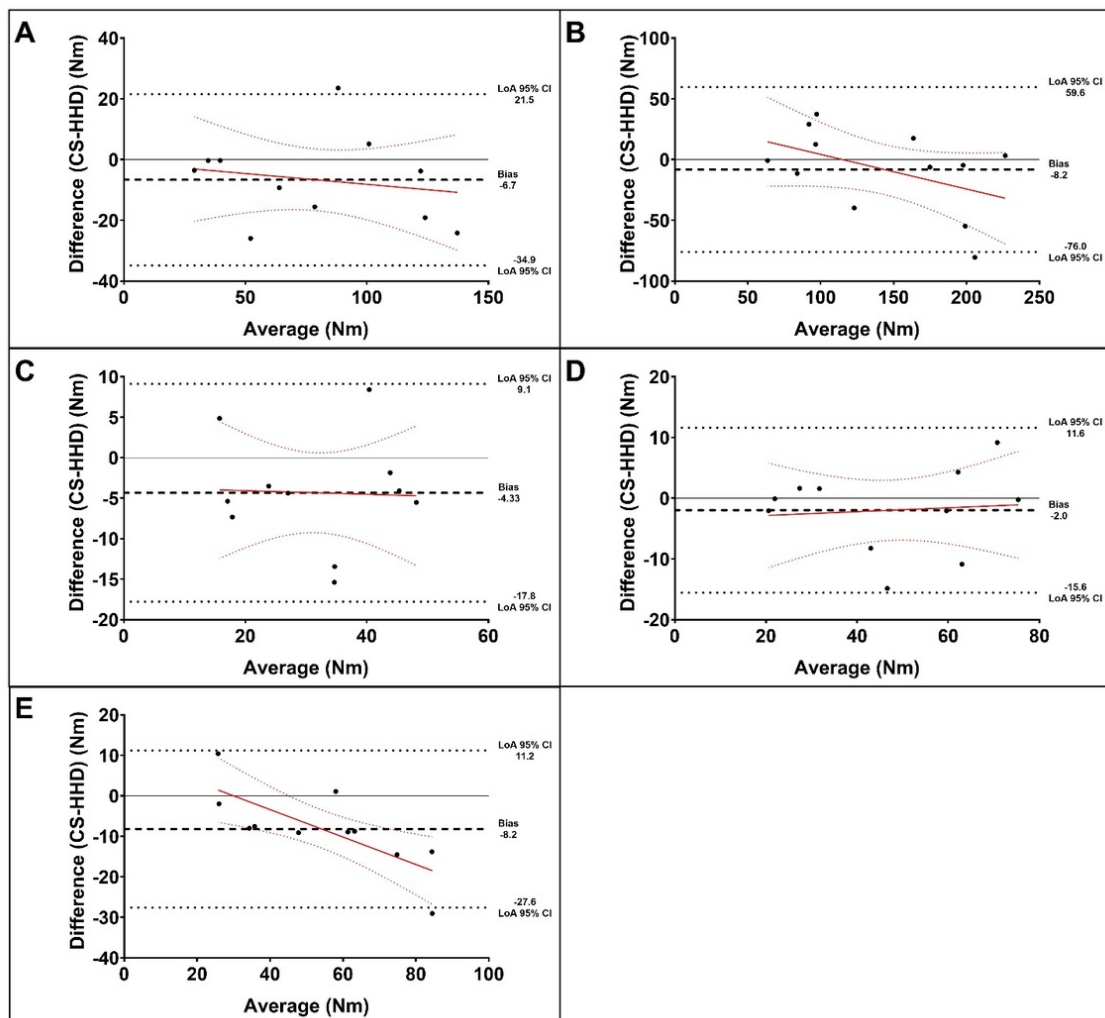


Figure 3. Simple linear regression (red solid line) with 95% confidence intervals (red dashed line) of the relationship between the difference and mean of torque values obtained with the crane scale (CS) and hand-held dynamometry (HHD) on Bland-Altman plots for A) knee flexion, B) knee extension, C) shoulder external rotation, D) shoulder internal rotation, and E) shoulder abduction. LOA = Limits of Agreement.

at a single time point, it was not possible to assess between-session reliability. As a result, the coefficients, SEM, and MDC values should not be generalized to measurements obtained across different testing sessions.

The hypotheses regarding concurrent validity were partially supported. Compared to the gold standard BIO, ICCs were poorer than expected (ICCs < 0.75) for all motions. However, concurrent validity with HHD was clinically acceptable, with excellent ICCs (ICCs > 0.9) for all motions. The CS underestimated torque values obtained by BIO and HHD for both comparisons. The magnitude of underestimation was much larger for BIO than for HHD. Additionally, there appeared to be a proportional bias for knee extension and shoulder abduction when compared to either alternative device. The CS tended to underestimate torque values to a greater degree as torque increased in these instances.

To the authors' knowledge, this is the first study to investigate the reliability and validity of a readily available CS for assessing knee and shoulder strength. One previous study investigated the concurrent criterion validity of an in-line dynamometer for knee extension strength in participants with and without anterior cruciate ligament reconstruction.²⁴ Validity compared to gold standard BIO was excellent (ICCs = 0.97 – 0.98) and there was no evidence of fixed or proportional bias. However, knee extension strength was assessed at 60° of knee flexion using a substantially more costly and less accessible in-line dynamometer specifically designed to measure muscle strength. Therefore, these results are not directly comparable.

In the absence of prior research exploring CSs as a method of strength assessment, the authors deemed it necessary to include an analysis of concurrent validity with HHD. HHD is more widely used in clinical settings than BIO. While it is unlikely that CSs will supplant BIO, they may yield results similar to those obtained through HHD. Consequently, comparing the current findings with existing literature comparing BIO and HHD provides valuable context for this study.

Most prior studies comparing BIO and HHD focused on the knee and evaluated validity using correlation coefficients other than ICCs, such as Pearson's or Spearman's coefficients) to assess validity. Additionally, there is substantial variability in the testing position and degree of stabilization during HHD across studies. Correlation coefficients ranged from 0.75 – 0.93 for knee extension and 0.69 – 0.88 for knee flexion.^{14,15,25-28} The observed correlation coefficients in the current study between the CS and BIO were lower for knee extension but similar for knee flexion. Two studies examined the concurrent validity of shoulder external and internal rotation with correlation coefficients ranging from 0.65 to 0.79 and 0.66 to 0.82, respectively.^{12,29} The observed correlation coefficients were lower for both motions. Although commonly used, correlation coefficients are a measure of relationship, not agreement, and do not account for fixed bias between two measures.³⁰ Only one study has utilized ICCs to examine concurrent validity between BIO and HHD.³¹ The authors found ICCs ranging from 0.72 to 0.87 for knee flexion and 0.40 to 0.92 for knee

extension. Our ICCs were similar for knee flexion but lower for knee extension. Combined, it appears that the concurrent validity of the CS is similar to that of HHD for knee flexion but lower for knee extension and shoulder motions.

A key decision in designing the study was to prioritize ecological validity and generalizability by selecting device set-ups and participant positioning that mimic how CSs and HHDs are typically used in clinical practice. Consequently, the CS and HHD assessments involved substantially less stabilization compared to the BIO assessments and to previous investigations of similar devices.^{14,24} For instance, thigh and trunk straps were not used during lower extremity testing with the CS and HHD, as such stabilization is rarely employed in clinical practice based on our experience. The reduced stabilization may have introduced extraneous movements and reduced force or torque production.^{32,33} These factors could partially explain why the CS significantly underestimated torque relative to the BIO, as well as the differing concurrent validity findings between the BIO and HHD and the inconsistencies with prior research. The authors suggest that clinicians use the maximum amount of stabilization that is feasible when using the devices to improve measurement accuracy.

Several limitations to this study must be acknowledged. First, the study was conducted in a relatively small sample of healthy individuals. As a result, there was low precision around the point estimates of the validity coefficients, a risk of type II error if interpreting the significance of linear regression, and limited generalizability to injured populations. The concerns regarding a small sample size were compounded by equipment malfunction, which further reduced the sample size for HHD measurements. To minimize the risk of type II error the authors relied on visual inspection of regression lines and Bland-Altman plots, rather than statistical significance, to identify fixed and proportional bias. A second limitation was the length of the data collection since testing was performed for five motions with three devices. This increased the potential for physical and mental fatigue. The order of device testing was randomized to limit bias introduced by physical fatigue. Finally, the participant positioning was not identical for all devices. The HHD set-up for shoulder motions deviated from the crane-scale set-up to mirror clinical use. Due to mechanical limitations, it was not possible to exactly recreate the CS set-up for shoulder motions on the BIO. These discrepancies in positioning likely contribute to the decreased concurrent validity.

Clinicians should carefully consider the affordability and convenience of commercially available CSs in comparison to their unknown longitudinal reliability and concerns regarding concurrent validity before integrating these devices into their clinical practice for measuring knee or shoulder strength. At the time of this study, the cost of the CS was approximately \$50, significantly lower than comparable HHDs (\$1,000) and IDs (\$40,000).

Although the CS exhibited excellent between-trials reliability, clinicians should be aware that these values were obtained without removing and reapplying the device, which would have likely resulted in lower reliability co-

efficients. Additionally, between-session and interrater reliability remain unexplored, requiring caution when comparing values over time or between different examiners. Finally, the minimal detectable change needs to be established over longer time intervals to determine whether perceived changes in strength are more than measurement error.

Despite lower-than-expected concurrent validity compared to BIO across all motions, the CS demonstrated excellent concurrent validity with HHD, suggesting its ability to produce values comparable to those obtained via HHD. However, clinicians should be aware that concerns have been raised about the reliability and validity of HHD for large muscle groups, such as the knee extensors.^{10,34} If clinicians intend to calculate a limb symmetry index based on these measurements, concerns regarding validity may be mitigated, as the index may minimize the impact of a fixed bias. However, limb symmetry indexes were not examined in this study, as only one side was tested per participant.

CONCLUSIONS

The results of this study indicate that a readily available CS demonstrates excellent between-trials test-retest reliability

for assessing isometric knee and shoulder strength. Concurrent validity with HHD is also excellent, but moderate to poor compared to BIO. Therefore, a CS may provide values similar to HHD but is not comparable to BIO. Future research should investigate longitudinal reliability and measurement properties when calculating limb symmetry indices in injured populations to provide additional context on the clinical utility of the device.

.....

DISCLOSURES

The authors have nothing to disclose.

ACKNOWLEDGEMENTS

This work was supported in part by the National Institutes of Health, National Institute of Arthritis and Musculoskeletal and Skin Diseases under award #K23AR080741.

Submitted: July 17, 2024 CDT. Accepted: February 13, 2025

CDT. Published: April 01, 2025 CDT.

© The Author(s)



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license's legal deed at <https://creativecommons.org/licenses/by-nc/4.0> and legal code at <https://creativecommons.org/licenses/by-nc/4.0/legalcode> for more information.

REFERENCES

1. Grindem H, Snyder-Mackler L, Moksnes H, Engebretsen L, Risberg MA. Simple decision rules can reduce reinjury risk by 84% after ACL reconstruction: the Delaware-Oslo ACL cohort study. *Br J Sports Med.* 2016;50(13):804-808. doi:[10.1136/bjsports-2016-096031](https://doi.org/10.1136/bjsports-2016-096031)
2. Vos RJD, Reurink G, Goudswaard GJ, Moen MH, Weir A, Tol JL. Clinical findings just after return to play predict hamstring re-injury, but baseline MRI findings do not. *Br J Sports Med.* 2014;48(18):1377-1384. doi:[10.1136/bjsports-2014-093737](https://doi.org/10.1136/bjsports-2014-093737)
3. Edouard P, Degache F, Beguin L, et al. Rotator cuff strength in recurrent anterior shoulder instability. *JBJS.* 2011;93(8):759. doi:[10.2106/jbjs.i.01791](https://doi.org/10.2106/jbjs.i.01791)
4. Greenberg EM, Greenberg ET, Albaugh J, Storey E, Ganley TJ. Rehabilitation practice patterns following anterior cruciate ligament reconstruction: A survey of physical therapists. *J Orthop Sports Phys Ther.* 2018;48(10):801-811. doi:[10.2519/jospt.2018.8264](https://doi.org/10.2519/jospt.2018.8264)
5. Gauthier ML, Unverzagt CA, Mendonça LDM, Seitz AL. Missing the forest for the trees: A lack of upper extremity physical performance testing in sports physical therapy. *Int J Sports Phys Ther.* 2023;18(2). doi:[10.26603/001c.73791](https://doi.org/10.26603/001c.73791)
6. Bohannon RW. Considerations and practical options for measuring muscle strength: A narrative review. *BioMed Res Int.* 2019;2019:1-10. doi:[10.1155/2019/8194537](https://doi.org/10.1155/2019/8194537)
7. Knepler C, Bohannon RW. Subjectivity of forces associated with manual-muscle test grades of 3+, 4-, and 4-. *Percept Mot Skills.* 1998;87(3 suppl):1123-1128. doi:[10.2466/pms.1998.87.3f.1123](https://doi.org/10.2466/pms.1998.87.3f.1123)
8. Mulroy SJ, Lassen KD, Chambers SH, Perry J. The ability of male and female clinicians to effectively test knee extension strength using manual muscle testing. *J Orthop Sports Phys Ther.* 1997;26(4):192-199. doi:[10.2519/jospt.1997.26.4.192](https://doi.org/10.2519/jospt.1997.26.4.192)
9. Nagatomi T, Mae T, Nagafuchi T, Yamada S ichi, Nagai K, Yoneda M. Shoulder manual muscle resistance test cannot fully detect muscle weakness. *Knee Surg Sports Traumatol Arthrosc.* 2017;25(7):1795. doi:[10.1007/s00167-016-4380-y](https://doi.org/10.1007/s00167-016-4380-y)
10. Stark T, Walker B, Phillips JK, Fejer R, Beck R. Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: A systematic review. *PM&R.* 2011;3(5):472-479. doi:[10.1016/j.pmrj.2010.10.025](https://doi.org/10.1016/j.pmrj.2010.10.025)
11. Popchak A, Poploski K, Patterson-Lynch B, Nigolian J, Lin A. Reliability and validity of a return to sports testing battery for the shoulder. *Phys Ther Sport.* 2021;48:1-11. doi:[10.1016/j.pts.2020.12.003](https://doi.org/10.1016/j.pts.2020.12.003)
12. Croteau F, Robbins SM, Pearsall D. Hand-held shoulder strength measures correlate with isokinetic dynamometry in elite water polo players. *J Sport Rehabil.* 2021;30(8):1233-1236. doi:[10.1123/jsr.2020-0277](https://doi.org/10.1123/jsr.2020-0277)
13. Riebe D, Franklin BA, Thompson PD, et al. Updating ACSM's Recommendations for exercise preparticipation health screening. *Med Sci Sports Exerc.* 2015;47(11):2473-2479. doi:[10.1249/MSS.0000000000000664](https://doi.org/10.1249/MSS.0000000000000664)
14. Bohannon RW, Bubela DJ, Wang YC, Magasi SR, Gershon RC. Adequacy of belt-stabilized testing of knee extension strength. *J Strength Cond Res.* 2011;25(7):1963-1967. doi:[10.1519/jsc.0b013e3181e4f5ce](https://doi.org/10.1519/jsc.0b013e3181e4f5ce)
15. Katoh M, Hiiragi Y, Hirano M, et al. Isometric knee muscle strength measurement using a belt-stabilized hand-held dynamometer and an isokinetic dynamometer with and without trunk fixation: investigation of agreement of measurement values and factors influencing measurement. *J Phys Ther Sci.* 2019;31(11):878-883. doi:[10.1589/jpts.31.878](https://doi.org/10.1589/jpts.31.878)
16. González-Rosalén J, Benítez-Martínez JC, Medina-Mirapeix F, Cuerda-Del Pino A, Cervelló A, Martín-San Agustín R. Intra- and inter-rater reliability of strength measurements using a pull hand-held dynamometer fixed to the examiner's body and comparison with push dynamometry. *Diagnostics.* 2021;11(7):1230. doi:[10.3390/diagnostics11071230](https://doi.org/10.3390/diagnostics11071230)
17. *Biodex Multi-Joint System: Setup/Operation Manual.*; 2016.
18. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15(2):155-163. doi:[10.1016/j.jcm.2016.02.012](https://doi.org/10.1016/j.jcm.2016.02.012)
19. Thomas JR, Nelson JK, Silverman SJ. *Research Methods in Physical Activity.* 6th ed. Human Kinetics; 2011.

20. Portney L, Watkins M. Statistical measures of reliability. In: *Foundations of Clinical Research*. ; 2000.
21. Schober P, Boer C, Schwarte LA. Correlation coefficients: Appropriate use and interpretation. *Anesth Analg*. 2018;126(5):1763-1768. doi:[10.1213/ANE.0000000000002864](https://doi.org/10.1213/ANE.0000000000002864)
22. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Int J Nurs Stud*. Published online 2010;6. doi:[10.1016/j.ijnurstu.2009.10.001](https://doi.org/10.1016/j.ijnurstu.2009.10.001)
23. Ho KM. Using linear regression to assess dose-dependent bias on a Bland-Altman plot. *J Emerg Crit Care Med*. 2018;2(8). doi:[10.21037/jeccm.2018.08.02](https://doi.org/10.21037/jeccm.2018.08.02)
24. Norris R, Morrison S, Price A, et al. Inline dynamometry provides reliable measurements of quadriceps strength in healthy and ACL-reconstructed individuals and is a valid substitute for isometric electromechanical dynamometry following ACL reconstruction. *The Knee*. 2024;46:136-147. doi:[10.1016/j.knee.2023.12.006](https://doi.org/10.1016/j.knee.2023.12.006)
25. Hirano M, Katoh M, Gomi M, Arai S. Validity and reliability of isometric knee extension muscle strength measurements using a belt-stabilized hand-held dynamometer: a comparison with the measurement using an isokinetic dynamometer in a sitting posture. *J Phys Ther Sci*. 2020;32(2):120-124. doi:[10.1589/jpts.32.120](https://doi.org/10.1589/jpts.32.120)
26. Hansen EM, McCartney CN, Sweeney RS, Palimenio MR, Grindstaff TL. Hand-held dynamometer positioning impacts discomfort during quadriceps strength testing: A validity and reliability study. *Int J Sports Phys Ther*. 2015;10(1):62-68.
27. Muff G, Dufour S, Meyer A, et al. Comparative assessment of knee extensor and flexor muscle strength measured using a hand-held vs. isokinetic dynamometer. *J Phys Ther Sci*. 2016;28(9):2445-2451. doi:[10.1589/jpts.28.2445](https://doi.org/10.1589/jpts.28.2445)
28. Martins J, da Silva JR, da Silva MRB, Bevilaqua-Grossi D. Reliability and validity of the belt-stabilized handheld dynamometer in hip- and knee-strength tests. *J Athl Train*. 2017;52(9):809-819. doi:[10.4085/1062-6050-52.6.04](https://doi.org/10.4085/1062-6050-52.6.04)
29. Chen B, Liu L, Chen LB, et al. Concurrent validity and reliability of a handheld dynamometer in measuring isometric shoulder rotational strength. *J Sport Rehabil*. 2021;30(6):965-968. doi:[10.1123/jsr.2020-0021](https://doi.org/10.1123/jsr.2020-0021)
30. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: Measures of agreement. *Perspect Clin Res*. 2017;8(4):187-191. doi:[10.4103/picr.picr_123_17](https://doi.org/10.4103/picr.picr_123_17)
31. Mentiplay BF, Perraton LG, Bower KJ, et al. Assessment of lower limb muscle strength and power using hand-held and fixed dynamometry: A reliability and validity study. Haddad JM, ed. *PLOS ONE*. 2015;10(10):e0140822. doi:[10.1371/journal.pone.0140822](https://doi.org/10.1371/journal.pone.0140822)
32. Kolber MJ, Beekhuizen K, Cheng MSS, Fiebert IM. The reliability of hand-held dynamometry in measuring isometric strength of the shoulder internal and external rotator musculature using a stabilization device. *Physiother Theory Pract*. 2007;23(2):119-124. doi:[10.1080/09593980701213032](https://doi.org/10.1080/09593980701213032)
33. Yasacı Z, Argut SK, Celik D. Reliability of a new stabilization device for measurement of muscle strength using a hand-held dynamometer. *Muscle Nerve*. 2024;70(6):1240-1246. doi:[10.1002/mus.28266](https://doi.org/10.1002/mus.28266)
34. Chamorro C, Armijo-Olivo S, De la Fuente C, Fuentes J, Javier Chiroso L. Absolute reliability and concurrent validity of hand held dynamometry and isokinetic dynamometry in the hip, knee and ankle joint: Systematic review and meta-analysis. *Open Med*. 2017;12:359-375. doi:[10.1515/med-2017-0052](https://doi.org/10.1515/med-2017-0052)

SUPPLEMENTARY MATERIALS

Supplemental Tables

Download: <https://ijspt.scholasticahq.com/article/132166-reliability-and-validity-of-a-crane-scale-for-isometric-knee-and-shoulder-strength-assessment/attachment/271097.docx>
