Academic Year
2023 - 2024

Reliability of the Athletic Shoulder test in overhead racquet athletes

Michelle Schellekens

Jonas Pittoors

Noa Verheyen

Master Thesis
Master of Rehabilitation Sciences and Physiotherapy

Promotor
Prof. dr. Filip Struyf

Supervisor
Claudia Cavaggion, PhD student

Acknowledgements

The completion of this reliability study was made possible through the collaborative efforts of Michelle Schellekens, Jonas Pittoors, and Noa Verheyen. We extend our sincere gratitude to our promotor, Prof. Dr. Filip Struyf, for affording us the invaluable opportunity to undertake this research and for providing guidance at every stage of its development.

Special thanks are due to Jeroen Beeusaert and Joren Huygen for their significant contributions during the methodological elaboration of our research, as well as for their assistance and training in the operation of essential software tools. Our appreciation also goes to Julien Prudhomme, co- owner of LAB Antwerp, for graciously providing the facilities necessary for conducting the measurements.

Our greatest appreciation goes to our supervisor, Ms. Claudia Cavaggion, for her consistent guidance, support, and patience throughout the entire two-year process. Finally, our acknowledgment is reserved for the subjects who generously allocated time for testing, and to our friends and family for their mental support and encouragement throughout this journey.

Table of content

List of tables and figures

List of abbreviations

RC	Rotator cuff
ROM	Range of motion
ER	External rotation
HHD	Hand-held dynamometer
ASH test	Athletic Shoulder test
GRRAS	Guidelines for Reporting Reliability and Agreement studies
VAS	Visual Analogue Scale
SPADI	Shoulder Pain and Disability Index
ICC	Intraclass Correlation Coefficient
SEM	Standard Error of Measurement
MDC_{95}	Minimal Detectable Change at 95% confidence level
M	Male
F	Female
BMI	Body Mass Index
N/A	Not applicable
Av. training volume	Average number of hours training per week
Max. training volume	Maximal number of hours training per week
Dom	Dominant arm
NDom	Non-dominant arm
N	Newton

Positioning

This reliability study was the second part of a two-year-long master's thesis for the degree of Master by three students during the academic year 2023-2024: Michelle Schellekens, Jonas Pittoors, and Noa Verheyen. All authors worked collaboratively throughout the entire process.

This thesis is promoted by Professor Doctor Filip Struyf and supervised by Miss Claudia Cavaggion. Pr. Dr. Filip Struyf is professor at the University of Antwerp (Belgium), where he teaches physiotherapy and general medicine students about the assessment and treatment of shoulder problems. He is also active as a physiotherapeutic consultant. Furthermore, expertscape[1] has awarded him as worldwide leading expert in shoulder pain. Miss Claudia Cavaggion is currently performing Ph.D. research at the University of Antwerp. Her expertise in research concerning the shoulder has been instrumental in shaping the research questions and methodology used in this review.

Both parts of this master thesis topic are placed within a broader context of overhead athletes with a specific interest of the shoulder. This second part focuses on the reliability of a new recent shoulder test[2] within an asymptomatic and symptomatic population of tennis, badminton and padel athletes. There is a need of reliable tests for assessing strength of the shoulder. This study aims to fill this gap and provide a reliable shoulder test to perform in clinical practice.

This reliability study was part of a larger investigation also involving another master thesis conducted by three other physiotherapy students. Their research focused on the shoulder endurance test with the same population of athletes. Some parameters (e.g. arm length) mentioned in the method/results of this thesis were measured and reported but mainly relate to the data analysis of the other master thesis.

Abstract in Dutch and keywords

Doelen: De intra- en interbeoordelaarsbetrouwbaarheid van de Athletic Shoulder test (ASH-test) onderzoeken, op een krachtplaat, in een populatie van asymptomatische en symptomatische bovenhandse racketatleten.

Ontwerp: Een test-hertest betrouwbaarheidsstudie werd uitgevoerd.

Setting: De metingen vonden plaats bij LAB Antwerp, een kinesitherapiepraktijk.

Deelnemers: Tweeënveertig bovenhandse racketatleten namen deel aan deze studie, waaronder 21 asymptomatische atleten en 21 symptomatische atleten met schouderklachten. Alle deelnemers speelden minimaal vier uur per week tennis, badminton en/of padel.

Uitkomstmaten: Intraklasse correlatiecoëfficiënten (ICC) werden berekend om intra- en interbeoordeelaarsbetrouwbaarheid te bepalen bij asymptomatische en symptomatische atleten, alsook de analyse van de standaardfout (SEM), minimaal waarneembare verandering (MDC) en metingen van overeenkomst (Bland-Altman grafieken).

Resultaten: De ASH-test toonde uitstekende intrabeoordelaarsbetrouwbaarheid voor zowel de dominante als de niet-dominante schouder in asymptomatische (ICC 0.93-0.98) en symptomatische atleten (ICC 0.91-0.98), met een interbeoordelaarsbetrouwbaarheid die varieerde van goed tot uitstekend voor beide groepen (asymptomatisch: ICC 0.80-0.95, symptomatisch: ICC 0.79-0.93).

Conclusies: De ASH test vertoonde uitstekende betrouwbaarheid wanneer deze werd uitgevoerd door één beoordelaar, wat wijst op het potentieel voor precieze isometrische krachtbeoordeling bij bovenhandse racketatleten, met en zonder schoudersymptomen. Ondanks de variërende interbeoordeelaarsbetrouwbaarheid, gaande van goed tot uitstekend, behaalden enkele specifieke testposities een uitstekende betrouwbaarheid, wat hun klinische bruikbaarheid bevestigt. Toekomstige toepassingen in de sportgeneeskunde kunnen onder meer blessurerevalidatie en verbetering van protocollen voor return-to-play omvatten.

Keywords: Reliability; Shoulder Assessment; overhead racquet athlete; Athletic Shoulder test; Isometric Strength; Force Plate.

Abstract and keywords

Objectives*:* To assess the intra- and inter-rater reliability of the Athletic Shoulder test, using a force plate, in a population of asymptomatic and symptomatic overhead racquet athletes.

Design*:* A test-retest reliability study was conducted.

Setting: The measures took place at LAB Antwerp within a physiotherapy practice environment.

Participants: Forty-two overhead racquet athletes participated in this study, including 21 asymptomatic athletes and 21 symptomatic athletes with shoulder complaints. All participants were actively engaged in tennis, badminton, and/or padel for a minimum of four hours weekly.

Outcome measures*:* Intraclass correlation coefficients (ICC) were calculated to assess the intra- and inter-rater reliability in asymptomatic and symptomatic athletes, along with analysis of standard errors of measurement (SEM), minimum detectable change (MDC) and measurements of agreement (Bland-Altman plots).

Results: The ASH test demonstrated excellent intra-rater reliability for both dominant and non-dominant shoulders in asymptomatic (ICC 0.93-0.98) and symptomatic athletes (ICC 0.91-0.98), with inter-rater reliability ranging from good to excellent for both groups (asymptomatic: ICC 0.80-0.95, symptomatic: ICC 0.79-0.93).

Conclusions*:* The ASH test demonstrated excellent reliability when conducted by a single novice rater, indicating its potential for precise isometric strength assessment in overhead racquet athletes, with and without shoulder complaints. Although inter-rater reliability varied from good to excellent, certain test positions showed excellent reliability, confirming their clinical utility. The test's potential future applications in sports medicine encompass assistance in injury rehabilitation and the improvement of return-to-play protocols.

Keywords: Reliability; Shoulder Assessment; overhead racquet athlete; Athletic Shoulder test; Isometric Strength; Force Plate.

Introduction

Shoulder injuries are notably prevalent among athletes engaging in racquet sports such as tennis, badminton, and padel[3, 4]. Specifically, among tennis players, shoulder injuries rank as the most common upper extremity injuries[5]. Similarly, in badminton, the shoulder stands out as the most affected upper limb region[3, 6, 7]. In padel, a notable 37.5% of all injuries are localised to the upper limb, largely attributed to the sport's emphasis on overhead strokes like smashes[8]. The spectrum of injuries observed in this population encompasses a wide range, from mild discomfort to more severe conditions including rotator cuff lesions and tears, glenohumeral capsular lesions, superior labrum from anterior to posterior tears, osseous lesions, and injuries to the biceps tendon[9].

This high prevalence of upper extremity injuries in racquet sports is attributed to the complex biomechanics involved in overhead throwing motions, which exert significant forces on the shoulder joint[10, 11]. For example, during the late cocking phase of a tennis serve, the shoulder joint undergoes maximal external rotation (ER), along with stretching and eccentric loading of the muscles. This is followed by phases of acceleration and deceleration, ultimately leading to maximum internal rotation of the shoulder joint. These repetitive, forceful overhead actions heighten the risk of shoulder overuse injuries[12, 13]. Therefore, achieving an optimal balance between the functional mobility and stability of the shoulder is essential[9, 14].

Although it is well-known that overhead racquet sports exert considerable strain on athletes' shoulders—impacting their ability to train and compete—this group of athletes is underrepresented in scientific literature[5]. Current research predominantly focuses on overhead athletes in sports such as American football, rugby, baseball, and water polo. Moreover, the rising popularity of sports such as padel underscores the urgent need for targeted research focused on this distinct group of athletes. Therefore, addressing this literature gap with focused research on overhead racquet athletes is crucial.

In Master Thesis Part 1, the association between alterations in shoulder range of motion (ROM) and the development of shoulder complaints in overhead athletes was explored. The findings were heterogeneous, with the identification of only one significant predictor: a deficit in ER ROM of the dominant shoulder. This deficit, defined as a dominant shoulder ER ROM less than 5° greater than that of the non-dominant shoulder, was reported in only one low-quality study, involving a small sample size of baseball players, resulting in a level 3 strength of conclusion[15]. However, a 5° difference in ROM might hold minimal clinical relevance[16] and most overhead athletes typically exhibit an increased ER ROM in their throwing shoulder as a natural adaptation, rather than a deficit, without a corresponding increase in shoulder complaints[17, 18]. Therefore, the recommendation was to avoid excessive focus on ROM and instead prioritise the exploration of other factors (Master Thesis, Part 1). Factors such as shoulder strength might play a more substantial role in predicting athletes at risk, making them a more relevant area for investigation[16, 19, 20]. With limited existing literature on strength assessment in overhead racquet athletes, it is essential to explore tools for directly assessing strength and to establish reliable screening tools in this population[21, 22].

While numerous tests for lower limb strength and function are valid and reliable, options for evaluating upper limb strength and function are relatively limited[23, 24]. Despite their proben reliability for strength assessments, the use of isokinetic devices is hindered by their high costs, extensive setup times, and significant space requirements, reducing their utility in clinical setings[25, 26]. Moreover, isokinetic movements are not commonly performed by upper extremity muscles in everyday activities[27]. Yen et al.[28] observed that in a patient population, some subjects are unable to reach the preset velocity, highlighting another limitation. Hand-held dynamometers (HHD) offer a more cost-effective solution, with proven reliability and validity for shoulder strength assessments[28, 29]. However, the force plate remains more reliable and is often considered the gold standard[2], thus making it the preferred measurement tool for the present study.

The Athletic Shoulder (ASH) test is an isometric shoulder test which was firstly created to measure force production and assess potential interlimb or post-injury deficits in elite-rugby players specific to their sport's shoulder positions[2]. This functional test targets long-lever positions at 180°, 135°, and 90° shoulder abduction. Ashworth et al.[2] recognised the ASH test as a valuable tool for rehabilitation and injury prevention, offering a comprehensive assessment of the athletes' ability to meet the functional demands of their sport[2]. Further validation of the ASH test's applicability was provided by Trunt et al.[30], who demonstrated its effectiveness in identifying bilateral shoulder strength adaptations in healthy baseball pitchers. Specifically, they observed significant differences in isometric peak force between the dominant and non-dominant arms at 90° and 135° shoulder abduction, aligning with the shoulder positions during pitching[30]. The ASH test could be valuable in overhead racquet sports as it replicates the torque and force from extended arm movements, such as the ball impact of a tennis serve[2]. By using force plates and various test positions, the ASH test can provide a valid and objective assessment of isometric strength in the upper extremities.

To date, the reliability of the ASH test in overhead racquet athletes remains unexplored in current literature. Moreover, most research on strength assessments predominantly focuses on healthy subjects, thus omitting data on symptomatic populations[27]. Bridging this knowledge gap, the current study marks the first investigation into the intra- and inter-rater reliability of the ASH test performed on a force plate in both asymptomatic and symptomatic populations of overhead racquet athletes. To our knowledge, this is also the first study to examine the reliability of a straight-arm strength test on a force plate in such population. By including a representative population, this study seeks to enrich the applicability of its findings for clinical use, providing valuable insights into the test's utility for healthcare professionals, even those with limited experience, following brief pilot training[26]. Demonstrating the ASH test's reliability could make it a valuable addition to sports medicine professionals' toolkit for assessing and managing the health of athletes.

Methodology

This reliability study was performed following the Guidelines for Reporting Reliability and Agreement studies (GRRAS)[31]. This study assessed the intra- and inter-rater reliability of the ASH test among both asymptomatic and symptomatic overhead racquet athletes. Approval was received from the Committee of the Antwerp University Hospital (ref: B3002023000165). Methods were performed in accordance with the Declaration of Helsinki and all participants provided written informed consent. The study took place at the physiotherapy practice LAB in Antwerp, which provided the expertise and facilities to conduct the study. Measurements were performed by M.S., N.V. and J.P., all 2nd Master physiotherapy students at the University of Antwerp.

Preparatory phase

Rater A (M.S.) and rater B (N.V.) lacked prior experience in using the force plate and the data analysis software. To address this, a preparatory phase was implemented, which included both training and pilot testing. A two-hour training session took place in November 2023 at LAB Antwerp. Sport performance coach J.H. with four years of experience in training athletes, and two years of expertise in operating the force platform and analysing data, taught the raters how to use the system. Subsequently, a pilot test was conducted on the same day, under the supervision of J.H., to ensure consistency between raters and verify the accuracy of data collection and interpretation. Two overhead racquet athletes participated in this pilot phase, one symptomatic shoulder patient and the other asymptomatic. Following the successful completion of the pilot phase, the main reliability study was initiated at the same physiotherapy practice and extended over four weeks.

Participants

Due to the similar overhead movement patterns in tennis, badminton, and padel, athletes engaged in these sports were recruited for this study. A convenience sample of both asymptomatic and symptomatic participants was recruited through social media, sport clubs, email, and among friends and colleagues of the investigators. Both male and female participants were eligible if they met the following criteria: aged between 18 and 60 years, engaged in a minimum of 4 hours of play per week, and participated in competitive or recreational levels within the specified sports. The distinction between recreational and competitive athletes was delineated following the criteria set forth by McKinney et al.[32], wherein recreational athletes are those allocating at least four hours per week to racquet sports for personal and fitness purposes, while competitive athletes devote more than six hours weekly to training designed for performance improvement and participation in official competitions. Participants were excluded if they had experienced traumatic injury or surgery involving the shoulder complex, upper thorax, back, or humerus within the past year; if a known pregnancy was present; or if they were unable to achieve 180° shoulder abduction for assuming the test position. All criteria for eligibility relied on participants' self-reported information.

If athletes reported shoulder pain during racquet sports in the questionnaire (appendix, figure 1), the Visual Analogue Scale[33] (VAS) was used to distinguish between asymptomatic and symptomatic participants. Athletes were defined as symptomatic if they experienced dominant shoulder pain with a reported VAS of at least three out of ten while playing racquet sports within the preceding month, as described in Almeida et al[19]. This classification solely considered the dominant arm, which is primarily utilised in the sport. Athletes with complaints on the non-dominant arm were categorised as asymptomatic. The SPADI-Dutch (Shoulder Pain and Disability Index) version[20], was solely employed to gain deeper insights into the severity of the shoulder complaint. The asymptomatic group included individuals who either experienced no shoulder pain in the dominant arm during racquet sports in the last 30 days or scored less than three out of ten on the VAS. Participants were unaware of this predetermined cut-off value or group categorisation.

Design

For this study, a repeated measures design was implemented. Rater A conducted the first two ASH tests, while rater B administered the last ASH test. The sequence of the six test positions were predetermined and maintained consistently for every athlete, rather than being randomised. This design assessed intra-rater reliability within a single rater (rater A) across two trials (A1 and A2), as well as inter-rater reliability between rater A and rater B. All measurements for each athlete were conducted within a single testing session.

Both raters A and B remained unaware of each other's results. A third non-blinded rater C (J.P.) managed participant intake, group classification (asymptomatic/symptomatic), allocation, and results' tracking to ensure blinding of the raters A and B. Athletes were asked to refrain from participating in overhead racquet sports on the testing day.

Procedure

First, athletes provided signed informed consent with rater C. Next, they completed a baseline questionnaire covering demographics, sport-specific features, and, if applicable, any shoulder existing complaints (appendix, figure 1). After the intake, rater C measured athletes' body weight and provided them with detailed instructions about the ASH procedure. Athletes were then directed to the testing room, ensuring that only one athlete was present at a time in the test zone.

ASH Test procedure on force platform

Athletes performed a maximal isometric strength test on both shoulders on the 'AMTI 3D force plate,' utilising the ASH Test Battery as described by Ashworth et al[2]. Tape marks served as reference points for arm placement (figure 2a). Measurements of both arms were independently taken in 180° abduction (I-position, figure 2b), 135° abduction (Y-position, figure 2c), and 90° abduction (T-position, figure 2d). To prevent compensatory trunk movements, the athlete's non-tested arm was held behind the athlete's back during testing. Irrespective of hand dominance or the side exhibiting symptoms, the test sequence for all participant proceeded as follows: left I-position, right I-position, left Y-position, left T-position, right T-position, right Y-position. Participants were given a minimum rest of 10 seconds between positions[34]. This fixed test sequence constituted one ASH test cycle and was performed three times for each participant. This sequence differed from Ashworth et al.'s original protocol, aiming for enhanced time-management and minimise patient displacements during test[2].

Athletes first underwent two warm-up attempts at 80%-90% intensity for each position, guided by rater A. Subsequently, the three actual ASH tests were performed: the initial and second assessments were administered by rater A (labelled as A1 and A2), while the third test was conducted by rater B. Athletes were instructed to exert maximal force for 3 seconds, gradually increasing force in the first second and then slowly decreasing after completing the test. Standardised instructions were given on positioning, contraction, and relaxation. Tests were repeated if technical malfunctions occurred or if compensatory behaviours such as using lower limbs, anterior scapular tilt, or elbow flexion were observed.

ASH Test outcome

Force data were analysed with Qualisys Track Manager software, which provided graphs of three force vectors (x, y, z), with the analysis focusing on the vertical (z) vector. The average maximum force over a 1-second interval was visually determined from the graphs for each measurement (appendix, figure 3). To ensure blinding, raters independently analysed their own measurement graphs.

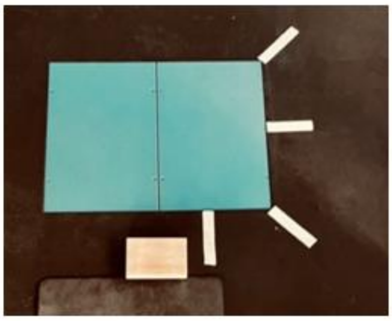


Figure 2a: Set-up of ASH test with tape marks

Figure 2b: ASH test left I-position (180° abduction) .

Figure 2c: ASH test in left Y-position (135° abduction)

Figure 2d: ASH test in left T-position (90° abduction), with rater A correcting the position.

Statistical analysis

A sample size was determined based on specified parameters for a pre-specified intraclass correlation coefficient (ICC) of 0.70, an expected ICC = 0.90, α = 0.05, and power = 0.80, considering two raters. The sample size calculation was based on the review of Bujang et al.[35] and was achieved by focussing on the inter-rater reliability, which is typically lower than intra-rater reliability. Since all measurements took place on the same day, no drop-out range was included in the sample size calculation. A total sample size of 38 participants was determined, with 19 participants in each group.

Data analysis was performed using IBM® SPSS® Version 28. The normality of all variables was assessed through the Kolmogorov-Smirnov test ($p > 0.05$). Differences in patient characteristics between the asymptomatic and symptomatic groups were analysed using unpaired t-tests for normally distributed data, Mann-Whitney U tests for non-normally distributed data, and Fisher's exact test for nominal variables. The interpretation of ICC values was as follows: > 0.90 = excellent, 0.75 to 0.90 = good, 0.50 to 0.75 = moderate, < 0.50 = poor[36].

The ICC model 2,2 (2-way random-effects model—absolute agreement, single measures) was applied to evaluate intra-rater reliability by rater A. To assess inter-rater reliability and generalise findings to raters with similar characteristics, the ICC model 2,3 (2-way random-effects model—absolute agreement, single measures) was employed[37]. Inter-rater reliability was assessed by comparing the measurements from rater A (A1 and A2) independently with the single measurement taken by rater B, resulting in two separate evaluations of inter-rater reliability. When faced with non-normally distributed ASH measurements, a natural logarithm transformation was employed to normalise the data for further analysis. However, since the differences in ICC values between raw and transformed data was minimal (less than 0.04), analysis focused on the non-transformed ICCs. Bland-Altman plots, containing 95% limits of agreement, were used to visually assess the agreement between the two raters and identify any potential bias[38]. Paired t-tests (for normally distributed data) or Wilcoxon signed-rank tests (for non-normally distributed data) were used to examine differences between raters and to confirm the presence of systematic bias when visually evident in the Bland-Altman plots. Additionally, if proportional bias was visually observed in the Bland-Altman plots, regression analysis was conducted. Specifically, the presence of proportional bias was confirmed by observing a non-zero slope in the regression of differences on means. Lastly, the Standard Error of Measurement ($SEM = SD \times \sqrt{(1-ICC)}$) and the Minimal Detectable Change at a 95% confidence level ($MDC_{95} = SEM \times 1.96 \times \sqrt{2}$) were calculated for both intra and inter-rater reliability assessments[39].

Results

A total of 47 overhead racquet athletes were examined, with 5 patients excluded due to insufficient playing hours (<4h/week). Consequently, the final analysis included 42 athletes, evenly distributed between asymptomatic (21) and symptomatic (21) athletes. No measurements were lost during data acquisition or processing. Significantly, the symptomatic group exhibited a notably higher maximal training volume (hours/week) compared to the asymptomatic group ($p < 0.05$). However, no significant differences were observed in the other parameters, as presented in table 1.

Table 1 Participant characteristics: Mean or Frequency

	ASYMPTOMATIC (n=21)	SYMTOMATIC (n=21)	TOTAL (n=42)
Sex, M/F	M=14; F=7	M=14; F=7	M=28; F=14
Age, y (SD)	25.80 (3.67)	35.51 (12.52)	28.66 (9.56)
Height, cm (SD)	176.14 (10.44)	177.24 (8.83)	176.69 (9.57)
Weight, kg (SD)	70.43 (11.63)	74.56 (14.69)	72.49 (13.26)
BMI, kg/m^2 (SD)	22.63 (2.61)	23.57 (3.24)	23.01 (2.94)
Arm length, cm (SD)	56.55 (4.34)	56.64 (2.79)	56.60 (3.61)
Right-hand dominance (%)	20 (95%)	18 (86%)	38 (90%)
Competitive athletes	5	11	16
Av. Training volume, h/wk (SD)	5.81 (3.67)	9.21 (8.19)	7.51 (6.50)
Max. Training volume, h/wk (SD)	8.60* (4.57)	14.19*(9.92)	11.39 (8.14)
Av. matches per year (SD)	28.76 (25.27)	21.24 (19.90)	25.00 (22.79)
SPADI-D score (SD)	N/A	28.00 (14.59)	N/A

Abbreviations:*SD*=Standard Deviation, *BMI* = Body Mass Index, *SPADI-D* = Dutch Shoulder Pain and Disability Index, *M* = Male, *F* = Female, *N/A* = Not applicable, * indicating p<0.05, *Av.and Max. training volume* = Average and Maximal number of hours training per week, *Av. Matches per year* = Average number of matches per year.

In both groups, tennis was the most prevalent overhead racquet sport. In the asymptomatic group, 9 athletes were involved in both tennis and padel, while 8 athletes exclusively played tennis. Additionally, 2 athletes were solely engaged in padel, and 1 player in badminton. In the symptomatic group, 12 players exclusively participated in tennis, while 7 played both tennis and padel. Moreover, 1 athlete participated in both badminton and padel, and 1 athlete engaged in both tennis and badminton. Means and standard deviations of the ASH measurements in each position for both raters A and B are shown in table 2.

Table 2 Mean values of peak forces

Position	Asymptomatic Mean force in N			Symptomatic Mean force in N		
	Rater A1	Rater A2	Rater B	Rater A1	Rater A2	Rater B
I, Dom (SD)	92.60 (35.37)	92.00 (35.51)	99.95 (39.18)	112.43 (38.84)	105.81 (36.52)	109.24 (33.22)
Y, Dom (SD)	83.67 (30.77)	87.81 (37.26)	90.38 (29.30)	90.24 (25.63)	87.52 (23.49)	90.38 (24.95)
T, Dom (SD)	81.90 (29.13)	81.14 (28.29)	84.62 (27.43)	83.14 (26.50)	81.62 (23.17)	82.90 (27.15)
I,NDom (SD)	84.62 (33.61)	81.90 (31.72)	84.76 (27.42)	94.71 (33.90)	92.81 (32.56)	99.10 (29.85)
Y,NDom (SD)	71.57 (23.25)	72.24 (24.27)	75.62 (25.41)	83.57 (23.76)	80.67 (26.87)	82.81 (24.97)
T,NDom (SD)	69.19 (20.88)	69.43 (21.29)	73.00 (23.04)	73.05 (25.71)	72.57 (24.72)	76.81 (24.00)

The mean forces of the first and second trials performed by rater A in each position are presented in columns "Rater A1" and "Rater A2", respectively. The mean forces of the third trial in each position, conducted by rater B, are represented in column "Rater B".

Abbreviations: *SD*=Standard Deviation, *Dom* = Dominant arm, *NDom* = Non-Dominant Arm, *N* = Newton

The intra-rater reliability of rater A was excellent (0.93-0.98), for both the dominant and non-dominant arm across all positions among asymptomatic athletes (table 3). This excellent intra-rater reliability was also preserved in the symptomatic group (0.91-0.98) (table 4).

Table 3 Intra and inter-rater reliability of the ASH test in the asymptomatic group.									
Position	Intra-rater reliability, Rater A			Inter-rater reliability, Rater A1 and B			Inter-rater reliability, Rater A2 and B		
	ICC (95% CI)	SEM, N	MDC_{95}, N	ICC (95% CI)	SEM, N	MDC_{95}, N	ICC (95% CI)	SEM, N	MDC_{95}, N
I, Dom	0.93 (0.84; 0.97)	9.31	25.81	0.83 (0.64; 0.93)	15.25	42.27	0.87 (0.69; 0.94)	13.69	37.95
Y, Dom	0.95 (0.87; 0.98)	7.87	21.81	0.95 (0.70; 0.98)	6.98	19.35	0.93 (0.84; 0.97)	8.68	24.06
T, Dom	0.97 (0.94; 0.99)	4.69	13.00	0.93 (0.84; 0.97)	7.54	20.90	0.95 (0.88; 0.98)	6.31	17.49
I, NDom	0.93 (0.83; 0.97)	8.77	24.31	0.80 (0.57; 0.92)	13.61	37.72	0.82 (0.61; 0.92)	12.54	34.76
Y, NDom	0.98 (0.95; 0.99)	3.44	9.54	0.92 (0.80; 0.97)	6.89	19.10	0.91 (0.80; 0.96)	7.33	20.32
T, NDom	0.97 (0.94; 0.99)	3.40	9.42	0.90 (0.76; 0.96)	6.95	19.26	0.91 (0.79; 0.96)	6.65	18.43

Abbreviations:*ICC* = Intraclass correlation coefficient (single measure); *CI*= confidence interval; SEM= standard error of measurement; MDC_{95}= minimal detectable change with 95% CI.

Table 4 Intra and inter-rater reliability of the ASH test in the symptomatic group.									
Position	Intra-rater reliability, Rater A			Inter-rater reliability, Rater A1 and B			Inter-rater reliability, Rater A2 and B		
	ICC (95% CI)	SEM, N	MDC_{95}, N	ICC (95% CI)	SEM, N	MDC_{95}, N	ICC (95% CI)	SEM, N	MDC_{95}, N
I, Dom	0.96 (0.82; 0.99)	7.54	20.90	0.88 (0.72; 0.95)	12.67	35.12	0.86 (0.69; 0.94)	13.02	36.09
Y, Dom	0.91 (0.80; 0.96)	7.38	20.46	0.87 (0.71; 0.95)	9.12	25.28	0.92 (0.82; 0.97)	6.85	18.99
T, Dom	0.93 (0.83; 0.97)	6.77	18.77	0.84 (0.65; 0.93)	10.63	29.26	0.88 (0.73; 0.95)	8.74	24.23
I, NDom	0.94 (0.87; 0.98)	7.94	22.01	0.79 (0.55; 0.91)	14.71	40.77	0.81 (0.60; 0.92)	13.51	37.45
Y, NDom	0.95 (0.87; 0.98)	5.84	16.19	0.83 (0.62; 0.93)	10.14	28.11	0.84 (0.65; 0.93)	10.41	28.86
T, NDom	0.98 (0.96; 0,99)	3.48	9.65	0.91 (0.78; 0.96)	7.67	21.26	0.93 (0.81; 0.97)	6.54	18.13

Abbreviations *ICC:* Intraclass correlation coefficient (single measure); *CI*= confidence interval; *SEM*= standard error of measurement; MDC_{95}= minimal detectable change with 95% CI.

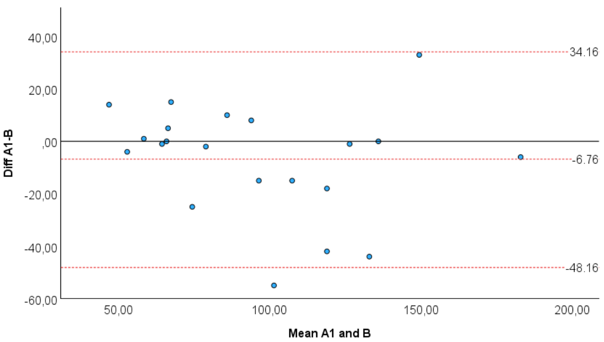
Regarding inter-rater reliability, the comparison between both measurements of rater A (A1 and A2) and rater B was good to excellent (0.80-0.95) across all positions for both arms in asymptomatic athletes. Similar findings were observed in symptomatic participants, with good to excellent inter-rater reliability (0.79-0.93). Generally, intra-rater reliability was consistently higher or equal to inter-rater reliability in all measurements for the asymptomatic group. In the symptomatic group, intra-rater reliability was higher in all measurements, except in the dominant Y-position.

Analysing inter-rater reliability between measurements of rater A (A1 and A2) and rater B, superior ICC values for A2 and B were observed across all positions in the asymptomatic group, except for the dominant and non-dominant Y-positions. Similarly, in the symptomatic group, A2 and B exhibited greater reliability in all positions, except for the dominant I-position.

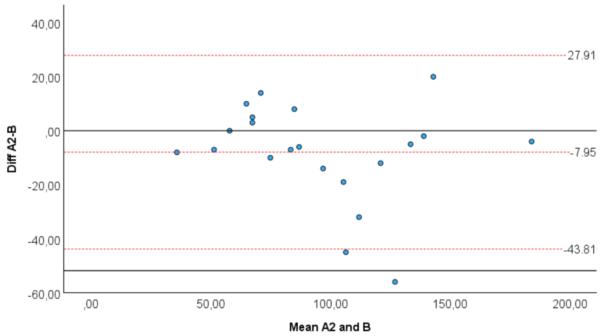
The SEM and MDC_{95} showed similar results in both asymptomatic and symptomatic groups: lower values for intra-rater reliability compared to inter-rater reliability, except for the dominant Y-position. Bland-Altman plots of both groups showed that the mean differences in maximal isometric strength between the two raters, when assessing the same position, were less than 7.95N. Among asymptomatic participants, neither systematic nor proportional bias was observed in either the I- and T-positions for both arms (Fig. 3a, 3b, 3e-l). However, in the dominant Y-position, systematic bias was evident (Fig. 3c), as confirmed by the paired t-test whereas rater B recorded significantly higher ASH values compared to measurement A1 ($p < 0.001$). Furthermore, proportional bias was observed in the dominant Y-position (Fig. 3d), where for higher force values, rater A2 recorded higher numbers than B ($p < 0.05$). Conversely, in symptomatic participants, systematic bias was only observed in the non-dominant T-position, where rater B evaluated significantly higher ASH values compared to measurement A2 (Fig. 3x), as confirmed by the paired t-test ($p<0.05$). Bland-Altman plots for all other positions in symptomatic participants (Fig. 3m-w) showed no systematic bias. Additionally, none of the positions exhibited proportional bias in the symptomatic group (Fig. 3a-x).

Figure 3: Bland Altman plots comparing measurements between rater A (A1 and A2) and B.

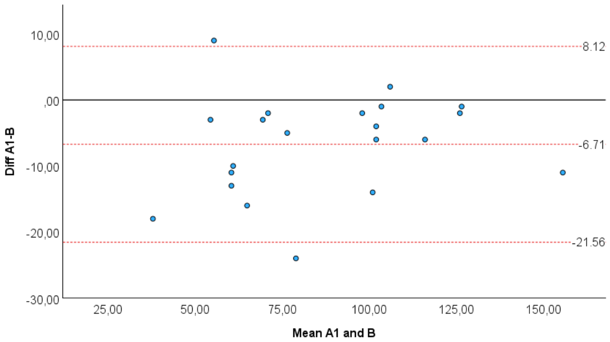
3a) I–position, dominant, asymptomatic, A1 and B



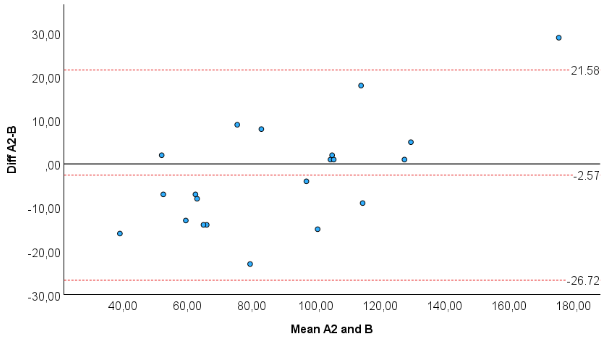
3b) I–position, dominant, asymptomatic, A2 and B



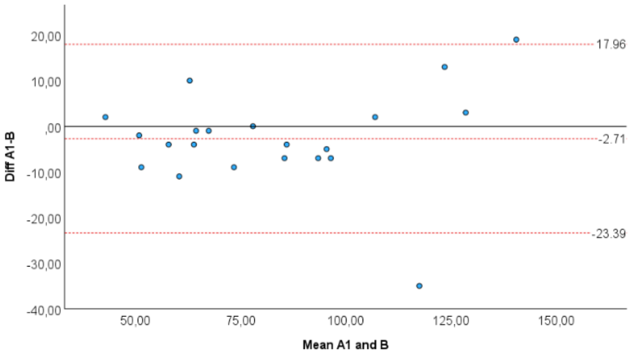
3c) Y-position, dominant, asymptomatic, A1 and B



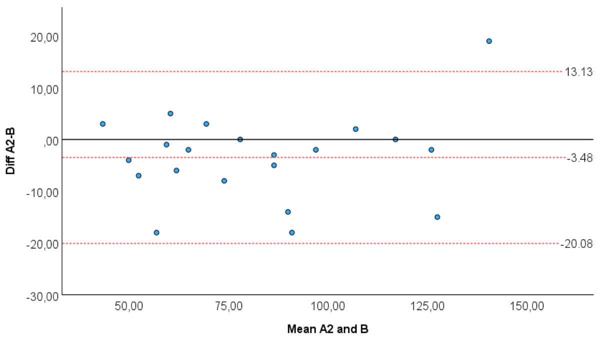
3d) Y-position, dominant, asymptomatic, A2 and B



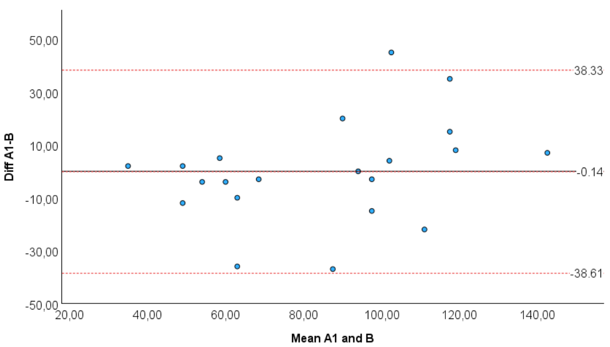
3e) T-position, dominant, asymptomatic, A1 and B



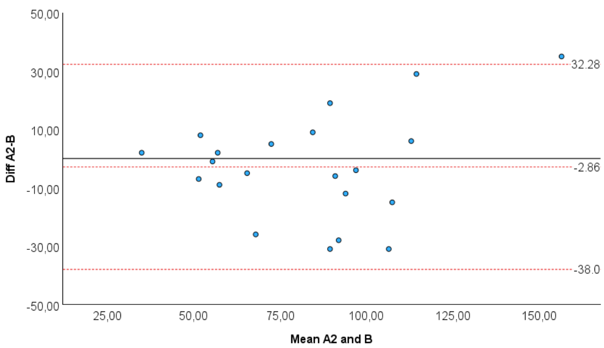
3f) T-position, dominant, asymptomatic, A2 and B



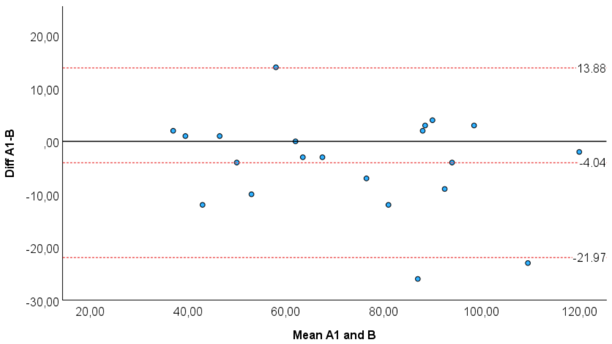
3g) I-position, non-dominant, asymptomatic, A1 and B



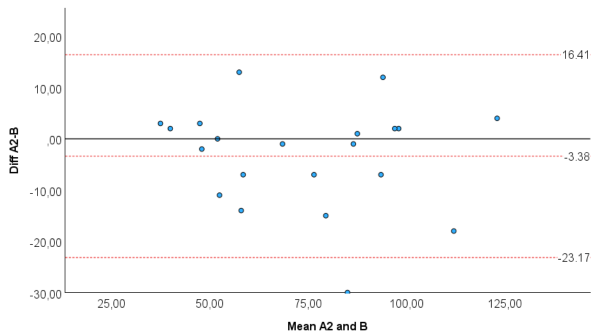
3h) I-position, non-dominant, asymptomatic, A2 and B



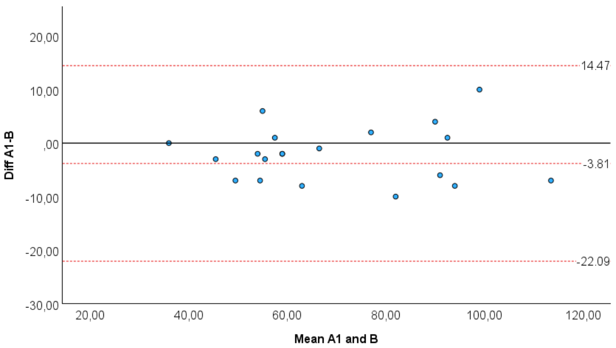
3i) Y-position, non-dominant, asymptomatic, A1 and B



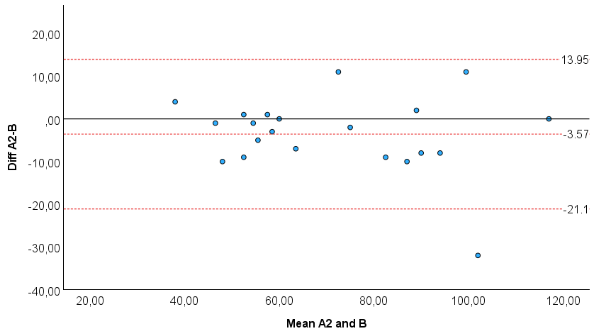
3j) Y-position, non-dominant, asymptomatic, A2 and B



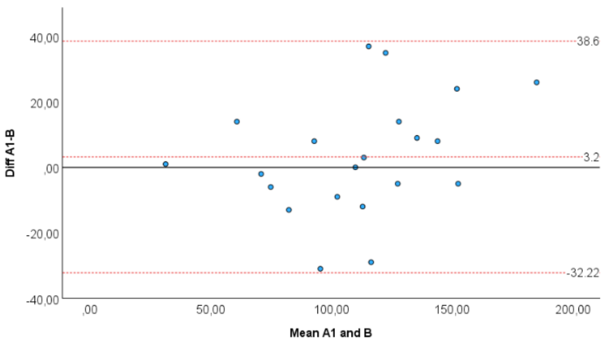
3k) T-position, non-dominant, asymptomatic, A1 and B



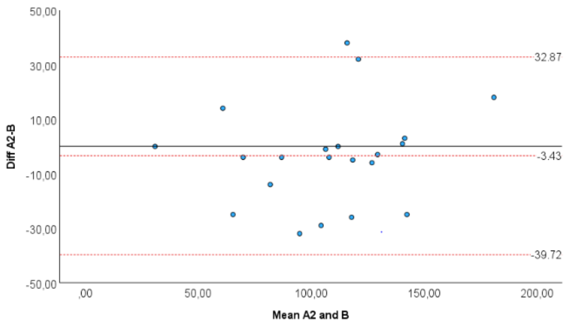
3l) T-position, non-dominant, asymptomatic, A2 and B



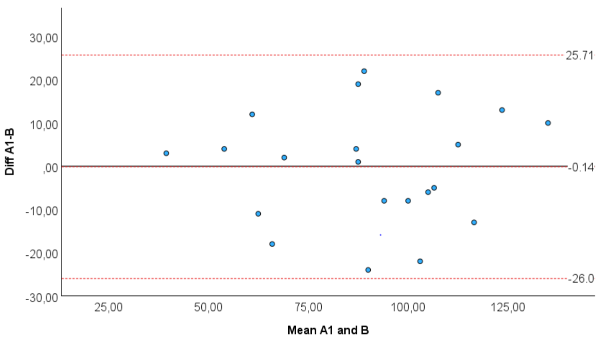
3m) I-position, dominant, symptomatic, A1 and B



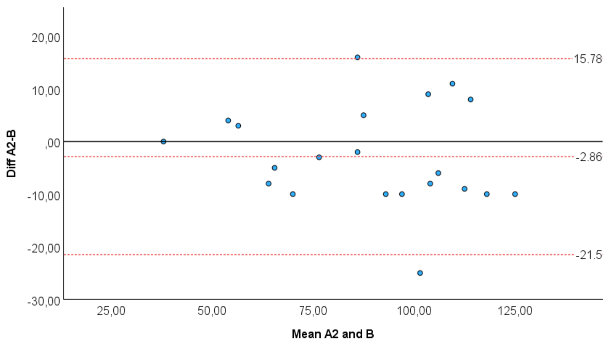
3n) I-position, dominant, symptomatic, A2 and B



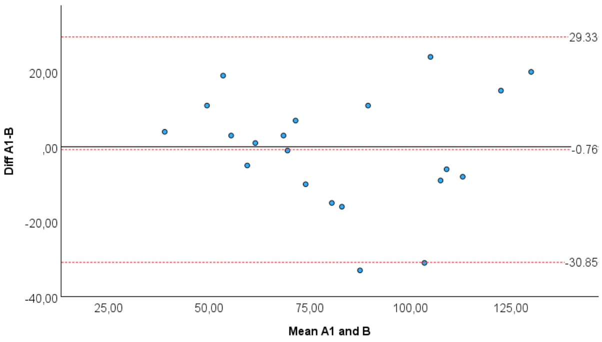
3o) Y-position, dominant, symptomatic, A1 and B



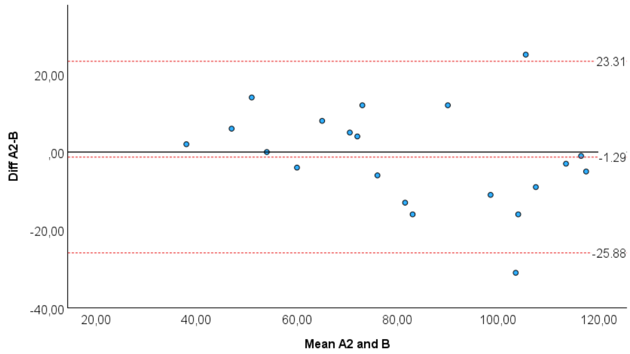
3p) Y-position, dominant, symptomatic, A2 and B



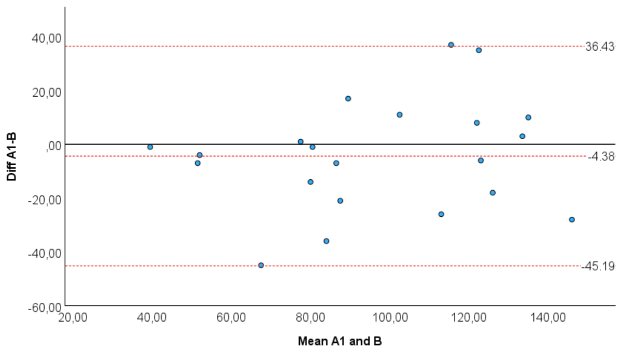
3q) T-position, dominant, symptomatic, A1 and B



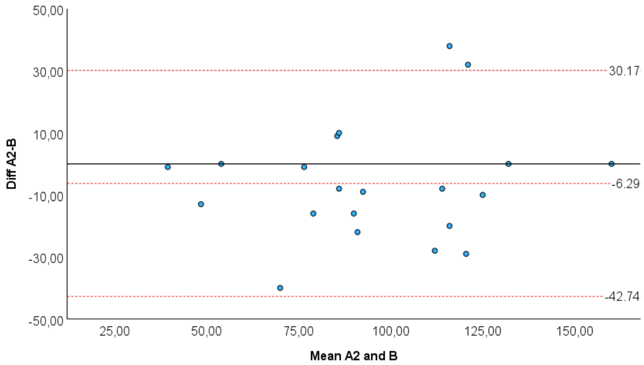
3r) T-position, dominant, symptomatic, A2 and B



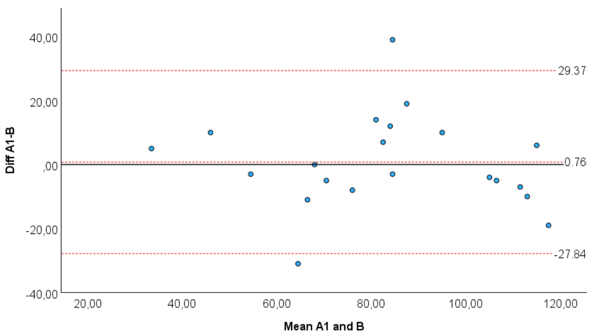
3s) I-position, non-dominant, symptomatic, A1 and B



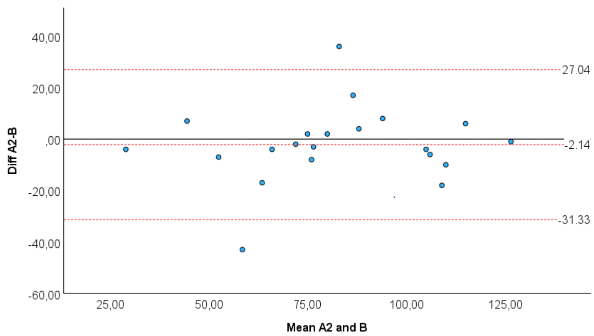
3t) I-position, non-dominant, symptomatic, A2 and B



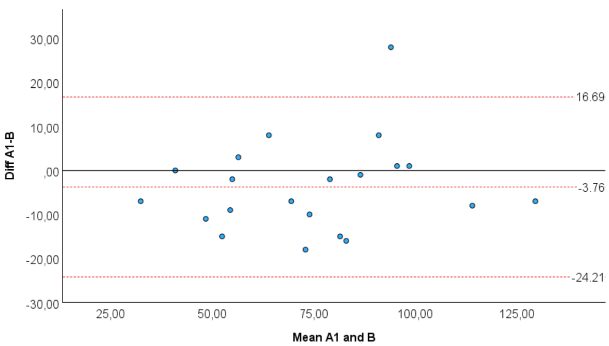
3u) Y-position, non-dominant, symptomatic, A1 and B



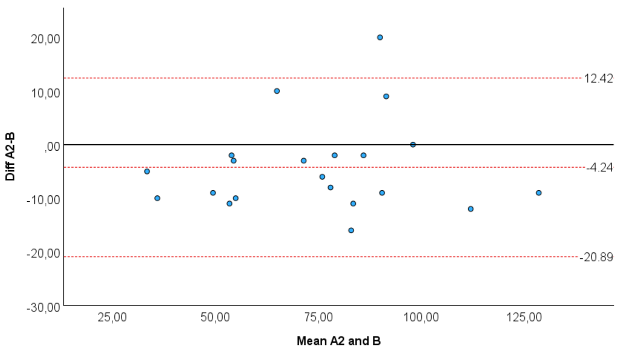
3v) Y-position, non-dominant, symptomatic, A2 and B



3w) T-position, non-dominant, symptomatic, A1 and



3x) T-position, non-dominant, symptomatic, A2 and B



Discussion

This study evaluated the reliability of the ASH test by novice examiners on overhead racquet athletes, divided into asymptomatic and symptomatic groups. Both groups exhibited excellent intra-rater reliability in all positions for dominant and non-dominant arms, while inter-rater reliability ranged from good to excellent. In the majority of test positions evaluated in both cohorts, comparison between A2 and B yielded higher ICC values than those between A1 and B.

Intra-rater reliability was consistently excellent in both groups. Similarly, Królikowska et al.'s[40] findings on asymptomatic recreational university athletes, demonstrated excellent intra-rater reliability across all ASH positions on a force plate. Furthermore, their study identified the lowest ICC values specifically for the dominant limb in the I-position, a result that aligns with the findings in the asymptomatic group of the current study. As this is the first study to investigate the ASH test in symptomatic athletes playing overhead racquet sport, comparative data from existing literature is lacking within this population. However, similar to the results observed in our symptomatic group, Beshay et al.'s[41] study on symptomatic shoulder patients also demonstrated excellent intra-rater reliability for isometric shoulder tests using HHD, although performed in positions other than the ASH test. Upon comparing intra-rater reliability between the two distinct groups in the current study, the asymptomatic group displayed higher reliability in three out of six positions (Y-dominant, T-dominant, and Y-Non-Dominant), while the reliability in the symptomatic group was higher in the remaining positions (I-Dominant, I-Non dominant, and T-Non-Dominant). The maximum variation in ICC values in the same testing position between both groups was only 0.04, indicating minor differences. This demonstrates that the ASH test was similar across asymptomatic and symptomatic overhead racquet athletes. Furthermore, intra-rater reliability showed minimal variation across different ASH positions within both asymptomatic and symptomatic groups, with ICC value differences not exceeding 0.05 and 0.07, respectively, indicating similar reliability across various positions.

Inter-rater reliability in the asymptomatic group exhibited good to excellent results across all positions. In contrast, Królikowska et al.[40] observed excellent inter-rater reliability in all positions of the ASH test among asymptomatic recreational university athletes, using K-force plates. Differences in reliability might arise from several factors. First, Królikowska's[40] population was not limited to overhead racquet athletes. Second, the duration of breaks between measurements varied; our study conducted consecutive ASH measurements, with transitions between rater A and rater B taking less than one minute, while Królikowska et al. reported breaks exceeding 90 minutes between rater sessions[40]. In this study, the dominant arm consistently showed higher mean peak force values and greater inter-rater reliability across all positions. This could be attributed to the asymmetric nature of the sports involved, where the dominant arm is more trained, less prone to fatigue, and capable of producing more consistent forces than the less-used non-dominant arm. Similarly, Królikowska et al.[40] found higher inter-rater reliability in the dominant arm for females across all positions, although this pattern did not hold for the male group in the I-position. The lowest inter-rater reliability in our asymptomatic group was in the non-dominant I-position, a trend also observed among female participants in Królikowska et al.'s[40] study.

Inter-rater reliability within the symptomatic group ranged from good to excellent, aligning with existing literature on HHD shoulder measurements in symptomatic patients[41]. These findings were consistent to those observed in the asymptomatic group, demonstrating similar inter-rater reliability across both groups. Despite being symptomatic, the dominant arm showed higher force values in all positions compared to the non-dominant arm. This strength discrepancy, even in the presence of pain, may be due to the asymmetric characteristics of the sports in this study. Moreover, inter-rater reliability was consistently higher for the dominant arm, except for the T-position where it was lower (ICC=0.84) than the non-dominant arm (ICC=0.91). This difference could potentially be attributed due to greater discomfort reported by in the dominant T-position during testing, affecting consistency.

SEM values indicate the precision of individual assessments, aiding in determining the clinical significance of measurements. MDC_{95} values, crucial for monitoring actual progress in muscle strength, indicate the minimum change needed to be 95% confident it exceeds the measurement error[36, 42]. In the asymptomatic group, both intra and inter-rater MDC_{95} and SEM values of the dominant arm were consistently higher than those of the non-dominant arm in the same test position. This trend was not observed in the symptomatic group. Furthermore, measures conducted by two different raters showed higher MDC_{95} and SEM values compared to those conducted by one rater, reflecting the lower ICCs in inter-rater reliability. For example, a significant change in the dominant I-position for an asymptomatic athlete should be at least 25.8N when assessed by one novice and 42.3N when assessed by two different examiners. Comparing our MDC values with those in other studies involving the ASH test was not possible due to discrepancies in the units used for the strength variable[2, 29, 43]. Furthermore, this study is the first to report and analyse systematic and proportional bias in the ASH test between two raters, supported by paired t-tests, Bland-Altman plots, and regression analysis, with no existing literature for direct comparison.

Limitations and strengths

This study acknowledges several limitations. Firstly, while raters were blinded to participants' health status, rater A often aware of the athlete’s sport type and level due to recruitment from her circle of acquaintances. Secondly, time constraints limited the feasibility of spreading measurements across multiple days per athlete. Initially, the plan was to evaluate intra-rater reliability twice with two ASH tests per rater. However, the pilot phase revealed that four ASH tests in one day was overly time-consuming and physically exhausting for participants, resulting in a reduction to three tests per athlete. Conducting all assessments on a single day, which involved 18 maximal strength tests per athlete, likely affected performance consistency. This compact scheduling may raise questions regarding athletes' continuous capability and psychological readiness to deliver maximal effort in each attempt. Peak force values showed that in 75% of the cases, the highest values occurred in the third test attempt, suggesting a warm-up, or learning effect rather than fatigue-induced performance decline. Additionally, the lack of randomisation among raters and test order, chosen for time efficiency and participant comfort, may have introduced order effects, influencing the interpretation of variations between positions[44]. Conducting a study with measurements over multiple days could offer advantages by allowing greater time flexibility, including the opportunity to assess intra-rater reliability twice, facilitate randomisation of test sequences and raters, and reduce learning effects. However, prolonging the period between assessments could increase the influence of extrinsic factors, such as changes in training routines or variations in shoulder complaints among symptomatic athletes. Thirdly, challenges arose regarding subjective aspects in both measurements and data interpretations. Athletes were instructed to exert “maximal pain-free force”, due to discomfort reported in both groups during the pilot phase. However, defining “pain-free” is subjective and challenging to standardise, particularly given this study’s novelty in including symptomatic individuals without established guidelines. Considering that asymptomatic athletes also reported discomfort during testing, it is likely that many athletes exceeded the pain threshold. Additionally, data interpretation relied on visually estimating maximal force values on graphs, which can vary among different raters and may have contributed to lower inter-rater reliability. Despite efforts for standardisation, using raw data could enhance objectivity in data interpretation. Future research should explore efficient and objective methods for data extraction and interpretation, considering constraints of time and resources in clinical practice.

This study has several strengths. Firstly, this study is the first of its kind to explore the reliability of the ASH test not only in a population of athletes engaging in overhead racquet sports but uniquely, it extends its focus to include both asymptomatic and symptomatic athletes in separate cohorts. This inclusion of a representative population for physical therapy practice, along with a representative sample of examiners, enriches the external validity of the study's findings[27]. Secondly, measurements were conducted blinded and independently. The assessors were blinded to the categorisation of participants as asymptomatic or symptomatic, as well as to each other's results. Participants were also blinded to their own test results. Moreover, using the same rater for both measurement and interpretation provided a comprehensive insight into the ASH test procedure. Thirdly, the study demonstrates comparability between the two groups across most demographic parameters. The only significant difference was the maximum training hours per week, being higher in the symptomatic group. This discrepancy was attributed to one single symptomatic participant counting all coaching hours. Future research could investigate the impact of training hours by comparing strength values and ASH test reliability between competitive and recreational athletes. This, however, was beyond the scope of the current study. Finally, the protocol was optimised for clinical use, emphasising time efficacy. While Ashworth et al.[2] reported a 6 minute-duration for one ASH test, this study completed it in 3 minutes. We recommend allocating 10 minutes for the entire process, encompassing instructions, warm-up, the ASH test, and data interpretation.

Practical relevance

This study found that novice examiners with minimal training can reliably conduct ASH measurements. These findings can be generalised to a wider group of inexperienced assessors with similar characteristics and training. In clinical practice, ICC values above 0.90 are considered suitable for confident decision-making, highlighting the precision required in measurements[45]. In this study, intra-rater reliability values ranged from 0.93 to 0.98 for asymptomatic athletes and from 0.91 to 0.98 for symptomatic athletes. For practical application, this study specifically analyses inter-rater reliability by comparing measurements A1 and B. Given that A1 represents the first ASH measures of rater A, this aligns with the initial and only assessment undertaken by rater B. In the asymptomatic group, inter-rater reliability varied from 0.80 to 0.95, with T- and Y-positions for both arms exceeding the critical value of 0.90. In the symptomatic group, these values ranged from 0.79 to 0.91, with only the non-dominant T-position reaching the critical threshold of 0.90. Although ICC values between A2 and B were slightly higher, this increment could be due to participants' growing familiarity with the protocol, suggesting a learning effect that may have inflated the reliability values.

Due to time constrains, this study deviated from Ashworth et al.'s[2] recommendation of holding familiarisation sessions prior to the study, which typically include a minimum of three complete ASH tests on separate days. While such familiarisation may not be feasible in certain clinical situations, such as following an injury[29], integrating the ASH test as a routine exercise during rehabilitation or training before the actual could minimise learning effects. This approach could enhance the reliability of the ASH test in clinical practice.

This research primarily focused on evaluating the ASH test's reliability utilising absolute force values recorded during a 1-second interval for analysis. This approach diverges from Ashworth et al.'s original methodology, which relied on maximal net peak force values[2]. Neither the original study by Ashworth et al.[2] nor this study normalised force values, reflecting the ongoing uncertainty regarding the best normalisation practices. However, normalization is vital for clinical interpretation to ensure equitable comparisons across subjects, as factors like limb length can affect force production. On the other hand, Tooth et al.[29]presented calculated joint moments by multiplying force values with arm length (N*m). Future research should aim to establish a standardised approach for normalisation and explore the influence of variables such as BMI on the ASH test results. These advancements could enhance interpretative accuracy, thereby improving the clinical utility of the ASH test.

The findings from this study lay a groundwork for future use of the ASH test in sports medicine. Firstly, the demonstrated reliability of the ASH test enables healthcare practitioners to confidently assess isometric shoulder strength and tailor programs for performance optimisation of overhead racquet athletes. Secondly, the reliability in symptomatic overhead racquet athletes supports the ASH test's use in rehabilitation and return-to-play protocols. Currently, no single test or test battery dictates return-to-play decisions after a shoulder injury[46]. By establishing baseline strength measurements using the ASH test, therapists can set objective criteria for athletes to meet before they are cleared to resume competitive play. This not only enhances the safety of the return-to-play process but also provides athletes with clear goals to work towards during their rehabilitation. For a comprehensive assessment of an overhead racquet athlete's readiness to return to play after a shoulder injury, we recommend combining the ASH test with other tests such as performance, ER/IR strength, and kinetic chain testing[46]. However, further investigation in subsequent studies is required to validate the potential application.

Future research should explore whether the ASH test results can predict shoulder injuries, which could improve its utility in assessing injury risk among overhead racquet athletes. Moreover, there's a legitimate question about the suitability of conducting the test in the prone position for these athletes. Assessing the reliability of the ASH test when performed in the standing position could address this concern, potentially making it more applicable to the specific demands of the sport.

Conclusion

Excellent intra-rater reliability of the ASH test, administered by a single novice rater, was demonstrated in both asymptomatic and symptomatic overhead racquet athletes, highlighting its consistency in assessing isometric upper limb strength. Despite this, variability was noted between two different raters, with inter-rater reliability ranging from good to excellent, not consistently meeting the standards required for clinical practice. Intra-rater measurements also showed lower SEM and MDC values compared to the inter-rater measurements. Therefore, it is recommended to consistently use the same rater across multiple ASH test administrations. Nonetheless, given the excellent inter-rater reliability observed in T- and Y-positions for both arms for asymptomatic athletes, and in the dominant Y-, and non-dominant T-positions for symptomatic athletes, these particular test positions can still be recommended for use across two raters in clinical practice. The ASH test holds valuable future applications in sports medicine. It can offer a precise assessment of isometric shoulder strength in overhead racquet athletes, aiding injury rehabilitation and potentially enhancing return-to-play guidelines. Future research is needed to explore its validity and predictive value for shoulder injuries and to assess the feasibility of conducting the test in a standing position, which may provide a more functional assessment of shoulder strength in overhead racquet athletes.

References

1. expertscape. expertise in shoulder pain: worldwide 2023 [Available from: https://expertscape.com/ex/shoulder+pain.
2. Ashworth B, Hogben P, Singh N, Tulloch L, Cohen DD. The Athletic Shoulder (ASH) test: reliability of a novel upper body isometric strength test in elite rugby players. BMJ Open Sport Exerc Med. 2018;4(1):e000365.
3. Marchena-Rodriguez A, Gijon-Nogueron G, Cabello-Manrique D, Ortega-Avila AB. Incidence of injuries among amateur badminton players: A cross-sectional study. Medicine (Baltimore). 2020;99(18):e19785.
4. Castillo-Lozano R, Casuso-Holgado MJ. Incidence of musculoskeletal sport injuries in a sample of male and female recreational paddle-tennis players. J Sports Med Phys Fitness. 2017;57(6):816-21.
5. Winge S, Jørgensen U, Lassen Nielsen A. Epidemiology of injuries in Danish championship tennis. Int J Sports Med. 1989;10(5):368-71.
6. Fahlström M, Yeap JS, Alfredson H, Söderman K. Shoulder pain -- a common problem in world-class badminton players. Scand J Med Sci Sports. 2006;16(3):168-73.
7. Guermont H, Mittelheisser G, Reboursière E, Gauthier A, Drigny J. Shoulder muscle imbalance as a risk factor for shoulder injury in elite badminton players: A prospective study. Phys Ther Sport. 2023;61:149-55.
8. Demeco A, de Sire A, Marotta N, Spanò R, Lippi L, Palumbo A, et al. Match Analysis, Physical Training, Risk of Injury and Rehabilitation in Padel: Overview of the Literature. Int J Environ Res Public Health. 2022;19(7).
9. Aldridge R, Stephen Guffey J, Whitehead MT, Head P. The effects of a daily stretching protocol on passive glenohumeral internal rotation in overhead throwing collegiate athletes. Int J Sports Phys Ther. 2012;7(4):365-71.
10. Ben Kibler W. Biomechanical Analysis of the Shoulder During Tennis Activities. Clinics in Sports Medicine. 1995;14(1):79-85.
11. Cools AM, Johansson FR, Borms D, Maenhout A. Prevention of shoulder injuries in overhead athletes: a science-based approach. Braz J Phys Ther. 2015;19(5):331-9.
12. Kovacs M, Ellenbecker T. An 8-stage model for evaluating the tennis serve: implications for performance enhancement and injury prevention. Sports Health. 2011;3(6):504-13.
13. Fleisig GS, Barrentine SW, Escamilla RF, Andrews JR. Biomechanics of Overhand Throwing with Implications for Injuries. Sports Medicine. 1996;21(6):421-37.
14. Wilk KE, Obma P, Simpson CD, Cain EL, Dugas JR, Andrews JR. Shoulder injuries in the overhead athlete. J Orthop Sports Phys Ther. 2009;39(2):38-54.
15. Wilk KE, Macrina LC, Fleisig GS, Aune KT, Porterfield RA, Harker P, et al. Deficits in Glenohumeral Passive Range of Motion Increase Risk of Shoulder Injury in Professional Baseball Pitchers: A Prospective Study. Am J Sports Med. 2015;43(10):2379-85.
16. McFarland EG, Tanaka MJ, Papp DF. Examination of the shoulder in the overhead and throwing athlete. Clin Sports Med. 2008;27(4):553-78.
17. Strama E, Keenan KA, Sell T, Faherty M, Rafferty D, Salesi K, et al. Throwing Shoulder Adaptations Are Not Related to Shoulder Injury or Pain: A Preliminary Report. Sports Health. 2023:19417381231197217.
18. Olivier B, Lala B, Gillion N. The cricketer's shoulder and injury: Asymmetries in range of movement and muscle length. S Afr J Physiother. 2020;76(1):754.
19. Almeida GP, Silveira PF, Rosseto NP, Barbosa G, Ejnisman B, Cohen M. Glenohumeral range of motion in handball players with and without throwing-related shoulder pain. J Shoulder Elbow Surg. 2013;22(5):602-7.
20. Breckenridge JD, McAuley JH. Shoulder Pain and Disability Index (SPADI). J Physiother. 2011;57(3):197.
21. Hoppe MW, Brochhagen J, Tischer T, Beitzel K, Seil R, Grim C. Risk factors and prevention strategies for shoulder injuries in overhead sports: an updated systematic review. Journal of Experimental Orthopaedics. 2022;9(1).
22. Bahr R. Why screening tests to predict injury do not work-and probably never will…: a critical review. Br J Sports Med. 2016;50(13):776-80.
23. Mentiplay BF, Perraton LG, Bower KJ, Adair B, Pua YH, Williams GP, et al. Assessment of Lower Limb Muscle Strength and Power Using Hand-Held and Fixed Dynamometry: A Reliability and Validity Study. PLoS One. 2015;10(10):e0140822.
24. Romero-Franco N, Fernández-Domínguez JC, Montaño-Munuera JA, Romero-Franco J, Jiménez-Reyes P. Validity and reliability of a low-cost dynamometer to assess maximal isometric strength of upper limb. J Sports Sci. 2019;37(15):1787-93.

25. Mavroidis C, Nikitczuk J, Weinberg B, Danaher G, Jensen K, Pelletier P, et al. Smart portable rehabilitation devices. J Neuroeng Rehabil. 2005;2:18.
26. Roy JS, MacDermid JC, Orton B, Tran T, Faber KJ, Drosdowech D, et al. The concurrent validity of a hand-held versus a stationary dynamometer in testing isometric shoulder strength. J Hand Ther. 2009;22(4):320-6; quiz 7.
27. Schrama PP, Stenneberg MS, Lucas C, van Trijffel E. Intraexaminer reliability of hand-held dynamometry in the upper extremity: a systematic review. Arch Phys Med Rehabil. 2014;95(12):2444-69.
28. Yen D. Limitations of isokinetic testing to determine shoulder strength after rotator cuff repair. Iowa Orthop J. 2005;25:141-4.
29. Tooth C, Forthomme B, Croisier JL, Gofflot A, Bornheim S, Schwartz C. The Modified-Athletic Shoulder Test: Reliability and validity of a new on-field assessment tool. Phys Ther Sport. 2022;58:8-15.
30. Trunt A, Fisher BT, MacFadden LN. Athletic Shoulder Test Differences Exist Bilaterally in Healthy Pitchers. Int J Sports Phys Ther. 2022;17(4):715-23.
31. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol. 2011;64(1):96-106.
32. McKinney J, Velghe J, Fee J, Isserow S, Drezner JA. Defining Athletes and Exercisers. Am J Cardiol. 2019;123(3):532-5.
33. Delgado DA, Lambert BS, Boutris N, McCulloch PC, Robbins AB, Moreno MR, et al. Validation of Digital Visual Analog Scale Pain Scoring With a Traditional Paper-based Visual Analog Scale in Adults. JAAOS Global Research & Reviews. 2018;2(3):e088.
34. Cools AM, De Wilde L, Van Tongel A, Ceyssens C, Ryckewaert R, Cambier DC. 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. 2014;23(10):1454-61.
35. Bujang MA. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review. Archives of Orofacial Sciences. 2017;12:1-11.
36. Portney LG. Foundations of Clinical Research: Applications to Evidence-Based Practice, 4e. New York, NY: F. A. Davis Company; 2020.
37. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155-63.
38. Giavarina D. Understanding Bland Altman analysis. Biochem Med (Zagreb). 2015;25(2):141-51.
39. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res. 2005;19(1):231-40.
40. Królikowska A, Mika A, Plaskota B, Daszkiewicz M, Kentel M, Kołcz A, et al. Reliability and Validity of the Athletic Shoulder (ASH) Test Performed Using Portable Isometric-Based Strength Training Device. Biology (Basel). 2022;11(4).
41. Beshay N, Lam P, Murrell G. Assessing the Reliability of Shoulder Strength Measurement: Hand-Held versus Fixed Dynamometry. Shoulder & Elbow. 2011;3:244-51.
42. Feijen S, Kuppens K, Tate A, Baert I, Struyf T, Struyf F. Intra- and interrater reliability of the 'lumbar-locked thoracic rotation test' in competitive swimmers ages 10 through 18 years. Phys Ther Sport. 2018;32:140-4.
43. Morrison G, Ashworth B, Taylor-Kaveney T. The validity of the sphygmomanometer for shoulder strength assessment in amateur rugby union players. Phys Ther Sport. 2021;47:59-65.
44. Lucas CP. The order effect: reflections on the validity of multiple test presentations. Psychol Med. 1992;22(1):197-202.
45. Polit DF, Beck CT. Nursing research : generating and assessing evidence for nursing practice. 8th ed ed. Philadelphia: Wolters Kluwer Health/lippincott Williams & Wilkins Philadelphia; 2008.
46. Schwank A, Blazey P, Asker M, Møller M, Hägglund M, Gard S, et al. 2022 Bern Consensus Statement on Shoulder Injury Prevention, Rehabilitation, and Return to Sport for Athletes at All Participation Levels. J Orthop Sports Phys Ther. 2022;52(1):11-28.
47. training C. Cochrane Handbook for Systematic Reviews of Interventions 2020 [Available from: https://training.cochrane.org/handbook/archive/v6.1.

Appendix

Figure 1: Demographic questionnaire

Vragenlijst versienummer
2 Datum 19/10/2023

Assessment of the Reliability of the Athletic Shoulder (ASH) Test and assessing differences in Athletic Shoulder Test and Shoulder Endurance Test (SET) between asymptomatic and symptomatic overhead racquet athletes.

1 Vragenlijst 1

Naam: ....
Voornaam: ....
Patiënten code (toegewezen door onderzoeker): ....

2 Vragenlijst 2

2.1 Persoonlijke informatie

Patiënten code:
Geslacht: M / V/ X
Email: .... @ ....
Geboortedatum: .... / .... / ....
Lengte:........................... cm
Gewicht: ....kg

2.2 Sport gerelateerde informatie

1. Bovenhandse sport(en): ....
2. Speelarm: RECHTS/LINKS
3. Gemiddelde aantal uren bovenhandse sport training per week: ....uren
4. Maximaal aantal uren training per week: ....uren/week
5. Competitie: JA/NEE
6. Indien ja op vraag 5, Ranking/Klassement? ....
7. Internationale toernooien? JA/NEE
8. Gemiddeld aantal wedstrijden per jaar: ....
9. Andere sport(en): ....
10. Aantal uren andere sport per week: ....uren/week

2.3 Bijkomende informatie

11. Behandelende arts: ………………………………………………….
12. Medicatiegebruik op dit moment? Zo ja, specifieer: JA/NEE ………………………………
13. Momenteel krachttraining schouderregio? JA/NEE
14. Momenteel in therapie voor schouderklachten en/of andere muscluloskeletale klachten? JA/NEE
15. Geschiedenis van traumatisch letsel, chirurgie aan de schoudergordel, bovenste borstkast, rug en bovenarm in het afgelopen jaar?

2.4 Informatie met betrekking tot schouderklachten

16. Heeft u in de afgelopen 30 dagen schouderklachten gehad tijdens het beoefenen van een racketsport? Zo ja, specifieer: JA/NEE……………………………………………………………
17. Indien het antwoord op deze vraag 'ja' is, gelieve vragenlijst 3 in te vullen.
18. Indien ja op vraag 16, hoelang zijn deze schouderklachten al aanwezig?……………………………

3 Vragenlijst 3

(Enkel indien 'ja' op vraag 16)

Patiënten code: …….
VAS tijdens sport (in te vullen op de voorziene vragenlijst)

SPADI (in te vullen op de voorziene vragenlijst)

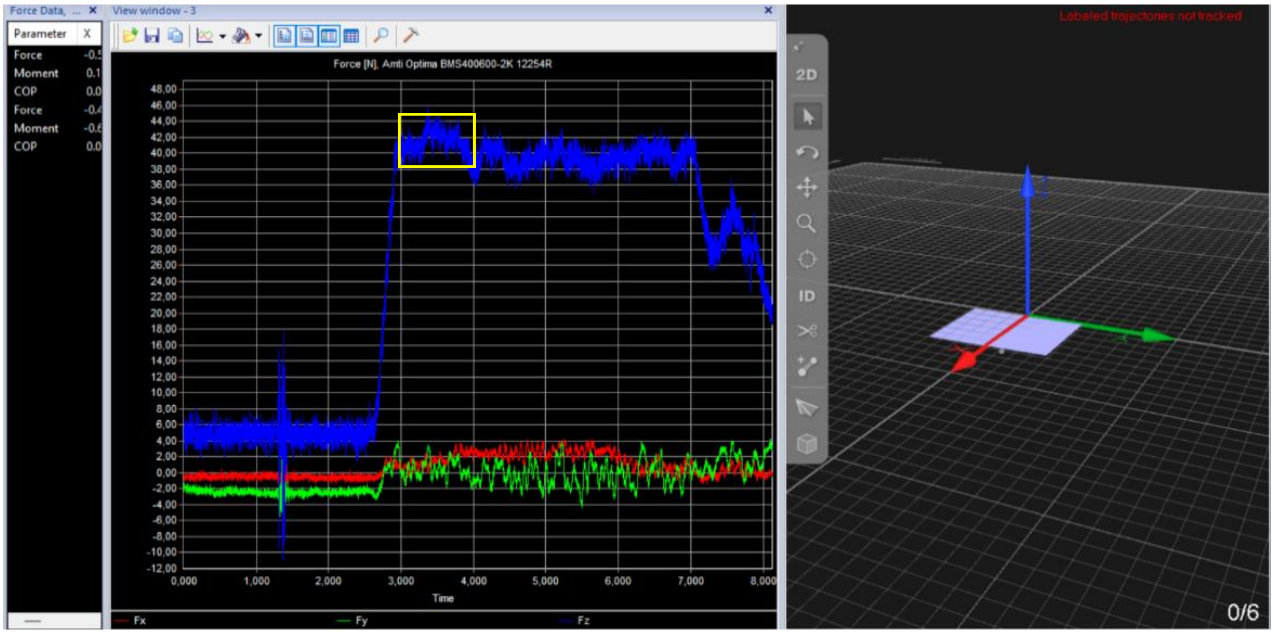


Figure 4: Example of the output window of one test position of the ASH test, generated by the Qualisys Track Manager software. In this window, forces were visually assessed using the blue graph, representing the vertical force axis (z). The highlighted area, indicated by the yellow arrow, represents a 1-second interval observed during the interpretation of force values. Within this interval, the average peak force was visually estimated for analysis.