Test-retest, intra- and inter-rater reliability of the reactive balance test in patients with chronic ankle instability

Alexandre Maricot[1], Elke Lathouwers[1], Jo Verschueren[1], Kevin De Pauw[1,2], Romain Meeusen[1,2,3], Bart Roelands[1,2] and Bruno Tassignon[1]*

[1]Human Physiology and Sports Physiotherapy Research Group, Faculty of Physical Education and Physiotherapy, Vrije Universiteit Brussel, Brussels, Belgium, [2]Brussels Human Robotics Research Center (BruBotics), Vrije Universiteit Brussel, Brussels, Belgium, [3]Department of Sports, Recreation, Exercise and Sciences (SRES), Faculty of Community and Health Sciences, University of the Western Cape, Cape Town, South Africa

Introduction: The Reactive Balance Test (RBT) could be a valuable addition to research on chronic ankle instability (CAI) and clinical practice, but before it can be used in clinical practice it needs to be reliable. It has already been proven reliable in healthy recreational athletes, but not yet in patients with CAI who have shown persistent deficits in dynamic balance. The study aimed to determine the test-retest, intra-, and inter-rater reliability of the RBT in patients with CAI, and the test-retest and inter-rater reliability of the newly developed RBT score sheet.

Methods: We used a repeated-measures, single-group design to administer the RBT to CAI patients on three occasions, scored by multiple raters. We included 27 participants with CAI. The study used multiple reliability measures, including Pearson r, intra-class correlations (ICC), standard error of measurement (SEM), standard error of prediction (SEP), minimal detectable change (MDC), and Bland–Altman plots, to evaluate the reliability of the RBT's outcome measures (visuomotor response time and accuracy). It also assessed the test-retest and inter-rater reliability of the RBT score sheet using the same measures.

Results: The ICC measures for test-retest reliability were similar for accuracy (0.609) and VMRT (0.594). Intra-rater reliability had high correlations and ICCs for accuracy ($r = 0.816$, ICC $= 0.815$) and VMRT ($r = 0.802$, ICC $= 0.800$). Inter-rater reliability had a higher ICC for VMRT (0.868) than for accuracy (0.690).

Conclusion: Test-retest reliability was moderate, intra-rater reliability was good, and inter-rater reliability showed moderate reliability for accuracy and good reliability for VMRT. Additionally, the RBT shows robust SEM and mean difference measures. The score sheet method also demonstrated moderate test-retest reliability, while inter-rater reliability was good to excellent. This suggests that the RBT can be a valuable tool in assessing and monitoring balance in patients with CAI.

1 Introduction

Chronic ankle instability (CAI) is a multifaceted clinical condition affecting approximately 40–45% of the adult population who have experienced a primary ankle sprain (1–3). The main symptoms of CAI are the prolonged symptoms of self-reported disability, local neuromuscular deficits and recurrent episodes of the ankle "giving way" (4–11). CAI negatively impacts health-related quality of life (12–16) and may alter physical activity levels (17–21). Moreover, an association between CAI and an early onset of osteoarthritis has been established (22–27). Due to the prevalence and long-lasting impairments of CAI, it is essential to develop testing tools to identify those at risk of (re) injury, to monitor rehabilitation progress and to make better-informed return to sport (RTS) decisions (28–34).

For these purposes, clinicians often use functional performance tests (FPTs) (28–31, 35). Two of the most reliable FPTs with excellent criterion validity for assessing dynamic balance of the lower extremities are the star excursion balance test (SEBT) and the Y-balance test (YBT) (35–40). Poor performance on these tests is associated with an increase in lower extremity injury risk (35, 37, 40–43). The SEBT measures maximum reach distance in eight directions whilst maintaining single-leg balance (44). A shorter and more reliable version, the YBT (45) uses the sum of three reach directions (anterior, posteromedial and posterolateral) to assess injury risk and identify patients with CAI (35, 41–43). Typically, individuals with CAI demonstrate a significantly lower reach distance on both FPTs (46–57).

However, The SEBT and YBT are limited in their ability to be applied to on-field sports contexts because they involve pre-planned tasks without considering dynamic environmental aspects of sports practice (35, 58). Functional exercises with neurocognitive components like decision-making and reaction times, in tandem with dynamic balance requiring an external focus of attention, are essential factors in sports injury risk and performance in sports such as basketball, football and table tennis (58–65).

Numerous studies have discovered that individuals with CAI have worse dynamic balance when performing tasks simultaneously than healthy individuals, suggesting that they require more attentional resources to stay balanced (66–75). Neurocognitive deficits, particularly reduced spatial awareness, have been observed in people with CAI assessed through computerized neurocognitive tests (76–78). This decline in spatial awareness may negatively impact their capacity to respond to environmental obstacles. Patients with CAI may consequently display less accuracy or slower reaction times when performing neurocognitive tasks that call for quick thinking and spatial awareness explaining these deficits.

The reactive balance test (RBT) was developed (79) to add neurocognitive components such as decision-making, visuomotor responses and environmental perception to the YBT in a standardized way. The RBT's primary outcome measures involve visuomotor response time (VMRT) and accuracy (79). Moreover, the RBT could be a valuable addition to research on CAI and clinical practice, but before it can be used in clinical practice it needs to be reliable. It has already been proven to reliably measure accuracy and VMRT in healthy recreational athletes (80), but not yet in patients with CAI. Therefore, the primary aim of this study was to determine the test-retest intra- and inter-rater reliability of the RBT in these patients through video-based assessment. The secondary objective of this study was to determine the test-retest, and inter-rater reliability of the RBT using a score sheet to assess the patients' accuracy in real-time. The score sheet was designed to enable rapid and efficient data collection, enabling evaluators to swiftly assess RBT performance without requiring time-consuming video playback. This approach could prove particularly beneficial when clinicians are working with extensive sample sizes or time-sensitive evaluations.

2 Materials and methods

We applied the guidelines for reporting reliability and agreement studies (GRRAS) (81) to report the RBT's reliability in individuals with CAI. We wanted to evaluate the reliability of the video-based analyzes and newly designed score sheet we developed. The video-based analysis entails the reviewer going through each trial afterward and scoring the RBT performance manually. In an effort to enhance efficiency and cut down on expenses, we introduced a score sheet that allows real-time scoring of the RBT performance. The consistency of the RBT outcome measures over time, under repeated and similar conditions, was defined as the test-retest reliability. The inter-rater reliability refers to how consistently different raters determine the accuracy and VMRT of patients of the same experimental trial. The intra-rater reliability of the RBT shows how consistently a rater can determine both RBT outcomes (82). The protocol was approved by the ethics committee of the UZ Brussel/Vrije Universiteit Brussel (B.U.N. 1,432,021,000,658) and was carried out following the "Code of Ethics of the World Medical Association" (Declaration of Helsinki).

2.1 Participants

Sample size calculations were based on the mathematics of Walter et al. (83). A reliability study with two experimental trials ($n = 2$), a null hypothesis of 0.5 and an alternative hypothesis of 0.8 based on an alpha of 0.05 and a beta error probability of 0.20 would require at least 22 participants. This reliability study included 27 individuals with CAI. Participants were recruited using convenience sampling through the Vrije Universiteit Brussel internal network and the social network of the researchers. Before participating in this study, all participants had to sign the informed consent form and complete three questionnaires Cumberland Ankle Instability Tool (CAIT), Foot and Ankle Ability Measure (FAAM) and the International Physical Activity Questionnaire (IPAQ). Depending on their MET/week the IPAQ assigns the participant to one of three classes: low physical activity level, moderate physical activity level and high physical activity level (84). Participant characteristics can be found in Table 1.

All participants were informed of the nature and procedures of the study. In- and exclusion criteria were based on the International Ankle Consortium guidelines (85). Forty-three eligible patients were screened on the following inclusion criteria: they (1) were between 18 and 35 years of age, (2) participate in at least one sport or physical workout at least once a week, (3) have a history of at least one significant ankle sprain that occurred at least 12 months before study enrolment that was associated with inflammatory symptoms (e.g., pain and swelling), (4) their initial sprain created at least one

TABLE 1 Patient characteristics.

Type	Mean ± SD	95% CI	Min	Max
Age (years)	22 (± 2.3)	21–23	18	28
Height (cm)	175.8 (± 10.3)	171.8–179.9	160	194
Weight (kg)	71.4 (± 10.1)	64.7–75.4	50.0	94.8
CAIT score	16.6 (± 3.6)	15.1–18.0	9	22
FAAM-ADL (%)	87.9 (± 11.3)	83.4–92.4	58	100
FAAM-S (%)	79.4 (± 15.4)	73.3–85.6	38	100
IPAq-score (MET/week)	4,560 (± 2,322)	3,641–5,478	1,080	9,172

ADL, activities of daily living; CAIT, Cumberland Ankle Instability Tool; FAAM, Foot and Ankle Ability Measure; IPAQ, International Physical Activity Questionnaire; MAX, maximum; MIN, minimum; S, Sport; SD, standard deviation; 95% CI, 95% Confidence Interval.

interrupted day of desired physical activity and (5) their most recent injury must have occurred more than 3 months before study enrolment, (6) experience recurrent sprains and 'feelings of giving way', (7) have at least two episodes of giving way in the 6 months before study enrolment, (8) and have a CAIT score of <24. All patients did not receive treatment during the study.

Potential participants with (1) a history of a fracture, (2) previous surgery to musculoskeletal structures (i.e., bones, joint structures, nerves) in the previous 2 years in either lower limb, (3) or any other relevant medical history, treatment or current condition (such as neurological diseases, inner ear disorders, color blindness etc.), which could affect balance ability or the action-perception pathways were excluded. Further exclusion criteria were evaluated in the pre-test checklist. All subjects were asked to limit alcohol and caffeine consumption the before and the day of each trial, not to take performance-enhancing medications, not to perform vigorous physical activities 24 h before the trial and to sleep at least 7 h the night before the trials. All participants confirmed to comply with the given instructions by completing the pre-test checklist.

2.2 Procedures

2.2.1 Test protocol

Patients visited the laboratory three times, once for a familiarization trial (± 1 h30) and twice for identical experimental trials (± 30 min). During the familiarization trial, the YBT was performed six times and the RBT two times on both legs. In each experimental trial, the YBT and the RBT were performed on both legs once. Figure 1 depicts an overview of a subject's participation protocol and timeline. The familiarization was conducted at least 1 week prior to the first experimental trial to ensure that participants had adequate time to get acquainted with the test procedures and to attenuate possible learning effects during the experimental trials. The duration and range of the test-retest time frame were chosen to reflect a clinically relevant period and lasted 22 (± 10) days on average (86). During the measurements, it was ensured that the conditions were always the same.

2.2.2 Y-balance test (YBT)

The Y-Balance Test (YBT Kit™, FunctionalMovement.com, Danville, VA) was performed on the Y-balance test kit. The YBT was performed according to the protocol and instructions described in earlier reliability research (40, 45, 80) provided guidelines for scoring

and evaluating the test results. The test determines the subject's maximum reach distance and ability to maintain balance while reaching in different directions. The maximal reach distance was measured by reading the reach distance at the proximal edge of the reach indicator. A reach was considered successful if the participant followed the instructions of Plisky et al. (45) and Gribble et al. (40). The trial was discarded and repeated if the subject; failed to maintain a one-sided stance on the platform, failed to maintain foot contact with the reach indicator on the target area while it was in motion (e.g., kicked the reach indicator), used the reach indicator for support of the stance, did not return to the starting position under control, or failed to keep hands on the iliac crest. The YBT was performed three times in each direction.
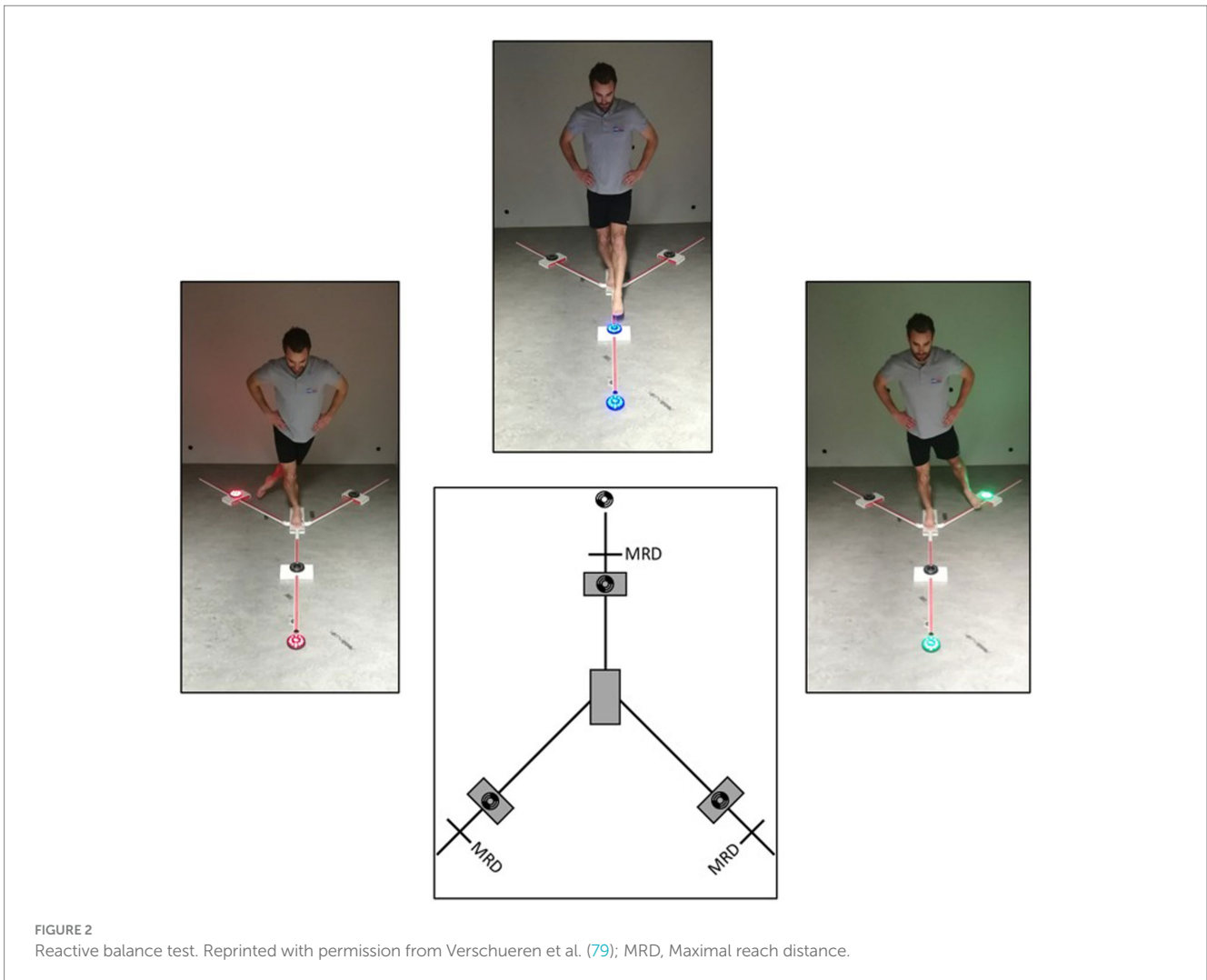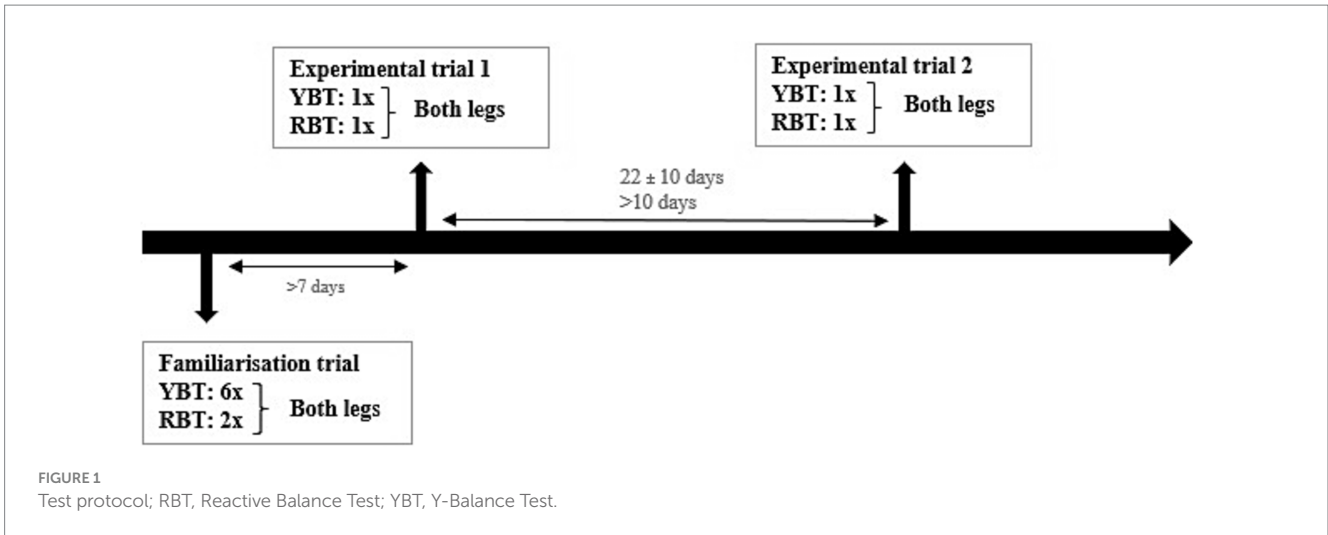
2.2.3 Reactive balance test (RBT)

The RBT is neurocognitive FPT that assesses VMRT and dynamic balance using visual stimuli. Its purpose is to extinguish a color-matched LED light on the YBT in response to an initial leading visual stimulus (Figure 2). The protocol, procedures, and scoring system are explained in detail in the study of Verschueren et al. (79).

The test uses Fitlight™ hardware, software, and four LED lights placed on the YBT Kit™. LED lights are positioned on each corresponding axis of the Y-balance Test Kit™ using 80% of the maximum reach distance plus six centimeters, as six centimeter corresponds with the radius of the LED lights housing. Each color is assigned to a specific axis and appears 12 times at random for 36 stimuli, with inter-stimulus times varying between 1.5, 2, and 2.5 s. Blue represented the front axis, green the posteromedial, and red the posterolateral. The Fitlight™ software were programmed to randomize colors and inter-stimulus times, making it difficult for participants to anticipate the timing of the following visual stimulus and the direction of the targeted motor response.

Accuracy was calculated using the following formula: Accuracy (%) = ((Total number of stimuli−(missed stimuli + multiple attempts needed + decision errors + balance errors))/Total number of stimuli) x 100. The definitions for each error are provided in Table 2. Each experimental trial was filmed with a video camera (Handycam 1,080 50i, HDR-CX105E, Sony Corporation, Japan) to manually analyze accuracy based on videorecording. When reviewing the RBT videos, raters recorded instances where participants missed stimuli, required multiple attempts, made decision errors, and exhibited balance errors.

To score the accuracy of the participant in real-time, we have created a score sheet with columns representing stimulus number,

FIGURE 1
Test protocol; RBT, Reactive Balance Test; YBT, Y-Balance Test.

FIGURE 2
Reactive balance test. Reprinted with permission from Verschueren et al. (79); MRD, Maximal reach distance.

target light, and four error types. The score sheets were adapted to the sequence of the stimuli used in each trial. Missed lights were given priority over stimuli that required multiple attempts or contained balance and decision errors. The VMRT was corrected for accuracy by removing the corresponding VMRT values from the original Fitlight™ dataset. The mean VMRT was obtained by only counting the correct extinguished LED lights. The score sheet can be found in Figure 3.

Error	Definition
Missed stimulus	The participant failed to extinguish the LED light
Multiple attempts	The participant reached from the standardized position but failed to extinguish the LED light the first time. The participant needed two or more attempts
Decision error	The participant initiated a movement in the wrong direction
Balance error	- The participant did not start from the standardized position at the stimulus onset - The participant is trying to find balance during the reach - The participant needs to put a hand or foot on the floor - The participant steps off the YBT Test kit - The participant is not able to keep the hands on the hips - The participant lifts the forefoot or heel off the testing surface

YBT, Y-Balance test.

2.3 Raters

One rater with a Master of Science degree in physiotherapy and rehabilitation sciences, and the other raters were master's students (E.D. and A.W.) in physiotherapy and rehabilitation sciences, with experience in YBT and RBT procedures evaluated RBT accuracy and VMRT. Raters were blinded to each other's evaluations. Test-retest reliability was determined by comparing rater 1 ratings from both experimental trials, intra-rater reliability was computed by comparing ratings of rater 1 from the second experimental trial. Inter-rater reliability was determined by comparing rater 1 and rater 2 ratings from the second experimental trial, and score sheet method reliability was compared using rater 2 ratings from both experimental trials and comparing them to rater 3 ratings.

2.4 Statistical analysis

Statistical analyzes were performed using the RStudio software version 4.2.2 (2022-10-31 ucrt). The mean, standard deviation and 95% confidence interval were calculated for accuracy and VMRT to illustrate the participants' characteristics. Reliability calculations were applied to the data of the self-reported CAI side confirmed by the lowest CAIT score. The Shapiro Wilk test was used to assess the normality of the data distribution, and this was supplemented by visual inspection using histograms and QQplots (87). To describe all reliability measures, we used the following indexes: Pearson r, intraclass coefficients (ICC), standard error of measurement (SEM), standard error of prediction (SEP) and minimal detectable change (MDC) (88–98). Additionally, we reported Bland–Altman plots to graphically render the agreement between two raters or ratings and the limits of the agreement (LOA) (99, 100). A paired t-test was used to detect any systematic biases between raters, ratings, or trials.

The guidelines of Koo and Li (88) were employed to select the appropriate model, type, and definition of ICC. ICC values less than 0.5 indicate poor reliability, values between 0.50 and 0.75

indicate moderate reliability, values between 0.75 and 0.90 indicate good reliability and values greater than 0.90 indicate excellent reliability (101). We interpreted Pearson r correlations between 0.1 and 0.3 as weak, between 0.3 and 0.5 as moderate and above 0.5 as strong. The statistical significance level was set at $p < 0.05$ for all statistical analyzes.

A two-way random absolute agreement, single-rater model (ICC 2.1), was used to estimate the inter-rater reliability. The test-retest and intra-rater reliability were calculated using a two-way mixed effect, consistency, single-rater model (ICC 3.1) as ICC equivalent.

3 Results

An overview of the relationship between variables (Pearson r and ICC), variability measures (SEM and SEP) and the ability of the RBT to detect true changes in score (MDC) for all reliability outcome measures are shown in Table 3.

3.1 Video-based RBT assessment

Patients with CAI scored an average accuracy of 81.20% (± 9.73%; CI: 77.27–85.13%) and 80.45% (± 10.60%; CI: 76.26–84.65%) for the first and second experimental trials, respectively. Corresponding VMRT were 779 ms (± 133 ms; CI: 725–832 ms) and 770 ms (± 93 ms; CI: 733–806 ms). The accuracy data were normally distributed. VMRT was normally distributed for the first, but not for the second experimental trial. The first and second trials showed a correlation of 0.612 for accuracy and 0.633 for VMRT. Similar ICC measures were found at 0.609 (0.297–0.804) for accuracy and 0.594 (0.276–0.795) for VMRT between both trials. We computed the SEM and SEP based on the 95% confidence interval (CI) and standard deviation (SD) for both outcomes: 6.6 and 8.4% for accuracy and 59 ms and 74 ms for VMRT. The MDC was 18.4% for accuracy and 164 ms for VMRT.

Intra-rater reliability was determined with a two-week interval by rater 1. Mean accuracy in the second evaluation was 79.42% (± 10.17%, CI: 75.40–83.45%) and mean VMRT was 761 ms (± 100 ms; CI: 721–800 ms). Both ratings were correlated with a Pearson r of 0.816 and ICC of 0.815 (0.635–0.911) for accuracy. Similar measures were found between the two VMRT ratings. The Pearson r was 0.802, and the ICC was 0.800 (0.608–0.904). The following variability measures were computed from both ratings for both outcomes. SEM was 4.37% and SEP 5.89% for accuracy, while it was 45 ms and 60 ms, respectively, for VMRT. The MDC of both outcomes were 12.11% and 123 ms.

Rater 2 estimated a mean accuracy of 83.33% (± 8.19%; CI: 80.09–86.57%) and a mean VMRT of 766 ms (± 109 ms; CI: 723–809 ms). When comparing both raters, we calculated the following correlation: 0.738 and 0.875 for accuracy and VMRT, respectively. While the degree of association for the accuracy outcome was reflected in an ICC of 0.690 (0.419–0.846), the VMRT ICC was 0.868 (0.731–0.937). The variability measures between both raters for accuracy were 4.56 and 5.93% for the SEM and SEP, while the same outcome measures were 40 ms and 54 ms, respectively, for the VMRT calculations. The MDC was 12.65% and 110 ms between both raters.

FIGURE 3
Reactive balance test score sheet. The first column indicates the stimulus number. The target color is indicated by the color on the score sheet and stated in column 2. The following columns indicate the possible results on a stimulus: successful, missed, multiple attempts, decision error and balance error (in order).

TABLE 3 Reliability outcome measures overview.

	Pearson r	ICC	[95% CI]	SEM	SEP	MDC	MEAN DIFF	LOA
Video-based analysis								
Test-retest reliability								
ACC	0.612	0.616	[0.305;0.808]	6.6%	8.4%	18.2%	−1.4%	[−23;20]
VMRT	0.633	0.600	[0.284;0.799]	59 ms	74 ms	162 ms	19 ms	[−149;187]
Intra-rater reliability								
ACC	0.816	0.815	[0.635–0.911]	4.4%	5.9%	12.1%	−2.7%	[−15;11]*
VMRT	0.802	0.800	[0.608–0.904]	45 ms	60 ms	123 ms	4 ms	[−167;175]
Inter-rater reliability								
ACC	0.738	0.690	[0.419–0.846]	4.6%	5.9%	12.6%	1.2%	[−12;14]
VMRT	0.875	0.868	[0.731–0.937]	40 ms	54 ms	110 ms	9 ms	[−157;176]
Score sheet method								
Test-retest reliability								
ACC	0.696	0.695	[0.413–0.856]	4.7%	6.1%	13.0%	0.1%	[−18;18]
VMRT	0.598	0.585	[0.245–0.796]	69 ms	87 ms	191 ms	38 ms	[−163;238]
Inter-rater reliability								
ACC	0.831	0.822	[0.651–0.915]	3.1%	4.2%	8.5%	1.4%	[−14;17]
VMRT	0.959	0.958	[0.911–0.981]	23 ms	32 ms	64 ms	7 ms	[−161;175]

*Statistically significant; RBT, Reactive balance test; VMRT, Visuomotor response time (in ms); ACC, accuracy (in %); ICC [95%CI], 95% Confidence Interval of the Intraclass Correlation Coefficient; LOA, Limit Of Agreement; MEAN DIFF, Mean Difference between raters/ratings; SEM, Standard Error of Measurement; SEP, Standard Error of Prediction; MDC, Minimal Detectable Change.

3.2 Score-sheet-based RBT assessment

To determine test-retest reliability, rater 2 rated the first and second experimental trials using the score-sheet assessment method. Patients with CAI scored an average accuracy of 83.4% (± 8.99%; CI: 79.7–87.3%) and 84.8% (± 8.5%; CI: 81.4–88.1%) for the first and second experimental trial, respectively. The accuracy outcome measure was normally distributed. To describe the test-retest reliability of the accuracy measure we have calculated the pearson r (= 0.696) and ICC (= 0.695 [0.413–0.856]). The SEM and SEP of the second experimental trial were 4.71 and 6.13%, respectively. The MDC was 13.04%. VMRT was also normally distributed. The average VMRT across participants was 779 ms (± 135 ms; CI: 722-836 ms) and 785 ms (± 107 ms; 743–828) for both experimental trials.

To determine the inter-rater reliability of the RBT using the score-sheet method, rater 3 outcomes were compared with rater 2 outcomes. Patients with CAI scored an average accuracy score of 85.6% (± 7.3%; CI: 82.7–88.5%) and VMRT as 781 ms (± 112; CI: 737–826). Accuracy data was normally distributed whereas VMRT data was not for both raters. Rater 3 ratings correlated with rater 2 with a pearson r of 0.831 for accuracy and 0.959 for VMRT. ICC were 0.823 (0.652–0.915) and 0.959 (0.913–0.981) for both outcomes. The raters showed a SEM of 3.1%, SEP of 4.2% and MDC of 8.5% for accuracy. The same reliability measures for VMRT were 23 ms, 32 ms and 63 ms, respectively.

4 Discussion

This study is the first to determine the test-retest, intra- and inter-rater reliability of a neurocognitive functional performance test in patients with CAI. The reliability of the video-based analysis is considered moderate to good. The data indicates the VMRT

performance was more robust than the accuracy measure across the trials. When the LOA are compared with the MDC, the data indicates the RBT is more precise and sensitive to changes than the raters' score. We are also the first to report on a RBT score sheet's test-retest and inter-rater reliability. The newly designed score sheet was created to make it more time-efficient to evaluate the accuracy of the RBT and, thus, facilitate the use of the RBT in clinical practice. The moderate to excellent reliability measures of the score sheet method show that this novel scoring system can be used as a valid substitute for video analyzes.

The video analysis showed that the test-retest reliability for accuracy and VMRT was moderate. However, there may still be some variability in scores that cannot be attributed to chance or measurement error, which may stem from individual differences, rater bias, or administration issues. It is worth noting, however, that its MDC is relatively large and, thus, that the RBT is not sensitive to detect small changes in performance over time.

Furthermore, when multiple raters are involved, the RBT's accuracy ratings may be influenced by chance or by variations in the raters' rigor. The observed variability is more likely due to chance as the SEM is relatively small (4.6%). The high MDCs (12.7%) and significant LOA (−15%;11%) might explain its moderate reliability for accuracy, however. VMRT outcomes were more robust when rated by multiple raters.

The RBT maintains consistency when assessing performance over time. Its wider LOA range, however, raises the possibility that accuracy ratings may change significantly over time. In contrast, the score sheets show good to excellent reliability for inter-rater and test-retest comparisons. However, some unexplained variability over time remains present.

Under the current conceptualisations of the PROMIS health organization to be able to use a measuring instrument, acceptable

minimal ICC standards for reliability coefficients are ≥0.70 for group comparisons and between 0.90 and 0.95 for individual comparisons (102). When we apply these recommendations to the current study, the test-retest reliability results of the RBT using either scoring method do not meet the criteria for group analyzes or follow-up measures in scientific research. These results are in contrast with the study of Tassignon et al. (80), which tested the reliability of the RBT in healthy recreational athletes and found an ICC ≥ 0.70. However, several factors may contribute to the increased variability of the results, and these will be explored in the limitation section. Additionally, other conclusions in line with the study of Tassignon et al. can be drawn from this study: the reliability increases if the same rater scores the RBT and another rater can vouch if it is not practically possible to always use the same rater, especially when interested in the VMRT outcome.

Our results imply there may be room for improvement when designing a reliable neurocognitive functional performance test. Changes in the test administration process may reduce variability and increase sensitivity to detect small changes in performance over time. The findings also highlight the importance of using different reliability outcome measures, such as the LOA, to interpret results accurately and account for variability between test sessions or raters. Overall, this emphasizes the need for continued research and refinement of assessment methods to measure neurocognitive components accurately and reliably in functional performance tests.

Despite the need for caution when interpreting results, the moderate to good reliability of the RBT, and, particularly, the good to excellent reliability of the score sheet method suggest the test can be useful to assess certain aspects of cognitive function in a clinical setting. However, it is important to keep in mind the limitations of the test, especially when the goal is to detect small changes over time. Thus, while the RBT and the newly developed score sheet method may be valuable tools for clinicians, continued efforts are necessary to improve their accuracy and sensitivity in measuring cognitive function and dynamic balance ability.

4.1 Limitations and future research

It is important to keep in mind that reliability is a measure of consistency, not accuracy. Even if there is some variability in the scores due to rater differences, this does not necessarily mean that the measure itself is inaccurate or invalid. It means that there may be some inconsistency in how the measure is being administered or scored across different raters or ratings which could potentially impact the interpretation of the scores. This is especially relevant as even small differences between accuracy ratings can have a significant impact on the overall test score because the MDC is relatively high (8.5–18%). Therefore, further defining errors could minimize rater variability and improve the reliability of the RBT test. Another potential reason to explain this increased variability compared to healthy recreational athletes (80), might be the inherent increased variability in postural control strategies used by patients with CAI (103–108). It might be driven by the neuroplastic changes they experience related to their ligamentous ankle injuries (109, 110). Additionally, the included participants had quite a large standard deviation in IPAQ score (SD = 2322MET), which is more than 5 times the standard deviation of the IPAQ score of the participants of the

study of Tassignon et al. This is relevant as the ICC estimates depend on the characteristics of the population and may explain the increased variability.

Despite these factors, the RBT still shows strong reliability measures regardless of reliability type. The SEM for RBT outcomes is small, especially considering the degree of variability of the data (VMRT: 40–59 ms, accuracy: 4.4–6.6%). This suggests an increased precision and validity of the RBT. In addition, the mean difference between tests, ratings or raters is even smaller than the SEM (VMRT: 4–19 ms, accuracy: −2.7–1.2%) indicating very little bias and an increased confidence of the results between measurements. The RBT may not be best suited for group analyzes in research according to the PROMIS health organization, but it remains a useful tool for patients with CAI. The addition of neurocognitive load to functional performance tests brings these exercises closer to the sports context due the demand of being able to respond and adapt to a changing environment (60, 79, 111). Evidence emerged suggesting that lower neurocognitive performance is associated with an increased risk of ankle injury (112, 113). As demonstrated in a systematic review, patients with CAI (110) show adaptations related to their injury and could benefit from neurocognitive training to cope with neuroplasticity related to ligamentous ankle injuries. However, there are no studies reporting on the benefits of neurocognitive training on LAS injury risk yet. Future studies should investigate whether neurocognitive training could positively influence and, thus, reduce injury risk as further research is needed to understand the different modalities of this approach.

5 Conclusion

The RBT is a neurocognitive functional performance test which can be analyzed either video-based or via score sheet method. The video-based analyzes revealed moderate test-retest reliability, but good intra-rater reliability in patients with CAI. When multiple raters evaluate the RBT performance, the VMRT outcome is more reliable than the accuracy outcome. Additionally, the RBT shows robust SEM and mean difference measures. The score sheet method showed similar reliability: over time the reliability was moderate, however, the reliability between raters was good to excellent. Also, the RBT can be a useful tool to assess and monitor balance in patients with CAI as it incorporates neurocognitive elements: environmental perception and decision-making.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics Committee of the UZ Brussel/Vrije Universiteit Brussel. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

AM: Writing – original draft, Writing – review & editing. EL: Writing – original draft, Writing – review & editing. JV: Writing – original draft, Writing – review & editing. KP: Writing – original draft, Writing – review & editing. RM: Writing – original draft, Writing – review & editing. BR: Writing – original draft, Writing – review & editing. BT: Writing – original draft, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Zhang C, Chen N, Wang J, Zhang Z, Jiang C, Chen Z, et al. The prevalence and characteristics of chronic ankle instability in elite athletes of different sports: A cross-sectional study. *J Clin Med*. (2022) 11:7478. doi: 10.3390/jcm11247478

2. Lin CI, Houtenbos S, Lu YH, Mayer F, Wippert PM. The epidemiology of chronic ankle instability with perceived ankle instability- a systematic review. *J Foot Ankle Res*. (2021) 14:41. doi: 10.1186/s13047-021-00480-w

3. Wikstrom EA, Cain MS, Chandran A, Song K, Regan T, Migel K, et al. Lateral ankle sprain and subsequent ankle sprain risk: A systematic review. *J Athl Train*. (2021) 56:578–8510. doi: 10.4085/1062-6050-168-20

4. Miklovic TM, Donovan L, Protzuk OA, Kang MS, Feger MA. Acute lateral ankle sprain to chronic ankle instability: a pathway of dysfunction. *Phys Sportsmed*. (2018) 46:116–22. doi: 10.1080/00913847.2018.1409604

5. Dejong AF, Koldenhoven RM, Hertel J. Proximal adaptations in chronic ankle instability: systematic review and Meta-analysis. *Med Sci Sports Exerc*. (2020) 52:1563–75. doi: 10.1249/MSS.0000000000002282

6. Alghadir AH, Iqbal ZA, Iqbal A, Ahmed H, Ramteke SU. Effect of chronic ankle sprain on pain, range of motion, proprioception, and balance among athletes. *Int J Environ Res Public Health*. (2020) 17:5318. doi: 10.3390/ijerph17155318

7. Michels F, Wastyn H, Pottel H, Stockmans F, Vereecke E, Matricali G. The presence of persistent symptoms 12 months following a first lateral ankle sprain: A systematic review and meta-analysis. *Foot Ankle Surg*. (2022) 28:817–26. doi: 10.1016/j.fas.2021.12.002

8. Lorenzo-Sanchez-Aguilera C, Rodriguez-Sanz D, Gallego-Izquierdo T, Lazaro-Navas I, Plaza-Rodriguez J, Navarro-Santana M, et al. Neuromuscular Mechanosensitivity in subjects with chronic ankle sprain: A cross-sectional study. *Pain Med*. (2020) 21:1991–810. doi: 10.1093/pm/pny299

9. Hertel J, Corbett RO. An updated model of chronic ankle instability. *J Athl Train*. (2019) 54:572–8810. doi: 10.4085/1062-6050-344-18

10. Emamvirdi M, Hosseinzadeh M, Letafatkar A, Thomas AC, dos'Santos T, Smania N, et al. Comparing kinematic asymmetry and lateral step-down test scores in healthy, chronic ankle instability, and patellofemoral pain syndrome female basketball players: a cross-sectional study. *Sci Rep*. (2023) 13:12412. doi: 10.1038/s41598-023-39625-1

11. Tajdini H, Mantashloo Z, Thomas AC, Letafatkar A, Rossettini G. Inter-limb asymmetry of kinetic and electromyographic during walking in patients with chronic ankle instability. *Sci Rep*. (2022) 12:3928. doi: 10.1038/s41598-022-07975-x

12. Houston MN, Van Lunen BL, Hoch MC. Health-related quality of life in individuals with chronic ankle instability. *J Athl Train*. (2014) 49:758–6310. doi: 10.4085/1062-6050-49.3.54

13. Gribble PA, Bleakley CM, Caulfield BM, Docherty CL, Fourchet F, Fong DT, et al. Consensus statement of the international ankle consortium: prevalence, impact and long-term consequences of lateral ankle sprains. *Br J Sports Med*. (2016) 50:1493–5. doi: 10.1136/bjsports-2016-096188

14. Al Mahrouqi MM, Mac Donald DA, Vicenzino B, Smith MD. Quality of life, function and disability in individuals with chronic ankle symptoms: a cross-sectional online survey. *J Foot Ankle Res*. (2020) 13:67. doi: 10.1186/s13047-020-00432-w

15. Lee I, Lee SY, Ha S. Alterations of lower extremity function, health-related quality of life, and spatiotemporal gait parameters among individuals with chronic ankle instability. *Phys Ther Sport*. (2021) 51:22–8. doi: 10.1016/j.ptsp.2021.06.006

16. Kosik KB, Johnson NF, Terada M, Thomas-Fenwick AC, Mattacola CG, Gribble PA. Health-related quality of life among middle-aged adults with chronic ankle instability, copers, and uninjured controls. *J Athl Train*. (2020) 55:733–810. doi: 10.4085/1062-6050-190-19

17. Raw V, De Keizer G, Van Dijk CN. Long-term follow-up of inversion trauma of the ankle. *Arch Orthop Trauma Surg*. (1995) 114:92–6. doi: 10.1007/BF00422833

18. Konradsen L, Bech L, Ehrenbjerg M, Nickelsen T. Seven years follow-up after ankle inversion trauma. *Scand J Med Sci Sports*. (2002) 12:129–35. doi: 10.1034/j.1600-0838.2002.02104.x

19. Hubbard-Turner T, Turner MJ. Physical activity levels in college students with chronic ankle instability. *J Athl Train*. (2015) 50:742–10. doi: 10.4085/1062-6050-50.3.05

20. Holland B, Needle AR, Battista RA, West ST, Christiana RW. Physical activity levels among rural adolescents with a history of ankle sprain and chronic ankle instability. *PLoS One*. (2019) 14:e0216243. doi: 10.1371/journal.pone.0216243

21. Donovan L, Hetzel S, Laufenberg CR, McGuine TA. Prevalence and impact of chronic ankle instability in adolescent athletes. *Orthop J Sports Med*. (2020) 8:2325967119900962. doi: 10.1177/2325967119900962

22. Golditz T, Steib S, Pfeifer K, Uder M, Gelse K, Janka R, et al. Functional ankle instability as a risk factor for osteoarthritis: using T2-mapping to analyze early cartilage degeneration in the ankle joint of young athletes. *Osteoarthr Cartil*. (2014) 22:1377–85. doi: 10.1016/j.joca.2014.04.029

23. Valderrabano V, Hintermann B, Horisberger M, Fung TS. Ligamentous posttraumatic ankle osteoarthritis. *Am J Sports Med*. (2006) 34:612–20. doi: 10.1177/0363546505281813

24. Kosik KB, Terada M, McCann R, Thomas A, Johnson N, Gribble P. Decreased perceived ankle and knee joint health in individuals with perceived chronic ankle instability. *Knee Surg Sports Traumatol Arthrosc*. (2020) 28:177–8310. doi: 10.1007/s00167-018-5163-4

25. Song K, Pietrosimone B, Blackburn JT, Padua DA, Tennant JN, Wikstrom EA. Acute Talar cartilage deformation in those with and without chronic ankle instability. *Med Sci Sports Exerc*. (2021) 53:1228–34. doi: 10.1249/MSS.0000000000002572

26. Hirose K, Murakami G, Minowa T, Kura H, Yamashita T. Lateral ligament injury of the ankle and associated articular cartilage degeneration in the talocrural joint: anatomic study using elderly cadavers. *J Orthop Sci*. (2004) 9:37–4310. doi: 10.1007/s00776-003-0732-9

27. Lee S, Song K, Lee SY. Epidemiological study of post-traumatic ankle osteoarthritis after ankle sprain in 195, 393 individuals over middle age using the National Health

Insurance Database: A retrospective design. *J Sci Med Sport*. (2022) 25:129–33. doi: 10.1016/j.jsams.2021.08.018

28. Tassignon B, Verschueren J, Delahunt E, Smith M, Vicenzino B, Verhagen E, et al. Criteria-based return to sport decision-making following lateral ankle sprain injury: a systematic review and narrative synthesis. *Sports Med*. (2019) 49:601–1910. doi: 10.1007/s40279-019-01071-3

29. Smith MD, Vicenzino B, Bahr R, Bandholm T, Cooke R, Mendonca LM, et al. Return to sport decisions after an acute lateral ankle sprain injury: introducing the PAASS framework-an international multidisciplinary consensus. *Br J Sports Med*. (2021) 55:1270–6. doi: 10.1136/bjsports-2021-104087

30. Powell C, Jensen J, Johnson S. Functional performance measures used for return-to-sport criteria in youth following lower-extremity injury. *J Sport Rehabil*. (2018) 27:581–90. doi: 10.1123/jsr.2017-0061

31. Picot B, Hardy A, Terrier R, Tassignon B, Lopes R, Fourchet F. Which functional tests and self-reported questionnaires can help clinicians make valid return to sport decisions in patients with chronic ankle instability? A narrative review and expert opinion. *Front Sports Act Living*. (2022) 4:902886. doi: 10.3389/fspor.2022.902886

32. Kalatakis-Dos-Santos AE, de Paula Gomes CAF, Pontes-Silva A, Mendes LP, de Oliveira SG, Goncalves MC, et al. Fear of return to sport scale (FRESS): a new instrument for use in injured professional or recreational athletes in rehabilitation. *Sport Sci Health*. (2022) 5:1–1010. doi: 10.1007/s11332-022-00975-4

33. Ardern CL, Glasgow P, Schneiders A, Witvrouw E, Clarsen B, Cools A, et al. Consensus statement on return to sport from the first world congress in sports physical therapy, Bern. *Br J Sports Med*. (2016) 50:853–64. doi: 10.1136/bjsports-2016-096278

34. Vereijken A, Aerts I, Jetten J, Tassignon B, Verschueren J, Meeusen R, et al. Association between functional performance and return to performance in high-impact sports after lower extremity injury: A systematic review. *J Sport Sci Med*. (2020) 19:564–76.

35. Hegedus EJ, McDonough SM, Bleakley C, Baxter D, Cook CE. Clinician-friendly lower extremity physical performance tests in athletes: a systematic review of measurement properties and correlation with injury. Part 2--the tests for the hip, thigh, foot and ankle including the star excursion balance test. *Br J Sports Med*. (2015) 49:649–56. doi: 10.1136/bjsports-2014-094341

36. Powden CJ, Dodds TK, Gabriel EH. The reliability of the star excursion balance test and lower quarter Y-balance test in healthy adults: A systematic review. *Int J Sports Phys Ther*. (2019) 14:683–94. doi: 10.26603/ijspt20190683

37. Plisky P, Schwartkopf-Phifer K, Huebner B, Garner MB, Bullock G. Systematic review and Meta-analysis of the Y-balance test lower quarter: reliability, discriminant validity, and predictive validity. Int J sports. *Phys Ther*. (2021) 16:1190–209. doi: 10.26603/001c.27634

38. Pierobon A, Raguzzi I, Solino S, Salzberg S, Vuoto T, Gilgado D, et al. Minimal detectable change and reliability of the star excursion balance test in patients with lateral ankle sprain. *Physiother Res Int*. (2020) 25:e1850. doi: 10.1002/pri.1850

39. Scinicarelli G, Trofenik M, Frobose I, Wilke C. The reliability of common functional performance tests within an experimental test battery for the lower extremities. *Sports (Basel)*. (2021) 9:100. doi: 10.3390/sports9070100

40. Gribble PA, Hertel J, Plisky P. Using the star excursion balance test to assess dynamic postural-control deficits and outcomes in lower extremity injury: a literature and systematic review. *J Athl Train*. (2012) 47:339–5710. doi: 10.4085/1062-6050-47.3.08

41. Chang WD, Chou LW, Chang NJ, Chen S. Comparison of functional movement screen, star excursion balance test, and physical fitness in junior athletes with different sports injury risk. *Biomed Res Int*. (2020) 2020:1–8. doi: 10.1155/2020/8690540

42. Dallinga JM, Benjaminse A, Lemmink KAPM. Which screening tools can predict injury to the lower extremities in team sports? A systematic review. *Sports Med*. (2012) 42:791–815. doi: 10.1007/BF03262295

43. Bennett H, Chalmers S, Milanese S, Fuller J. The association between Y-balance test scores, injury, and physical performance in elite adolescent Australian footballers. *J Sci Med Sport*. (2022) 25:306–11. doi: 10.1016/j.jsams.2021.10.014

44. Olmsted LC, Carcia CR, Hertel J, Shultz SJ. Efficacy of the star excursion balance tests in detecting reach deficits in subjects with chronic ankle instability. *J Athl Train*. (2002) 37:501–6.

45. Plisky PJ, Gorman PP, Butler RJ, Kiesel KB, Underwood FB, Elkins B. The reliability of an instrumented device for measuring components of the star excursion balance test. *N Am J Sports Phys Ther*. (2009) 4:92–9.

46. McCann RS, Crossett ID, Terada M, Kosik KB, Bolding BA, Gribble PA. Hip strength and star excursion balance test deficits of patients with chronic ankle instability. *J Sci Med Sport*. (2017) 20:992–6. doi: 10.1016/j.jsams.2017.05.005

47. Doherty C, Bleakley C, Hertel J, Caulfield B, Ryan J, Delahunt E. Dynamic balance deficits 6 months following first-time acute lateral ankle sprain: A laboratory analysis. *J Orthop Sports Phys Ther*. (2015) 45:626–33. doi: 10.2519/jospt.2015.5653

48. Doherty C, Bleakley CM, Hertel J, Caulfield B, Ryan J, Delahunt E. Laboratory measures of postural control during the star excursion balance test after acute first-time lateral ankle sprain. *J Athl Train*. (2015) 50:651–6410. doi: 10.4085/1062-6050-50.1.09

49. Doherty C, Bleakley C, Hertel J, Caulfield B, Ryan J, Delahunt E. Dynamic balance deficits in individuals with chronic ankle instability compared to ankle sprain copers 1

year after a first-time lateral ankle sprain injury. *Knee Surg Sports Traumatol Arthrosc*. (2016) 24:1086–9510. doi: 10.1007/s00167-015-3744-z

50. Kosik KB, Johnson NF, Terada M, Thomas AC, Mattacola CG, Gribble PA. Decreased dynamic balance and dorsiflexion range of motion in young and middle-aged adults with chronic ankle instability. *J Sci Med Sport*. (2019) 22:976–80. doi: 10.1016/j.jsams.2019.05.005

51. Bastien M, Moffet H, Bouyer LJ, Perron M, Hébert LJ, Leblond J. Alteration in global motor strategy following lateral ankle sprain. *BMC Musculoskelet Disord*. (2014) 15:436. doi: 10.1186/1471-2474-15-436

52. Gribble PA, Hertel J, Denegar CR, Buckley WE. The effects of fatigue and chronic ankle instability on dynamic postural control. *J Athl Train*. (2004) 39:321–9.

53. Plante JE, Wikstrom EA. Differences in clinician-oriented outcomes among controls, copers, and chronic ankle instability groups. *Phys Ther Sport*. (2013) 14:221–6. doi: 10.1016/j.ptsp.2012.09.005

54. Linens SW, Ross SE, Arnold BL, Gayle R, Pidcoe P. Postural-stability tests that identify individuals with chronic ankle instability. *J Athl Train*. (2014) 49:15–2310. doi: 10.4085/1062-6050-48.6.09

55. Sefton JM, Hicks-Little CA, Hubbard TJ, Clemens MG, Yengo CM, Koceja DM, et al. Sensorimotor function as a predictor of chronic ankle instability. *Clin Biomech (Bristol, Avon)*. (2009) 24:451–8. doi: 10.1016/j.clinbiomech.2009.03.003

56. Ko J, Rosen AB, Brown CN. Functional performance deficits in adolescent athletes with a history of lateral ankle sprain (s). *Phys Ther Sport*. (2018) 33:125–32. doi: 10.1016/j.ptsp.2018.07.010

57. Ko J, Rosen AB, Brown CN. Comparison between single and combined clinical postural stability tests in individuals with and without chronic ankle instability. *Clin J Sport Med*. (2017) 27:394–9. doi: 10.1097/Jsm.0000000000000354

58. Chimera NJ, Warren M. Use of clinical movement screening tests to predict injury in sport. *World J Orthop*. (2016) 7:202–17. doi: 10.5312/wjo.v7.i4.202

59. Nuri L, Shadmehr A, Ghotbi N, Attarbashi MB. Reaction time and anticipatory skill of athletes in open and closed skill-dominated sport. *Eur J Sport Sci*. (2013) 13:431–6. doi: 10.1080/17461391.2012.738712

60. Glasgow P, Bleakley CM, Phillips N. Being able to adapt to variable stimuli: the key driver in injury and illness prevention? *Br J Sports Med*. (2013) 47:64–510. doi: 10.1136/bjsports-2012-091960

61. Wilke J, Groneberg DA. Neurocognitive function and musculoskeletal injury risk in sports: A systematic review. *J Sci Med Sport*. (2022) 25:41–5. doi: 10.1016/j.jsams.2021.07.002

62. Hulsdunker T, Struder HK, Mierau A. Neural correlates of expert Visuomotor performance in badminton players. *Med Sci Sports Exerc*. (2016) 48:2125–34. doi: 10.1249/MSS.0000000000001010

63. Voss MW, Kramer AF, Basak C, Prakash RS, Roberts B. Are expert athletes 'expert' in the cognitive laboratory? A meta-analytic review of cognition and sport expertise. *Appl Cogn Psychol*. (2010) 24:812–26. doi: 10.1002/acp.1588

64. Tanja van de WaterHuijgen B, Faber I, Elferink-Gemser M. Assessing cognitive performance in badminton players: A reproducibility and validity study. *J Hum Kinet*. (2017) 55:149–59. doi: 10.1515/hukin-2017-0014

65. Piccoli A, Rossettini G, Cecchetto S, Viceconti A, Ristori D, Turolla A, et al. Effect of attentional focus instructions on motor learning and performance of patients with central nervous system and musculoskeletal disorders: a systematic review. *J Funct Morphol Kinesiol*. (2018) 3:40. doi: 10.3390/jfmk3030040

66. Burcal CJ, Wikstrom EA. Cognitive loading-induced sway alterations are similar in those with chronic ankle instability and uninjured controls. *Gait Posture*. (2016) 48:95–8. doi: 10.1016/j.gaitpost.2016.05.004

67. Burcal CJ, Wikstrom EA. Examining the relationship between chronic ankle instability symptoms and dual-task balance performance. *Int J Athl Ther Train*. (2017) 22:34–9. doi: 10.1123/ijatt.2016-0033

68. Elshorbagy RT, Balbaa AE-DA-H, Ayad KE-S, Allam NM, Eladl HM, Allah WRA. Cognitive task versus focus of attention on dynamic postural control in recurrent ankle sprains. Journal of advanced pharmacy education and. *Research*. (2022) 12:6–1010. doi: 10.51847/OYUrFadR58

69. McGrath ML, Yentes JM, Rosen AB. Cognitive loading produces similar change in postural stability in patients with chronic ankle instability and controls. Athletic Training & Sports. *Health Care*. (2020) 12:249–5610. doi: 10.3928/19425864-20200610-02

70. Mohamadi S, Ebrahimi I, Salavati M, Dadgoo M, Jafarpisheh AS, Rezaeian ZS. Attentional demands of postural control in chronic ankle instability, copers and healthy controls: A controlled cross-sectional study. *Gait Posture*. (2020) 79:183–8. doi: 10.1016/j.gaitpost.2020.03.007

71. Rahnama L, Salavati M, Akhbari B, Mazaheri M. Attentional demands and postural control in athletes with and without functional ankle instability. *J Orthop Sports Phys Ther*. (2010) 40:180–7. doi: 10.2519/jospt.2010.3188

72. Rosen AB, Than NT, Smith WZ, Yentes JM, McGrath ML, Mukherjee M, et al. Attention is associated with postural control in those with chronic ankle instability. *Gait Posture*. (2017) 54:34–8. doi: 10.1016/j.gaitpost.2017.02.023

73. Shiravi Z, Talebian Moghadam S, Hadian MR, Olyaei G. Effect of cognitive task on postural control of the patients with chronic ankle instability during single and

double leg standing. *J Bodyw Mov Ther*. (2017) 21:58–6210. doi: 10.1016/j.jbmt.2016.05.001

74. Taghavi Asl A, Shojaedin SS, Hadadnezhad M. Comparison of effect of wobble board training with and without cognitive intervention on balance, ankle proprioception and jump landing kinetic parameters of men with chronic ankle instability: a randomized control trial. *BMC Musculoskelet Disord*. (2022) 23:888. doi: 10.1186/s12891-022-05706-x

75. Tohidast SA, Bagheri R, Safavi-Farokhi Z, Khaleghi Hashemian M, Delkhosh CT. The effects of acute and long-term whole-body vibration training on the postural control during cognitive task in patients with chronic ankle instability. *J Sport Rehabil*. (2021) 30:1121–8. doi: 10.1123/jsr.2021-0034

76. Rosen AB, McGrath ML, Maerlender AL. Males with chronic ankle instability demonstrate deficits in neurocognitive function compared to control and copers. *Res Sports Med*. (2021) 29:116–28. doi: 10.1080/15438627.2020.1723099

77. Mohammadi N, Hadian MR, Olyaei G. Information processing speed and accuracy in basketball players with and without functional ankle instability. *J Mod Rehabil*. (2020) 13:179–84. doi: 10.32598/JMR.13.3.179

78. Song K, Hoch JM, Quintana C, Heebner NR, Hoch MC. Slower visuomotor reaction time in division-I collegiate athletes with a history of ankle sprain. *Res Sports Med*. (2021) 31:1–910. doi: 10.1080/15438627.2021.1996361

79. Verschueren J, Tassignon B, Pluym B, Van Cutsem J, Verhagen E, Meeusen R. Bringing context to balance: development of a reactive balance test within the injury prevention and return to sport domain. *Arch Physiother*. (2019) 9:6. doi: 10.1186/s40945-019-0057-4

80. Tassignon B, Verschueren J, De Wachter J, Maricot A, De Pauw K, Verhagen E, et al. Test-retest, intra- and inter-rater reliability of the reactive balance test in healthy recreational athletes. *Phys Ther Sport*. (2020) 46:47–5310. doi: 10.1016/j.ptsp.2020.08.010

81. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hrobjartsson A, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol*. (2011) 64:96–106. doi: 10.1016/j.jclinepi.2010.03.002

82. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. *Br J Math Stat Psychol*. (2008) 61:29–48. doi: 10.1348/000711006X126600

83. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med*. (1998) 17:101–10. doi: 10.1002/(sici)1097-0258(19980115)17:1<101::aid-sim727>3.0.co; 2-e

84. Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc*. (2003) 35:1381–9510. doi: 10.1249/01.MSS.0000078924.61453.FB

85. Gribble PA, Delahunt E, Bleakley C, Caulfield B, Docherty CL, Fourchet F, et al. Selection criteria for patients with chronic ankle instability in controlled research: a position statement of the international ankle consortium. *J Orthop Sports Phys Ther*. (2013) 43:585–91. doi: 10.2519/jospt.2013.0303

86. Greenberg ET, Barle M, Glassmann E, Jung M-K. Interrater and test-retest reliability of the Y balance test in healthy, early adolescent female athletes. International journal of sports. *Phys Ther*. (2019) 14:204–13. doi: 10.26603/ijspt20190204

87. Mishra P, Pandey CM, Singh U, Gupta A, Sahu C, Keshri A. Descriptive statistics and normality tests for statistical data. *Ann Card Anaesth*. (2019) 22:67–72. doi: 10.4103/aca.ACA_157_18

88. Koo TK, Li MY. A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. *J Chiropr Med*. (2016) 15:155–63. doi: 10.1016/j.jcm.2016.02.012

89. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*. (2005) 19:231–4010. doi: 10.1519/15184.1

90. McManus IC. The misinterpretation of the standard error of measurement in medical education: a primer on the problems, pitfalls and peculiarities of the three different standard errors of measurement. *Med Teach*. (2012) 34:569–76. doi: 10.3109/0142159X.2012.670318

91. Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *BMJ*. (1992) 304:1491–4. doi: 10.1136/bmj.304.6840.1491

92. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med*. (2000) 30:1–15. doi: 10.2165/00007256-200030010-00001

93. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clin Rehabil*. (1998) 12:187–99. doi: 10.1191/026921598672178340

94. Stratford PW, Riddle DL. When minimal detectable change exceeds a diagnostic test–based threshold change value for an outcome measure: resolving the conflict. *Phys Ther*. (2012) 92:1338–47. doi: 10.2522/ptj.20120002

95. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. *Phys Ther*. (2006) 86:735–43. doi: 10.1093/ptj/86.5.735

96. Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL. Smallest real difference, a link between reproducibility and responsiveness. *Qual Life Res*. (2001) 10:571–8. doi: 10.1023/A:1013138911638

97. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med*. (1998) 4:217–38. doi: 10.2165/00007256-199826040-00002

98. de Vet HCW, Bouter LM, Bezemer PD, Beurskens AJHM. Reproducibility and responsiveness of evaluative outcome measures. *Int J Technol Assess Health Care*. (2001) 17:479–87. doi: 10.1017/S0266462301107038

99. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. (1999) 8:135–60. doi: 10.1177/096228029900800204

100. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. (1986) 1:908–9. doi: 10.1016/S0140-6736(86)90837-8

101. Portney L, Mary W. *Foundations of clinical research: Applications to practice. 3rd* ed. Upper Saddle River, New Jersey: Pearson Education, Inc. (2009). 7458 p.

102. Aaronson N, Alonso J, Burnam A, Lohr KN, Patrick DL, Perrin E, et al. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res*. (2002) 11:193–205. doi: 10.1023/a:1015291021312

103. Esteves J, Dinis R, Fernandes O, Castro MA, Oliveira R, Pezarat-Correia P. Differences in postural control between healthy and subjects with chronic ankle instability. *Phys Ther Sport*. (2022) 56:8–1410. doi: 10.1016/j.ptsp. 2022.05.014

104. dos Santos MJ, Gorges AL, Rios JL. Individuals with chronic ankle instability exhibit decreased postural sway while kicking in a single-leg stance. *Gait Posture*. (2014) 40:231–6. doi: 10.1016/j.gaitpost.2014.04.002

105. Chan LYT, Sim YTN, Gan FK. Bin Abd Razak HR. effect of chronic ankle instability on lower extremity kinematics, dynamic postural stability, and muscle activity during unilateral jump-landing tasks: A systematic review and meta-analysis. *Phys Ther Sport*. (2022) 55:176–88. doi: 10.1016/j.ptsp.2022.04.005

106. Kim H, Son SJ, Seeley MK, Hopkins JT. Altered movement biomechanics in chronic ankle instability, Coper, and control groups: energy absorption and distribution implications. *J Athl Train*. (2019) 54:708–1710. doi: 10.4085/1062-6050-483-17

107. Xa X, Wang Y, Xu X, Li H, Li Q, Na Y, et al. Postural control deficits during static single-leg stance in chronic ankle instability: a systematic review and Meta-analysis. *Sports Health*. 16:29–37. doi: 10.1177/19417381231152490

108. Song K, Burcal CJ, Hertel J, Wikstrom EA. Increased visual use in chronic ankle instability: A Meta-analysis. *Med Sci Sports Exerc*. (2016) 48:2046–56. doi: 10.1249/MSS.0000000000000992

109. Kim KM, Kim JS, Cruz-Diaz D, Ryu S, Kang M, Taube W. Changes in spinal and corticospinal excitability in patients with chronic ankle instability: A systematic review with Meta-analysis. *J Clin Med*. (2019) 8:1037. doi: 10.3390/jcm8071037

110. Maricot A, Dick E, Walravens A, Pluym B, Lathouwers E, De Pauw K, et al. Brain neuroplasticity related to lateral ankle ligamentous injuries: A systematic review. *Sports Med*. (2023) 53:1423–4310. doi: 10.1007/s40279-023-01834-z

111. Millikan N, Grooms DR, Hoffman B, Simon JE. The development and reliability of 4 clinical neurocognitive single-leg hop tests: implications for return to activity decision-making. *J Sport Rehabil*. (2019) 28:536–4410. doi: 10.1123/jsr. 2018-0037

112. CH D, Zaremski JL, Vincent HK, Vincent KR. Effect of neurocognition and concussion on musculoskeletal injury risk. *Curr Sports Med Rep*. (2015) 14:194–9. doi: 10.1249/JSR.0000000000000157

113. Griffin LY, Agel J, Albohm MJ, Arendt EA, Dick RW, Garrett WE, et al. Noncontact anterior cruciate ligament injuries: risk factors and prevention strategies. *J Am Acad Orthop Surg*. (2000) 8:141–50. doi: 10.5435/00124635-200005000-00001