Physiotherapy 131 (2026) 101868

The ExtRA Capacity Test: Reliability, validity and normative data of a new clinical tool for assessing shoulder muscle performance

Harry Ford[a,*], Jeremy Lewis[b], Vasileios Tyros[c], Daniel C. Low[a], William R. Bateman[d], Youmin Lee[e], Raj Sakadasariya[a], Joshua Rex[a], Matthew Oldfield[f], Marco Davare[g], Aliah Faisal Shaheen[a,*]

[a] Department of Sport, Health and Exercise Sciences, College of Health, Medicine and Life Sciences, Brunel University of London, Kingston Lane, Uxbridge, Middlesex, United Kingdom
[b] Central London Community Healthcare National Health Service Trust, Therapy Department, Finchley Memorial Hospital, Granville Road, London, United Kingdom
[c] PCRF Pirbright, Physiotherapy Department, Pirbright, United Kingdom
[d] Surrey Physio, 409 London Road, Mitcham, United Kingdom
[e] Brunel Medical School, College of Health, Medicine and Life Sciences, Brunel University of London, Kingston Lane, Uxbridge, Middlesex, United Kingdom
[f] School of Mechanical Engineering Sciences, University of Surrey, Guildford, Surrey, United Kingdom
[g] Faculty of Life Sciences and Medicine, King's College London, London, United Kingdom

Abstract

Objectives Introduce The ExtRA Capacity Test, a measure for assessing shoulder muscle performance. Assess its reliability, validity and present normative scores in a large sample of asymptomatic adults.
Design Cross-sectional observational study with test–retest.
Setting Community.
Participants Volunteers (n = 344, age 20–90 years).
Interventions The ExtRA Capacity Test involves two capacity tests completed to a 30 beats per minute metronome: maximal scapular plane lateral raises to 90° abduction with 2.5 kg of external load, and maximal external rotations in unsupported prone lying with the shoulder at 90° abduction.

Reliability was assessed in 30 asymptomatic participants, tested by two raters over two sessions, one week apart, using Bland–Altman analysis to determine mean bias and 95 % limits of agreement (LoA) as measures of error. Criterion validity was evaluated in 20 participants using Pearson correlation to examine the relationship between ExtRA and isokinetic dynamometry measures. A normative dataset was also established from 344 asymptomatic individuals across a range of ages, physical activity levels, and both sexes.
Results The intra-rater and inter-rater agreement for the ExtRA Capacity Test was assessed in a sample of 30 participants. The 95 % LoA for abduction and external rotation measurements ranged from 2.9 to 13.1 repetitions. In a sample of 20 participants, the abduction test showed good/moderate correlation with muscle strength measures but not with the external rotation test. Older age, female sex and not achieving the WHO activity guidelines have a negative impact on ExtRA performance.
Conclusions Within the caveats discussed in this paper, ExtRA can be considered a reasonably reliable tool for assessing shoulder strength and control in a clinical setting. The normative database will help clinicians set rehabilitation or return-to-play targets based on sex, age, and physical activity level.
Contributions of Paper

- This study introduces the ExtRA Capacity Test as a reliable tool for assessing shoulder muscle performance in both sporting and non-sporting populations.

* Corresponding authors.
E-mail addresses: Harry.Ford@brunel.ac.uk (H. Ford), aliah.shaheen@brunel.ac.uk (A.F. Shaheen).

https://doi.org/10.1016/j.physio.2025.101868

- The test demonstrates clinically acceptable intra- and inter-rater reliability, with the abduction component showing a strong correlation with strength measures from Isokinetic Dynamometry.
- The normative database established in this study facilitates the evaluation of shoulder performance relative to reference values stratified by age, sex, and physical activity level.
- Given its high reliability, the ExtRA Capacity Test can be used to monitor performance changes over time, providing valuable insights for clinical practice and rehabilitation planning.

Keywords: Shoulder; Capacity testing; Strength; Muscular control; Isokinetic dynamometry

Introduction

Reduction in muscle performance may result from injury, trauma, disease, co-morbidities, immobilisation, kinesiophobia, or sedentary lifestyle [1,2]. Shoulder strength, range of motion (ROM), and functional outcome measures are used to guide safe return to activity or sport [3]. Current methods for measuring shoulder strength vary in cost and clinical utility. Isokinetic dynamometry (IKD) is the gold standard for assessing muscle strength and endurance through a specified ROM [4,5]. However, its high cost and time requirements limit clinical application.

Faster, cost-effective methods like handheld dynamometry (HHD) and manual muscle testing (MMT) are widely used [3,4,6,7]. HHD demonstrates excellent reliability for isometric strength measurement [6], but isometric tests may not reflect functional demands [3]. On the other hand, MMT lacks precision, detecting changes only when strength is reduced by 15–25 % compared to the contralateral side [7]. Capacity testing, which measures the maximal number of repetitions before fatigue, may offer a more functional and practical alternative. It replicates real-world movements, provides objective outcomes, and avoids costly equipment or lengthy setup [8]. Open-chain capacity testing may also assess shoulder muscular control, a critical component for addressing shoulder pain [9–11]. Lower limb capacity tests with normative values already guide rehabilitation and return-to-play criteria [12,13].

Currently, upper limb capacity tests are limited to specific groups, such as elite athletes or those with neurological conditions, restricting their utility for broader musculoskeletal care [14–16]. The Shoulder Arm Return to Sports battery consists of 8 tests including push-up-claps, 5 kg overhead snatch and line hold rotations, to determine readiness to return to sport after upper extremity injury [16]. These tests require a high level of strength, power and neuromuscular control and would be inappropriate for those with lower functional demands [17]. No upper limb capacity test exists for the general population [18–20].

The aim of this study was to introduce the ExtRA Capacity Test, a new measure for assessing shoulder muscle performance (strength and control) using a functional capacity approach [11,21,22]. Secondary objectives included evaluating the inter- and intra-rater reliability of the tests, assessing their criterion validity, and examining normative scores in a large asymptomatic adult sample, accounting for age, sex, and adherence to the World Health Organization (WHO) Physical Activity Guidelines [23].

Materials and methods

Participants

The study was approved by the Institutional Research Ethics Committee. The study sampling was carried out between April 2022 and August 2025. Participants aged ≥ 20 years old with a QuickDASH score of < 5 (indicating excellent upper limb function) were included [24]. Exclusions were systemic illness, cervical/shoulder pain at rest or movement, cervical/shoulder pain or treatment in the past 12 months, history of spinal/upper limb fractures or surgery, and pregnancy.

Written informed consent was obtained. Data from 344 participants were collected, 30 of whom took part in the reliability testing and 20 in the validity testing. A further 174 were rejected at initial sampling due to a QuickDASH score > 5. At least 10 participants per decade (20s–80 +), for both sexes, meeting/not meeting the World Health Organisation (WHO) activity guidelines, provided a heterogeneous sample for generalizable normative data [24]. Patients/public were not involved in the research design, conduct, or reporting.

Procedures

Baseline measures

Height (m), weight (kg), age (years) and biological sex were recorded on the day of testing. Self-reported physical activity levels were assessed based on participants' adherence to the WHO Physical Activity Guidelines [23].

Participants indicated agreement or disagreement with two standardised statements: "*You complete at least 150–300 min of moderate-intensity aerobic physical activity; or at least 75–150 min of vigorous- intensity aerobic physical activity per week"* and *"you complete muscle-strengthening activities at moderate or greater intensity that involve all major muscle groups on 2 or more days a week"* [23]. Meeting activity guidelines, required agreement with both statements.

The ExtRA Capacity Test

The ExtRA (External Rotation/Abduction) Capacity Test measures the maximum number of repetitions performed before fatigue or a loss of control of scapular plane shoulder abduction in standing and prone-lying external rotation.

For the abduction capacity test, participants stood with their back against a wall, feet shoulder-width apart, with the elbow fully extended, abducting their arm to 90° in the scapular plane while holding a 2.5 kg dumbbell. Pilot testing confirmed this weight was challenging yet tolerable across age groups. Movement rate was set at 30 beats per minute (bpm), with 2 s each for concentric and eccentric phase. Humeral rotation was controlled by keeping the thumb parallel to the floor [25]. The test ended if participants failed to: (1) reach 90° abduction; (2) maintain pace; or (3) stay within the required plane.

External rotation capacity was assessed in prone, lying with the arm positioned at 90° abduction, elbow at 90° flexion, and forearm pronated with the palm facing the floor. The examiner initially supported the participant's arm, passively moving the shoulder into maximal passive shoulder external rotation, visually observing the point that medial side of the wrist reached in this position. The examiner passively moved the participant's arm through the required range twice. During testing, the examiner placed a finger at the point where the medial side of the wrist (overlying the ulnar head) would align at maximum external rotation. Another finger was positioned at the region overlaying the olecranon. The participant's upper arm remained unsupported on the treatment couch, during explanation, passive movements, and active testing.

Participants performed the movement actively from maximal external rotation. A repetition was counted each time the ulnar head touched the examiner's finger. Pace was controlled by a 30-bpm metronome, with 2 s for each the concentric and eccentric phase. The test ended if they failed to: (1) achieve full range (ulnar head did not reach the examiner's marker); (2) maintain the set pace; or (3) control the upper arm position (olecranon shifted by > 1 cm) (Fig. 1).

Prior to testing, participants completed a 2-min warmup including shoulder and cervical range of motion, resistance banded shoulder abductions, external rotation, and overhead press [26]. Following this, participants rested for 5 min before testing. Both dominant and non-dominant shoulders were tested. Arm dominance was assessed by participant self-report and defined as the limb primarily used to carry out common activities of daily living (e.g., brushing teeth or writing).

Reliability/validity

Thirty participants took part in the reliability study and 20 in the criterion validity study. Participants attended two testing sessions, separated by one week. During one session, each participant completed the ExtRA assessment (administered by the lead researcher) and isokinetic dynamometry (IKD) shoulder strength testing, with a 30-min rest interval between assessments. In the other session, participants completed the ExtRA assessment twice—once evaluated by the lead researcher and once by an independent physiotherapist—with a 30-min rest period between tests. To assess intra-rater reliability, the lead researcher's findings were compared across the two testing sessions. The order of testing was randomised, and results were withheld from all assessors until data collection was complete to minimise the risk of reporting bias.

When assessing the criterion validity of ExtRA, isokinetic testing was performed using a Biodex System 4 Isokinetic System [27]. The test was completed for scapular plane shoulder abduction capacity (from 0 to 90° abduction) and external rotation capacity at 90° abduction (from maximal external rotation-0° internal rotation). The Biodex System 4 used 'Passive Mode' at 30°/s for 20 repetitions [27]. For the external rotation capacity testing, participants were instructed to maximally push into external rotation and then to maximally resist the lever returning to the starting position. A similar instruction was given when monitoring abduction capacity: pull up as hard as possible and try to prevent the arm moving down. Prior to maximal testing, the participant was allowed to complete 5 practice repetitions. They were then given a 2-min break between the practice repetitions and the maximal test, and a 5-min rest between isokinetic testing in different positions [4].

Pearson correlation analyses examined the relationship between shoulder strength measures obtained by IKD (concentric/ eccentric values of peak torque, average peak torque, total work, work fatigue and average power) and ExtRA scores assessed by the lead researcher [4,28].

Inter- and intra-rater reliability were evaluated using Bland–Altman analysis. The mean bias (d) was calculated to assess systematic differences between paired measurements, and the 95 % limits of agreement (LoA; $d \pm 1.96 \times SD$) were used to quantify random error. This approach quantifies both systematic bias and random error between raters or sessions. [29,30].

There is no universally accepted threshold for defining "acceptable" bias or limits of agreement in Bland–Altman analysis, as this depends on the scale of measurement and the clinical or functional implications of observed differences, [30] Previous studies of strength and functional performance testing have emphasised the importance of interpreting measurement error in absolute terms, relative to the typical range of scores, rather than as a fixed percentage. In this context, the limits of agreement should be considered in relation to the expected performance range to determine whether the observed differences are meaningful for practical or clinical interpretation [31,32].

Normative data

Data was screened following collection, this included checking the ExtRA scores for abduction and external rotation (dominant and non-dominant sides) for normal distribution using Shapiro-Wilk tests. Data was also screened using visual inspections of scatterplots to identify outliers, relationships between ExtRA scores and age (as a continuous variable).

The data was found to be non-normally distributed ($p < 0.001$) with a positive skew due to a small number of participants scoring much higher than the means of the groups. Consequently, median values are used in reporting normative ranges for the subgroups. Similarly, the reported upper (90th) and lower (10th) bounds of confidence intervals are based on the group median.

Results

Study population

In the reliability study (n = 30), 43 % of the participants were female. The mean age was 37.2 (SD = 17.6) years, ranging from 20 to 65. The mean weight was 75.5 kg (SD = 16.8 kg) and mean height was 174.4 cm (SD = 10.4 cm). Of the participants in the reliability study, 53% self-reported adherence to the WHO Activity Guidelines. [23].

In the criterion validity study (n = 20), 40 % of the participants were female. The mean age was 28 (SD = 13.3) years, ranging from 20 to 58. The mean weight was 78.9 kg (SD = 18.1 kg) and mean height was 173.9 cm (SD = 11.6 cm). Among participants in the criterion validity study, 60 % self-reported meeting the WHO Activity Guidelines [23].

In the normative database study (n = 344), 53 % of the participants were female with a mean age of 53.5 (SD = 20.1) years, ranging from 20 to 90. The mean weight was 74.5 kg (SD = 16.1 kg) and mean height was 170 cm (SD = 9.8 cm). Self-reported compliance with the WHO Activity Guidelines was noted in 54 % of participants in the normative database study [23].

Reliability

Bland–Altman analysis of inter-rater reliability demonstrated minimal bias between raters across all measures (Fig. 2). In the abduction task, the mean bias between raters was 1.07 repetitions for the dominant arm and 0.82 repetitions for the non-dominant arm. The 95 % limits of agreement showed that the second rater's measurements were within − 9.56 to + 11.69 repetitions of the first rater's for the dominant arm, and within − 11.46 to + 13.06 repetitions for the non-dominant arm. For the external rotation task, the limits of agreement ranged from − 6.32 to + 6.89 repetitions for the dominant arm and from − 5.76 to + 7.01 repetitions for the non-dominant arm, with mean biases of 0.28 and 0.63 repetitions, respectively.

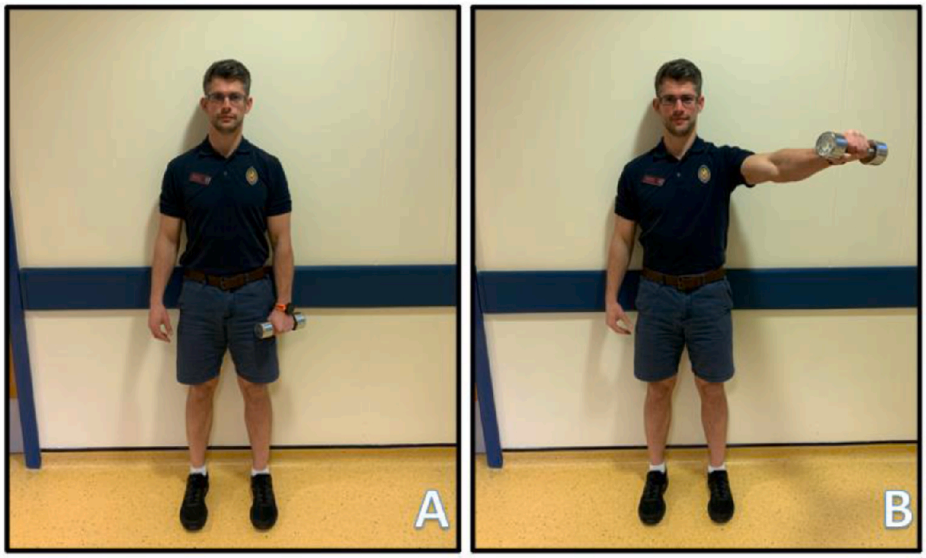
When considering intra-rater reliability for the abduction task, the mean bias was 0.17 repetitions for the dominant arm and − 0.38 repetitions for the non-dominant arm. On 95 % of occasions, repeated measurements by the same rater were within − 7.46 to + 7.79 repetitions for the dominant arm and − 6.42 to + 5.65 repetitions for the non-dominant arm. For the external rotation task, the limits of agreement ranged from − 5.56 to + 5.04 repetitions for the dominant arm and from − 4.14 to + 2.91 repetitions for the non-dominant arm, with small mean biases of − 0.26 and − 0.67 repetitions, respectively. This data is displayed in Supplementary material 1.

Validity tests

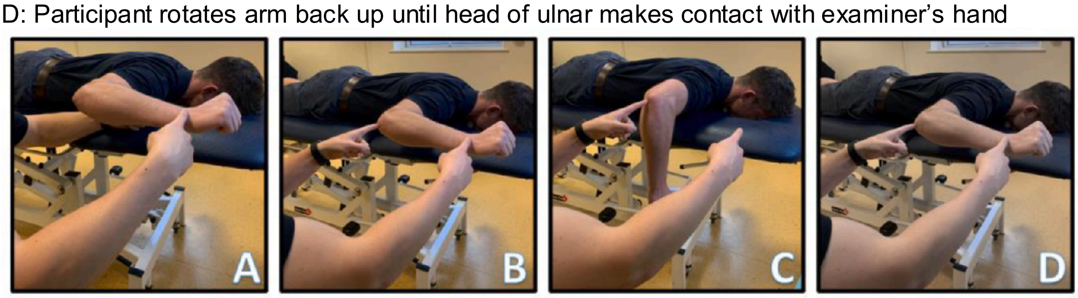
ExtRA abduction measures correlated to 33 of 40 measurements obtained using IKD ($p < 0.05$). Six IKD measures – concentric abduction total work and average concentric abduction power for the dominant and non-dominant arm, concentric average peak abduction torque for the non-dominant arm and dominant external rotation concentric peak torque – had a strong correlation with ExtRA abduction scores (0.7–0.89). Twenty six of the 40 IKD strength measures demonstrated moderate correlation to the abduction strength measure (0.50–0.69) and one showed low correlation (0.26–0.49) [33].

For external rotation, only 2 of the 40 measurements of IKD correlated with ExtRA external rotation ($p < 0.05$).

ExtRA Abduction Capacity Test
A: Starting position of capacity test
B: Participant raised 2.5kg dumbbell to 90° abduction in scapular plane



ExtRA External Rotation Capacity Test
A: Examiner to support participant's upper arm so it is in line with their upper back and noting position of the head of the ulnar when in maximal external rotation
B: Examiner to mark olecranon
C: Participant internally rotates arm around point marked on olecranon until fingers point to the floor
D: Participant rotates arm back up until head of ulnar makes contact with examiner's hand



Demonstration of the area that the researcher's index finger, marking the olecranon, should remain in during the ExtRA External Rotation Capacity Test

Fig. 1. Step by step instructions on how to complete The ExtRA Capacity Test for abduction and external rotation testing.

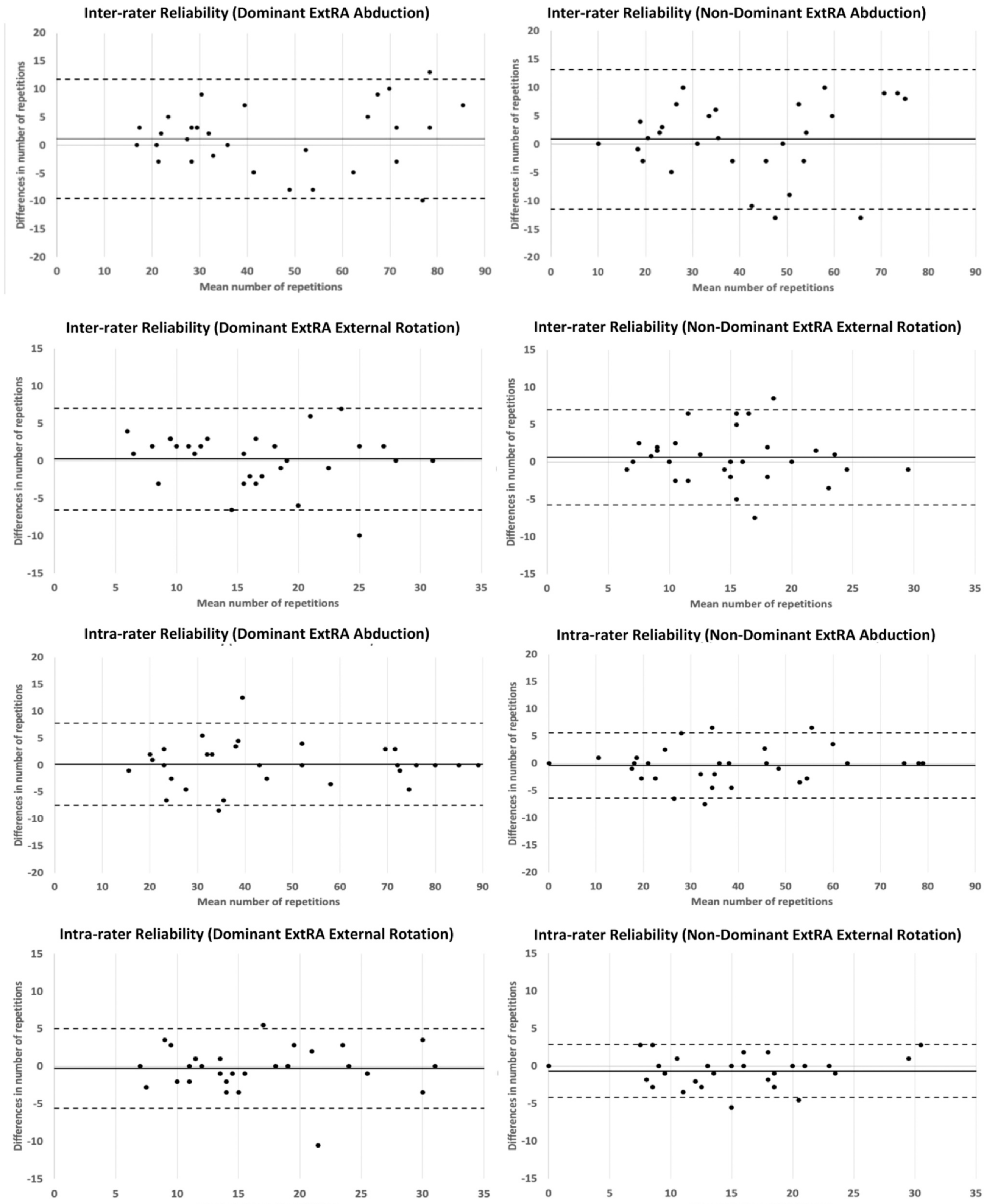


Fig. 2. Bland-Altman plots illustrating inter-rater and intra-rater reliability for the ExtRA Abduction and External Rotation Tests. The dark solid line represents the bias (mean difference) in measurements, while the dashed lines indicate the limits of agreement.

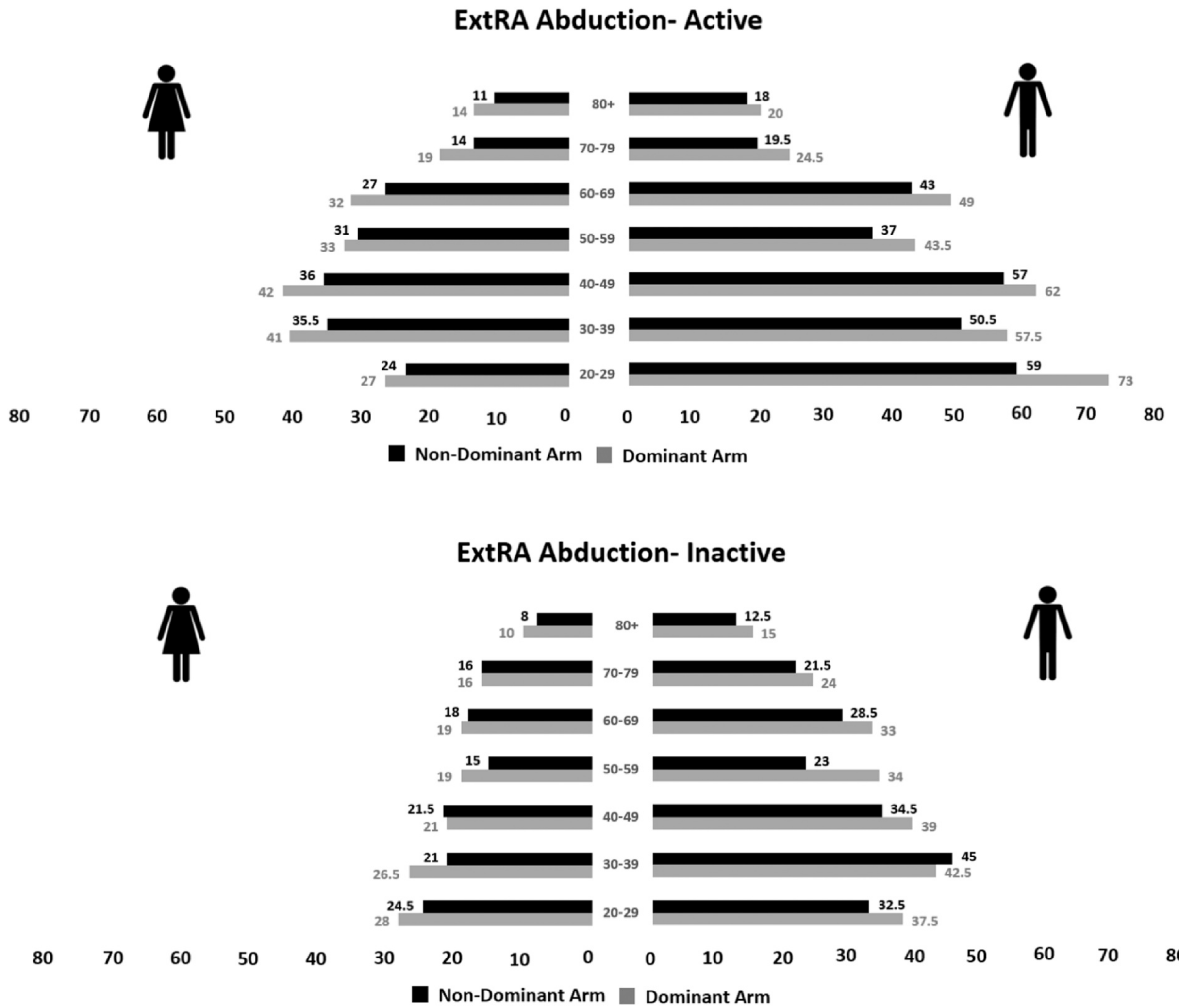


Fig. 3. Histogram displaying median repetitions of the abduction ExtRA capacity test by sex, age group, arm dominance, and whether they do/ do not meet the WHO physical activity guidelines.

The 2 measurements were non-dominant concentric abduction total work and non-dominant concentric external rotation peak torque. However, both demonstrated a low correlation (0.26–0.49) [33]. A table showing the results for all the correlational analyses for the validation study is available in the supplementary material (Supplementary material 2).

Normative values

Figs. 3 and 4 demonstrates the median number of repetitions of the ExtRA Capacity Test by age group, dominant versus non- dominant arm, and WHO physical activity guideline compliance. This data is also available to view in Supplementary materials 3 and 4.

Discussion

The Bland–Altman analysis indicated that the 95 % limits of agreement for the abduction measure were approximately ± 13 repetitions for inter-rater reliability and ± 8 repetitions for intra-rater reliability. According to this approach, the limits of agreement define the range within which most differences between repeated measures are expected to lie if no true change/difference has occurred. Therefore, only differences that exceed these limits can be interpreted as reflecting a real change in the underlying construct, rather than random error or variability inherent to the measurement process.

When this measurement error is interpreted in relation to the range of repetitions achieved, between 3 and 145, the inter-rater limits of agreement of 13 repetitions, indicate

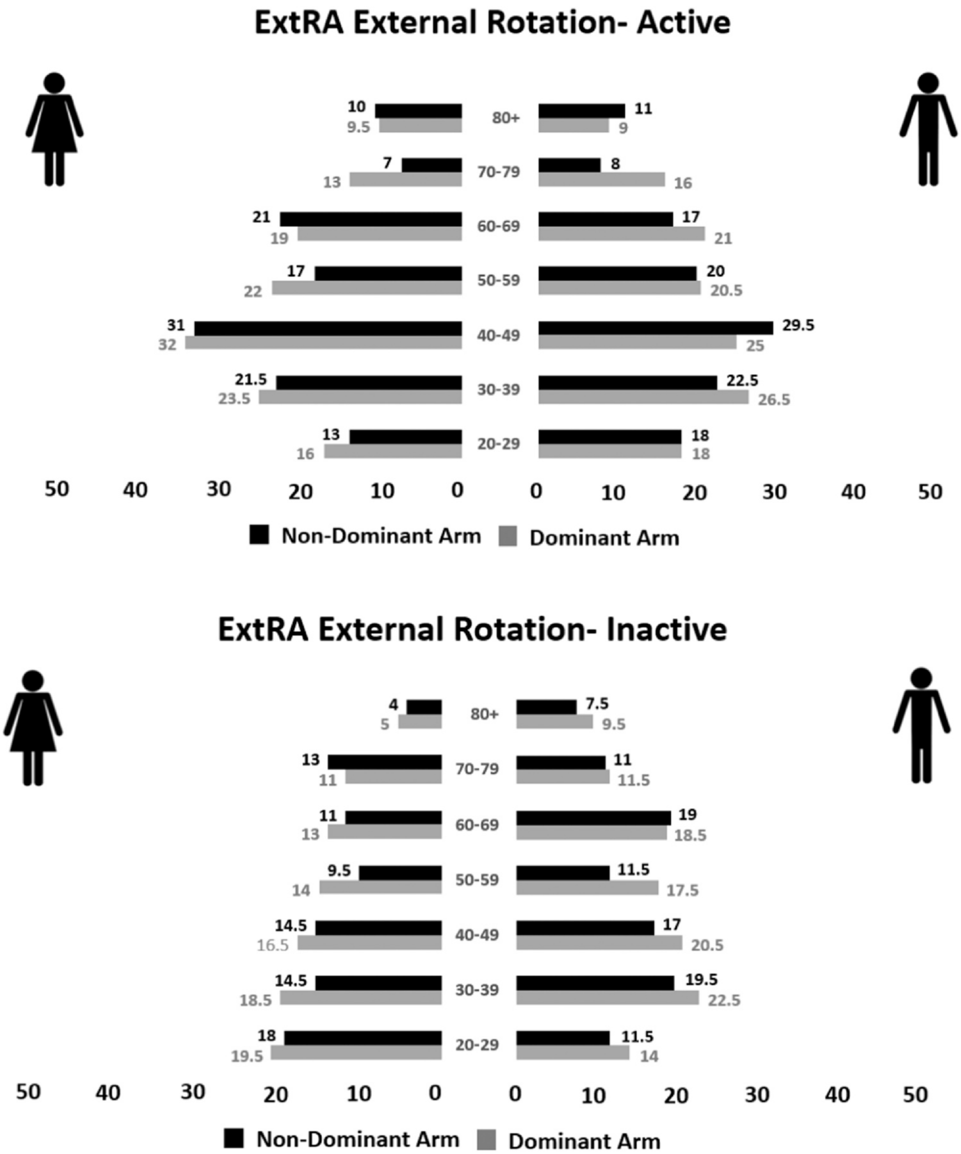


Fig. 4. Histogram displaying median repetitions of the external rotation ExtRA capacity test by sex, age group, arm dominance, and whether they do/ do not meet the WHO physical activity guidelines.

that for participants achieving fewer repetitions, the error could represent a substantial proportion of their performance. In contrast, among individuals with higher repetition counts, the relative impact of this measurement error would be less significant.

For the external rotation measure, the 95 % limits of agreement were approximately ± 7 repetitions for inter-rater reliability and ± 5.5 repetitions for intra-rater reliability. Accordingly, changes exceeding 7 repetitions when assessed by two independent raters, or greater than 5.5 repetitions when assessed by the same rater, can be interpreted as representing a true change/difference, rather than variability attributable to measurement error. Again, interpreting this in light of the scores achieved (1–80 repetitions), for those with lower scores this error could be significant, although less so in those with higher scores. Previous studies investigating alternative shoulder assessment tools, such as hand-held dynamometry and the Closed Kinetic Chain Upper Extremity Stability Test (CKCUEST), have reported absolute reliability metrics comparable to those observed for the ExtRA measures [34,35]. Unlike ExtRA, these tests either do not provide functional measures of shoulder performance or are tailored for specific athletic cohorts. The normative database showed that ExtRA effectively captures differences in shoulder muscle performance across sex, activity levels, and age groups.

Measured agreement between the ExtRA and IKD assessed the criterion validity. The abduction component of ExtRA correlates closely to the strength values obtained using IKD, both for concentric and eccentric abduction and external rotation strength. Contrastingly, ExtRA external rotation poorly correlated to shoulder strength values obtained using IKD. The result may be because whilst IKD measured rotational strength using a supported upper arm, ExtRA external rotation is performed with an unsupported upper arm, which means that rotator cuff recruitment is likely to be different [36]. Poor correlation may also be because ExtRA external rotation is limited by poor movement control of the shoulder rather than muscular fatigue. This may suggest that the test measures an aspect of shoulder function that is not currently captured by other measures or IKD. Muscular control training has been found to reduce pain and increase function for patients with shoulder pain and thus a measure of this aspect of muscle performance for the shoulder may be useful [9,10,36].

The results from the large cohort of asymptomatic participants provide population-based normative values, which can be used as a point of reference in a clinical setting when assessing the functional capacity of individuals. Older age, female sex, and not achieving WHO activity guidelines negatively impact ExtRA performance in both abduction and external rotation tests. Similar age-related patterns for the median number of repetitions for the abduction measure is seen for both sexes. Active individuals perform significantly better until ages 70–79, after which performance drops, aligning with previous findings of reduced shoulder strength above the age of 65 [37]. It is unclear why the abduction data demonstrate no significant change in performance in the active group compared to the inactive group over the age of 70. One possible explanation is that older adults are more likely to engage in submaximal or lower intensity resistance training. Thus, despite reporting that they meet the WHO activity guidelines, their activity level may not induce physiological adaptations for improved shoulder muscle performance during maximal testing.

Median scores in the external rotation capacity test show that men and women who meet WHO Activity Guidelines outperform those who do not, especially in younger age groups. Beyond ages 60–69 in men and 70–79 in women, achieving activity guidelines does not significantly impact performance. In both sexes, performance peaks in the active group in the 40–49 age group with a gradual decline observed from 20 to 29 and 80+ in the inactive group (Fig. 3, Supplementary material 3/4). The peak in median performance in the 40–49-year active groups is likely a result of high movement literacy in combination with minimal age-related decline in muscular performance [38]. By their 40s, physically active individuals likely accumulate experience that refines neuromuscular coordination. In older groups, performance is limited by gradual declines in muscle mass, strength, and neuromuscular control [38].

Limitations

One of the study limitations is that participants' activity prior to testing was not controlled. Therefore, the scores of some participants may have been affected by their engagement in strenuous exercise, potentially leading to fatigue. This nonetheless reflects a realistic scenario that a healthcare practitioner might encounter in a clinical setting.

Unlike other database studies which use an activity level scale [13,14], activity levels were classified into either 'achieves WHO activity Guidelines' or 'does not achieve WHO activity guidelines' in order to ensure resistance training was accounted for. The classification into subgroups based on these criteria results in considerable variance. For example, someone meeting the minimum WHO guidelines is grouped with an elite strength athlete doing daily resistance training. This leads to many high-performing outliers in the raw data, causing a positive skew in the active group.

Conclusion

This study introduces ExtRA, an external rotation and abduction capacity test for the shoulder. The test is devised for use in clinical settings and can provide a functional insight into shoulder muscle performance specific to a real world or sporting environment. The normative database, derived from a large sample of asymptomatic individuals, enables clinicians to assess shoulder muscle performance by comparing individual results to reference values matched for age, sex, and physical activity level. An estimate of the error of measurement has been given, and within the caveats given, has reasonable reliability. It can therefore be useful in monitoring performance over time whereby changes in ExtRA performance are likely to indicate changes in functional capacity.

CRediT authorship contribution statement

HF designed the data collection tools, held the role of lead researcher during data collection. He was involved in data collection, processing and analysis, and also drafted and revised the paper. He is guarantor. JL was influential in the design of the study as well as drafting and revising the paper. VT, WB, YL, RS, JR assisted with data collection. MD and MO were involved in data analysis and drafted and revised the paper. DL was influential in participant recruitment, statistical analysis of the data and also drafted and revised the paper. AS acted as the lead supervisor for this project. She was involved in study design, participant recruitment and also drafted and revised the paper.

Ethical Approval

Ethical approval was granted by the Brunel University of London, College of Health, Medicine and Life Sciences Research Ethics Committee (35944-LR-Mar/2022-38654-1).

Funding

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

Conflict of interest

The authors declare no conflicts of interest.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.physio.2025.101868.

References

[1] Campbell EL, Seynnes OR, Bottinelli R, *et al.* Skeletal muscle adaptations to physical inactivity and subsequent retraining in young men. Biogerontology 2013;14(3):247–59.

[2] Greenhaff PL. The molecular physiology of human limb immobilization and rehabilitation. Exerc Sport Sci Rev 2006;34(4):159–63.

[3] Cools AM, Maenhout AG, Vanderstukken F, *et al.* The challenge of the sporting shoulder: from injury prevention through sport-specific rehabilitation toward return to play. Ann Phys Rehabil Med 2020:101384.

[4] Bagordo A, Ciletti K, Kemp-Smith K, *et al.* Isokinetic dynamometry as a tool to predict shoulder injury in an overhead athlete population: a systematic review. Sports 2020;8(9):124.

[5] Osternig LR. Isokinetic dynamometry: implications for muscle testing and rehabilitation. Exerc Sport Sci Rev 1986;14:45–80.

[6] Dollings H, Sandford F, O'Conaire E, *et al.* Shoulder strength testing: the intra- and inter-tester reliability of routine clinical tests, using the PowerTrackTM II Commander. Shoulder Elb 2012;4(2):131–40.

[7] Nagatomi T, Mae T, Nagafuchi T, *et al.* Shoulder manual muscle resistance test cannot fully detect muscle weakness. Knee Surg Sports Trauma Arthrosc 2017;25(7):2081–8.

[8] Kollock R, Van Lunen BL, Ringleb SI, *et al.* Measures of functional performance and their association with hip and thigh strength. J Athl Train 2015;50(1):14–22.

[9] Ginn K, Cohen M. Exercise therapy for shoulder pain aimed at restoring neuromuscular control: a randomized comparative clinical trial. J Rehabil Med 2005;37(2):115–22.

[10] Magarey ME, Jones MA. Dynamic evaluation and early management of altered motor control around the shoulder complex. Man Ther 2003;8(4):195–206.

[11] Reed D, Cathers I, Halaki M, *et al.* Does changing the plane of abduction influence shoulder muscle recruitment patterns in healthy individuals? Man Ther 2016;21:63–8.

[12] Culvenor AG, Collins NJ, Guermazi A, *et al.* Early patellofemoral osteoarthritis features one year after anterior cruciate ligament reconstruction: symptoms and quality of life at three years. Arthritis Care Res 2016;68(6):784–92.

[13] Hébert-Losier K, Wessman C, Alricsson M, *et al.* Updated reliability and normative values for the standing heel-rise test in healthy adults. Physiotherapy 2017;103(4):446–52.

[14] Bento TP, Genebra CV, Cornélio GP, Biancon RD, Simeão SF, Vitta AD. Prevalence and factors associated with shoulder pain in the general population: a cross-sectional study. Fisioter Pesqui 2019;26:401–6.

[15] Earhart GM, Cavanaugh JT, Ellis T, *et al.* The 9-hole PEG test of upper extremity function: average values, test-retest reliability, and factors contributing to performance in people with Parkinson disease. J Neurol Phys Ther 2011;35(4):157–63.

[16] Olds M, Coulter C, Marant D, *et al.* Reliability of a shoulder arm return to sport test battery. Phys Ther Sport 2019;39:16–22.

[17] Hughes RE, Johnson ME, O'Driscoll SW, *et al.* Age-related changes in normal isometric shoulder strength. Am J Sports Med 1999;27(5):651–7.

[18] Lewis JS. Rotator cuff tendinopathy/subacromial impingement syndrome: is it time for a new method of assessment? Br J Sports Med 2009;43(4):259–64.

[19] Littlewood C, Bateman M, Connor C, *et al.* Physiotherapists' recommendations for examination and treatment of rotator cuff-related shoulder pain: a consensus exercise. Physiother Pract Res 2019;40(2):87–94.

[20] Richardson E, Lewis JS, Gibson J, *et al.* Role of the kinetic chain in shoulder rehabilitation: does incorporating the trunk and lower limb into shoulder exercise regimes influence shoulder muscle recruitment patterns? A systematic review of electromyography studies. BMJ Open Sport Exerc Med 2020;6(1).

[21] Reinold MM, Wilk KE, Fleisig GS, *et al.* Electromyographic analysis of the rotator cuff and deltoid musculature during common shoulder external rotation exercises. J Orthop Sports Phys Ther 2004;34(7):385–94.

[22] Tardo DT, Halaki M, Cathers I, *et al.* Rotator cuff muscles perform different functional roles during shoulder external rotation exercises. Clin Anat 2013;26(2):236–43.

[23] World Health Organization. WHO guidelines on physical activity and sedentary behaviour: at a glance. 2020.

[24] Phadnis J, Trompeter A, Gallagher K, *et al.* Mid-term functional outcome after the internal fixation of distal radius fractures. J Orthop Surg Res 2012;7(1):1–8.

[25] Coratella G, Tornatore G, Longo S, *et al.* An electromyographic analysis of lateral raise variations and frontal raise in competitive bodybuilders. Int J Environ Res Public Health 2020;17(17):6015.

[26] Wattanaprakornkul D, Cathers I, Halaki M, *et al.* The rotator cuff muscles have a direction-specific recruitment pattern during shoulder flexion and extension exercises. J Sci Med Sport 2011;14(5):376–82.

[27] Harding FV, Ghatan R, Chaney C, *et al.* Clinical case study: shoulder injury – senior tennis. Am Coll Sports Med Annu Meet 2015.

[28] Ahlgren P, Jarneving B, Rousseau R. Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. J Am Soc Inf Sci Technol 2003;54(6):550–60.

[29] Gerke O. Reporting standards for a Bland–Altman agreement analysis: a review of methodological reviews. Diagnostics 2020;10(5):334.

[30] Hopkins WG. Measures of reliability in sports medicine and science. Sports Med 2000;30(1):1–15.

[31] Królikowska A, Reichert P, Karlsson J, Mouton C, Becker R, Prill R. Improving the reliability of measurements in orthopaedics and sports medicine. Knee Surg Sports Traumatol Arthrosc 2023;31(12):5277–85.

[32] Giavarina D. Understanding bland altman analysis. Biochem Med 2015;25(2):141–51.

[33] Munro BH. Statistical methods for health care research. 5th ed. Lippincott Williams & Wilkins; 2005.

[34] Chamorro C, Arancibia M, Trigo B, Arias-Poblete L, Jerez-Mayorga D. Absolute reliability and concurrent validity of hand-held dynamometry in shoulder rotator strength assessment: systematic review and meta-analysis. Int J Environ Res Public Health 2021;18(17):9293.

[35] Kardor S, Gorji Z, Ghotbi N, Attarbashi-Moghadam B, Shadmehr A, Gorji M. Upper extremity physical performance tests in female overhead athletes: a test–retest reliability study. J Orthop Surg Res 2023;18(1):489.
[36] Worsley P, Warner M, Mottram S, *et al.* Motor control retraining exercises for shoulder impingement: effects on function, muscle activation, and biomechanics in young adults. J Shoulder Elb Surg 2013;22(4).
[37] Pike JM, Singh SK, Barfield WT, *et al.* Impact of age on shoulder range of motion and strength. JSES Int 2022;6(6):1029–33.
[38] Mendias CL, Roche SM, Harning JA, *et al.* Reduced muscle fiber force production and disrupted myofibril architecture in patients with chronic rotator cuff tears. J Shoulder Elb Surg 2015;24(1):111–9.