Title Intra-Rater Reliability of the Multiple Single-Leg Hop-Stabilization Test and Relationships with Age, Leg Dominance and Training

Author names and affiliations Dr Leanne Sawle, PhD[1,2] , Dr Jennifer Freeman PhD[1], Professor Jonathan Marsden PhD[1]

[1] Plymouth University. Faculty of Health and Human Sciences, Peninsula Allied Health Centre, Derriford Road, Plymouth, Devon PL6 8BH

[2] Cardiff University. School of Engineering. The Parade, Cardiff CF24 3AA

Corresponding author Dr Leanne Sawle. Cardiff University. School of Engineering. The Parade, Cardiff CF24 3AA. 07903731001 SawleL1@cardiff.ac.uk

Background: Balance is a complex construct, affected by multiple components such as strength and co-ordination. However, whilst assessing an athlete's dynamic balance is an important part of clinical examination there is no gold standard measure. The multiple single-leg hop-stabilization test (MSLHST) is a functional test which may offer a method of evaluating the dynamic attributes of balance, but it needs to show adequate intra-tester reliability.

Purpose: The purpose of this study was to assess the intra-rater reliability of a dynamic balance test, the multiple single-leg hop-stabilization test (MSLHST) on the dominant and non-dominant legs.

Design: Intra-rater reliability study

Methods: Fifteen active participants were tested twice with a 10-minute break between tests. The outcome measure was the multiple single-leg hop-stabilization test score, based on a clinically assessed numerical scoring system. Results were analysed using an Intraclass Correlations Coefficient ($ICC_{2,1}$) and Bland-Altman plots. Regression analyses explored relationships between test scores, leg dominance, age and training (an alpha level of $p = 0.05$ was selected).

Results: ICCs for intra-rater reliability were 0.85 for the dominant and non-dominant legs (confidence intervals = 0.62-0.95 and 0.61-0.95 respectively). Bland-Altman plots showed scores within two standard deviations. A significant correlation was observed between the dominant and non-dominant leg on balance scores (R^2=0.49, $p<0.05$), and better balance was associated with younger participants in their non-dominant leg (R^2=0.28, $p<0.05$) and their dominant leg (R^2=0.39, $p<0.05$) and a higher number of hours spent training for the non-dominant leg R^2=0.37, $p<0.05$).

Conclusion: The multiple single-leg hop-stabilisation test demonstrated strong intra-tester reliability with active participants. Younger participants who trained more, have better balance scores. This test may be a useful measure for evaluating the dynamic attributes of balance.

Level of Evidence: 3

Key words: Assessment, balance, reliability, hop testing

INTRODUCTION

Normal balance requires the interaction between multisensory organ systems (proprioceptive, visual and vestibular[1]) and the brain and spinal cord, which ultimately control the multi-joint musculoskeletal system. [2-4] These systems can be affected by factors such as nutrition, [5] age, [6] injury [7] and disease. [8] At an optimal level they work to maintain the center of gravity within a defined base of support, as well as the task specific orientation of body parts. [9]

Within sports medicine, assessing an athlete's balance is an important part of a clinical examination. [10] It is within this domain that an emphasis is placed upon proprioceptive / balance exercises as both a tool for injury prevention [11] and as a rehabilitation strategy. [10] However, the physical demands of sport are extremely diverse, and balance and postural control appear to be influenced by other performance attributes. For example, strength training programs lead to significant improvements in both static (Romberg) and dynamic (Star Excursion Balance Test) measures of balance. [12]

Despite the implementation of balance training for both injury prevention and rehabilitation, no gold standard outcome measure exists with which to quantify balance within the athletic population. [10] While it is acknowledged that balance can be measured statically or dynamically, [12] the population being examined should direct the nature of the test selected. Furthermore it should not be assumed that static balance ability is positively correlated with dynamic balance performance. [13] Therefore it appears appropriate to use a dynamic measure of balance when examining the athletic population, as all sports require a "dynamic" attribute of balance in some way.

The purpose of looking at athletic balance stems from the results of a series of single case studies evaluating the use of clinically targeted compression in athletes, whereby compression was delivered to the pelvic girdle via a customised orthosis in the form of shorts. Questionnaire responses from the participating athletes suggested that this type of external pelvic compression [14] may have had a positive effect upon balance. [15] In order to investigate whether this is the case, the intention was to incorporate a functional measure of athletic balance in future clinical trials. On the basis of the current literature [10] and discussion with clinical colleagues, it is anticipated that a functional single leg test may be an appropriate measure of dynamic balance.

Previous researchers have found that knee instability is positively correlated with one-legged tests, [16] and that a single leg hopping test can demonstrate good test re-test reliability . [17] The multiple single-leg hop-stabilization test (MSLHST) is a single leg dynamic measure , [18] involving forwards, and diagonal movements in a unipedal stance, that incorporates periods of statically maintaining this stance. Athletes are scored on both a balance and landing scale, according to the errors that they commit in each period of the test; these scores are summed to give the total error score. It has been argued that this type of functional test is important because it challenges athletes in a way which reflects the forces and directions of movement that are integral to sport. [18]

Although this test has been reported to have very good inter-tester reliability (ICC values 0.70-0.92), [18] intra-rater reliability was shown to be lacking. [10] Closer inspection of the intra-rater reliability reveals that this lack of reliability only refers to the balance scores which significantly differed between tests; no significant difference was observed with the landing scores. [10] Further, this study [18] assessed three test sessions, each 48 hours apart; a different scenario to the current intra-rater reliability study in which the testing was completed in one session.

A further consideration for any balance study involving athletes with a lower limb injury is the influence of lower limb dominance. In football, a players' dominant (preferred kicking leg) has been shown to be significantly stronger than their non-dominant leg in terms of hip adductor strength, [19] and hip flexor strength, [20] but not in all muscle groups. [19] It has been suggested that any rehabilitation of injury needs to take leg dominance into consideration. [19]As a strength deficit may potentially contribute to poor balance, it is important that a study considers the role of limb dominance, and examines how this may influence the reliability of the balance measure used.

The purpose of this study was to assess the intra-rater reliability of a dynamic balance test, the multiple single-leg hop-stabilization test (MSLHST) on the dominant and non-dominant legs.

A secondary purpose was to explore whether relationships exist between the MSLHST scores and leg dominance, age, and time spent engaging in exercise (training).

METHODS

Design

An intra-rater reliability study was undertaken. All of the testing was undertaken by a single investigator, using portable equipment; the test was scored in "real time" while the balance measure was being performed.

Participants

A convenience sample of volunteers was recruited from Plymouth University staff and students, and from local sports clubs. To maximise recruitment the study was conducted at the University (Human Movement Laboratory) to accommodate the staff and student participants. Ethical approval was gained from a local University Ethics Commitee (Plymouth University).

Eligibility Criteria

To be included, subjects had to be over the age of 18, and able to give informed consent, be self-declared as healthy, and have sustained no lower limb musculoskeletal injuries in the prior three months. Subjects were exluded if they were pregnant, had a current illness / unresolved condition , or had any neurological, musculoskeletal or cardiorespiratory impairment.

Sample Size

Reliability coefficients greater than 0.7 are deemed to be acceptable for most clinical trials. [21] A power calculation indicated that 15 people were needed to be recruited in order to demonstrate an ICC of >0.7 (power = 0.88; $\alpha = 0.05$). This is in keeping with the work of Fleiss [22] and their discussion of the numbers required for a reliability study involving quantitative measures.

Participant Characteristics

Participant demographics (age, gender, height, weight), their leg dominance (as defined by which side they would kick a ball), and the average number of hours spent training / performing sports in a week were recorded.

Measurement of the MSLHST

Testing was undertaken in standard sports attire (shorts, t shirt and athletic shoes) and conducted in the same undisturbed environment, in order to minimise external influences and allow for standardization. Standardized written instructions were given to all participants prior to testing; this included photographs of stances. Participants also received verbal instructions from the researcher while viewing the MSLHST set up, and before completing their practice attempts.

The distances between each of the boxes (Table 1) were standardised according to the participants' height. Diagonal distances represented 45% of the participants' height (wearing athletic shoes), and Pythagoras Theorem used to calculate the distances in the frontal plane, for the adjacent boxes. The mat was labelled according to the height related distances prior to testing to ensure that during testing, there was minimal delay in setting up the mat. This was achieved using hook and loop combinations of numbered Velcro® squares.

Table 1. *Hop distances according to height* [23]

Height in Centimetres (cm)	Diagonal Distance (cm)	Adjacent Distance (cm)
150-159.9	70	49
160-169.9	74	53
170-179.9	79	58
180-189.9	83	59
190-199.9	88	62
200-209.9	92	66

One practice attempt on each leg was undertaken for familiarization of the procedure while avoiding fatigue. Both the dominant leg (as defined as the leg that people would prefer to kick a ball with) and the non-dominant leg were tested in a randomized order (randomization was undertaken using the Microsoft Excel 2010 randomization function). After a 10 minute rest, participants were asked to complete the MSLHST again on both legs, in the same order.

The starting position was standardised with the participants standing on one leg with both hands on their iliac crests and eyes facing forwards. Participants were asked to hop to a series of numbered boxes; each with an area of 2.5cm^2 (Figures 1a, 1b). Arm position was standardized throughout the test, with participants asked to keep their hands on their iliac crests. The task was paced by a metronome (with an auditory cue every one second). On landing on each box, participants were asked to maintain their position for five seconds (counted aloud by the investigator). The balance period was defined as the period prior to undertaking each jump and the period one to five seconds after landing and stabilizing the position. The landing period was defined as the one second period immediately after landing, when the participant attempted to stabilize their position.

Previous work [18] has described how any error in either a landing or balance phase was counted as a failure. [18] Errors were scored according to the period in the test in which they were committed i.e. 3 points for an error in a balance period, and 10 points for a landing period error. Testing did not stop following an error; participants continued with the test and all errors were scored.The final test score was the sum of the balance and landing error scores. The MSLHST scoring was defined as:

Balance score. 3 error marks were given for participants committing the following in any balance period:

- Touching the floor with the non-weight bearing limb;
- Removing hands from iliac crests;
- Non-weight bearing limb touching the weight bearing limb;
- Non-weight bearing limb moving into excessive flexion, extension or abduction

(this was defined as movement beyond the predetermined stance (>30 degrees of movement); displayed to the participants in a photographical format).

Landing score. 10 error marks were given for participants committing the following in any landing period:

- Removing hands from iliac crests;
- Foot not covering the numbered square;
- Stumbling on landing;
- Landing foot not facing forwards with 10 degrees of inversion or eversion.

Therefore potential test scores could range from 0 -130 (0-100 for the landing component, and, 0-30 for the balance element).

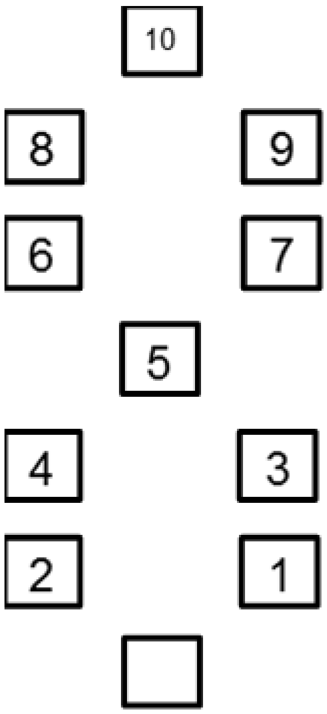

Figure 1a. *A representation of the boxes marked out for the multiple single-leg hop-stabilisation test*

Figure 1b. *A photograph of the testing mat being prepared for variable distances*

Statistical Analyses

Statistical analyses were performed using SPSS 20 for Windows (IBM). Two-way random absolute agreement intra-class correlation (ICC $_{2,1}$) and 95% confidence intervals were used to assess the intra-rater reliability. [24]

Bland Altman plots were presented to show a visual representation of intra-rater reliability. Using more than one measure of reliability has been advised as no one measure is suitable for all reliability studies. [25] ICCs give a relative view of reliability, therefore it has been advised not to draw conclusions before using methods of examining the absolute reliability. [26]

A paired t-test was used to ascertain if there was a significant difference between the balance ability of the dominant and non-dominant leg ($p = <0.05$). Regression analyses were undertaken to explore possible relationships between balance ability on the dominant and non-dominant leg, age and time spent training each week. The strength of the correlation coefficients were interpreted as: 0 = zero, 0.1-0.3 = weak, 0.4-0.6 = moderate, 0.7-0.9 = strong and 1 = perfect. [27]

The time spent training each week was further explored using t tests to determine the possibility of predicting test performance according to the amount of training undertaken (< or > five hours per week). Such a relationship has been observed in previous work, showing that lifelong football trained men demonstrated significantly superior balance to age matched untrained men. [28]

RESULTS

Fifteen participants (males = 8), aged 22-57 participated in the study. The demographics of the tested population are presented in Table 2.

Table 2. *Demographical data*

	Age (yrs)	Weight (kg)	Height (cm)	Gender	Dominant Leg	Average Weekly Training Hours
Mean	32.8	71.4	174.2	Female = 7 Male = 8	Left = 2 Right = 13	5.5
SD	9.2	9.5	7.5			4.3
Range	22-57	53.8-88	162.5-184.5			0.3-14

Table 3 presents the MSLHST score inter-rater reliability ICCs for the dominant and non-dominant leg, along with the 95% CI's. ICCs for both legs = 0.85.

Tables 4 and 5 present the ICCs for the balance and landing scores on each leg. For the non-dominant leg, balance and landing score ICCs were 0.87 and 0.78 respectively. For the dominant leg, ICCs were 0.88 for the balance score, and 0.72 for the landing score.

Table 3. *Intra-rater reliability results. ICC* $(_{2,1})$

		95% Confidence Intervals	
	Intraclass Correlation Coefficient	**Lower Bounds**	**Upper Bounds**
Dominant Leg	0.85	0.62	0.95
Non-Dominant Leg	0.85	0.61	0.95

Table 4. *Intra-rater reliability results for the non-dominant leg balance and landing scores. ICC* $(_{2.1})$

Non-Dominant Leg		**95% Confidence Intervals**	
	Intraclass Correlation Coefficient	**Lower Bounds**	**Upper Bounds**
Landing Score	0.78	0.47	0.92
Balance Score	0.87	0.64	0.95

Table 5. *Intra-rater reliability results for the dominant leg balance and landing scores. ICC (2.1)*

Dominant Leg		95% Confidence Intervals	
	Intraclass Correlation Coefficient	Lower Bounds	Upper Bounds
Landing Score	0.72	0.34	0.90
Balance Score	0.88	0.83	0.96

Figures 2 and 3 present visual representations of the intra-rater differences in scores for the dominant and non-dominant legs. Offer a summary statement here too.

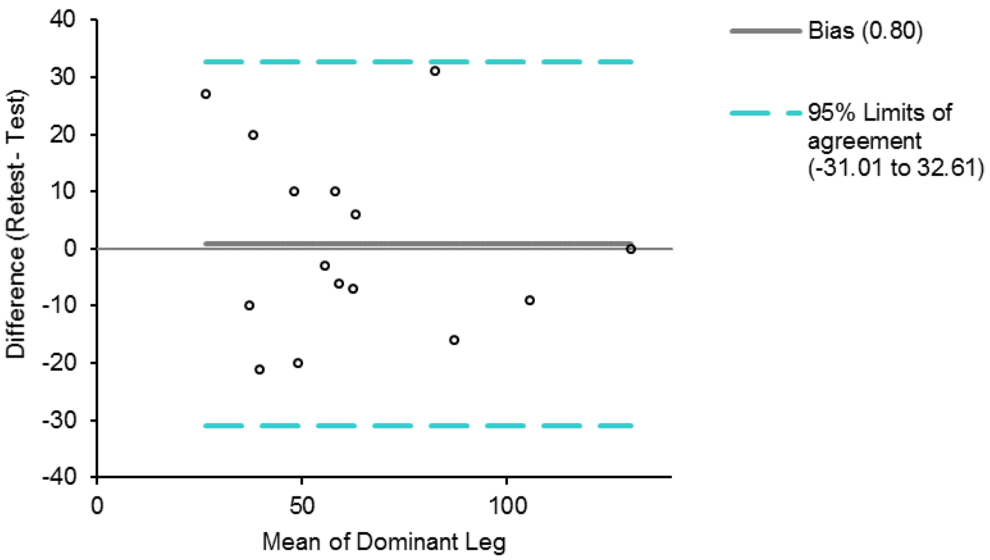

Figure 2. *Bland Altman plot of the intra-rater differences when the MSLHST is performed on the dominant leg*

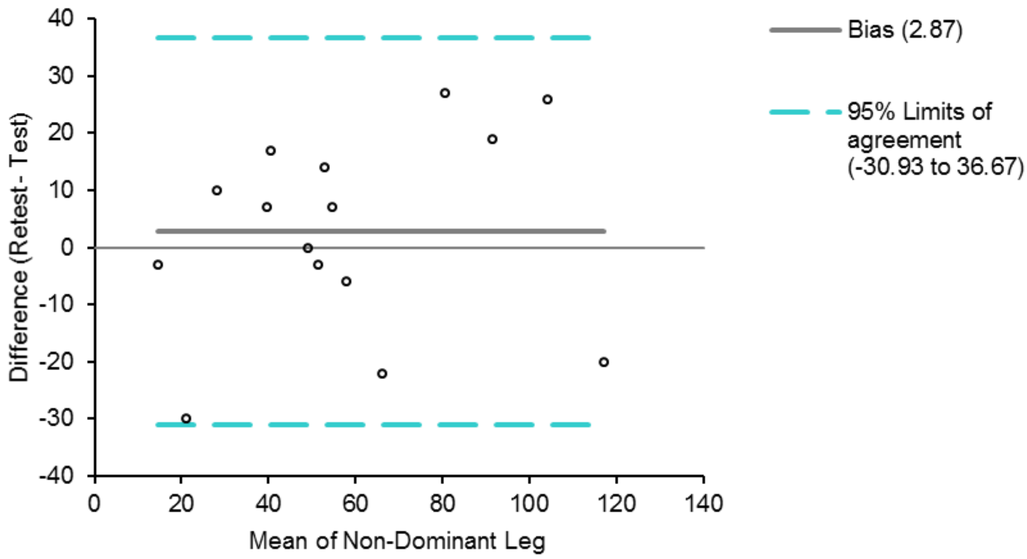

Figure 3. *Bland Altman plot of the intra-rater differences when the MSLHST is performed on the non-dominant leg*

Paired t-tests revealed no significant differences between performance of the dominant and non-dominant legs inthe first or second performance of the test (p = >0.05), therefore the scores for the dominant and non-dominant legs were averaged across the two tests (Figure 4).

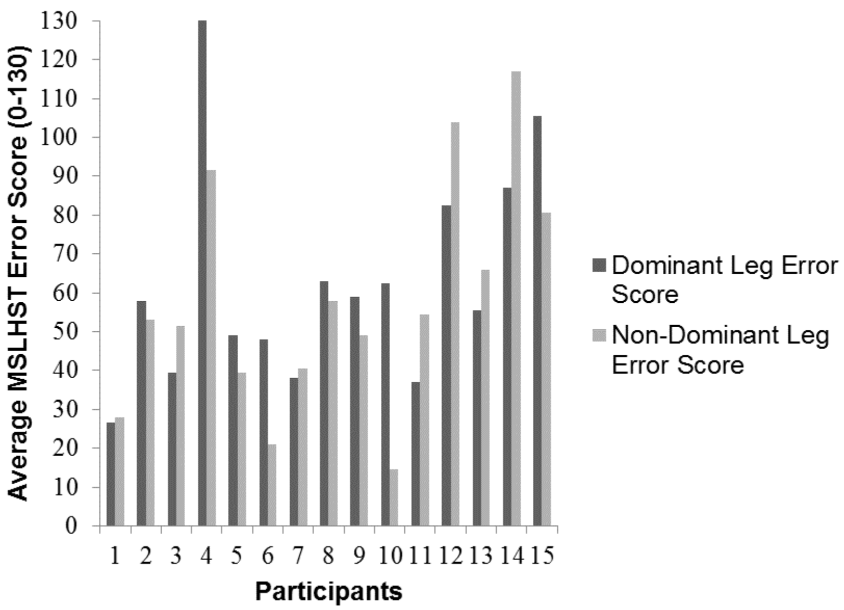

Figure 4 . *Mean error scores for the dominant and non-dominant leg*

There was a significant positive and strong relationship [29] between the scores obtained on the dominant and non-dominant legs; higher scores on one leg were associated with higher scores on the other leg (R^2=0.49 P<0.05; Figure 5).

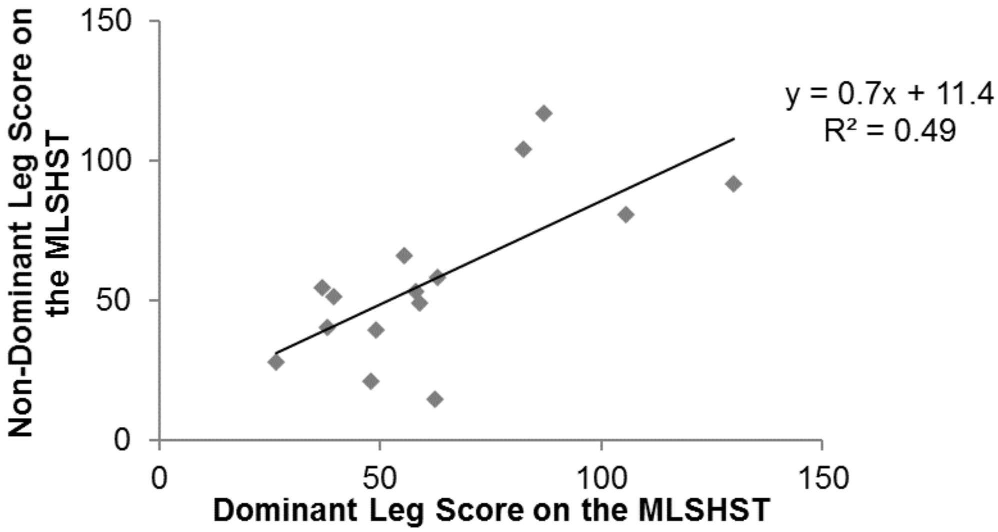

Figure 5. *A scatterplot showing the linear relationship between the average dominant and non-dominant leg scores on the multiple single-leg hop-stabilization test*

There was a significant positive and moderate relationship [29] between the scores obtained on both the dominant / non-dominant legs and the age of the participant. Higher scores (indicating more errors) were associated with advancing age The relationship was stronger on the dominant leg (non-dominant leg R^2 = 0.28, p<0.05, Figure 6; dominant leg R^2=0.39, p<0.05, Figure 7).

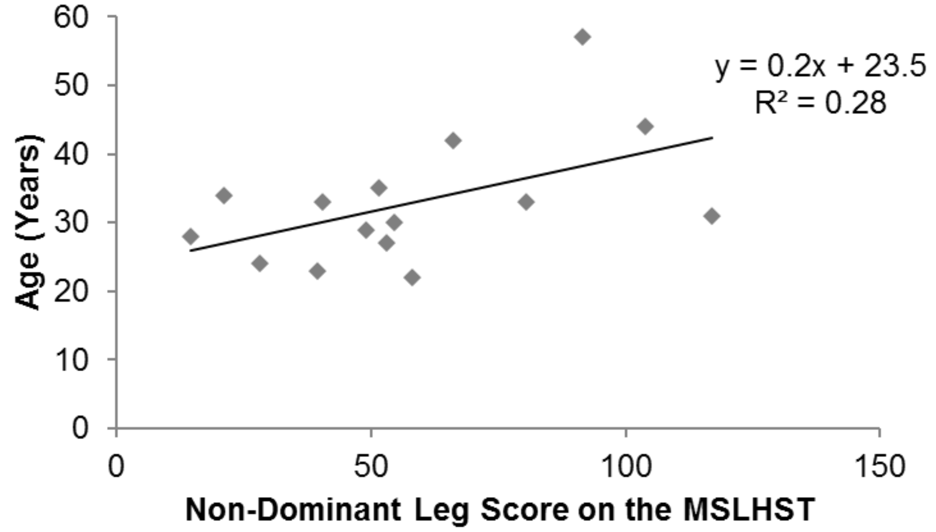

Figure 6. *A scatterplot showing the linear relationship between the average non-dominant leg scores on the multiple single-leg hop-stabilisation test and age*

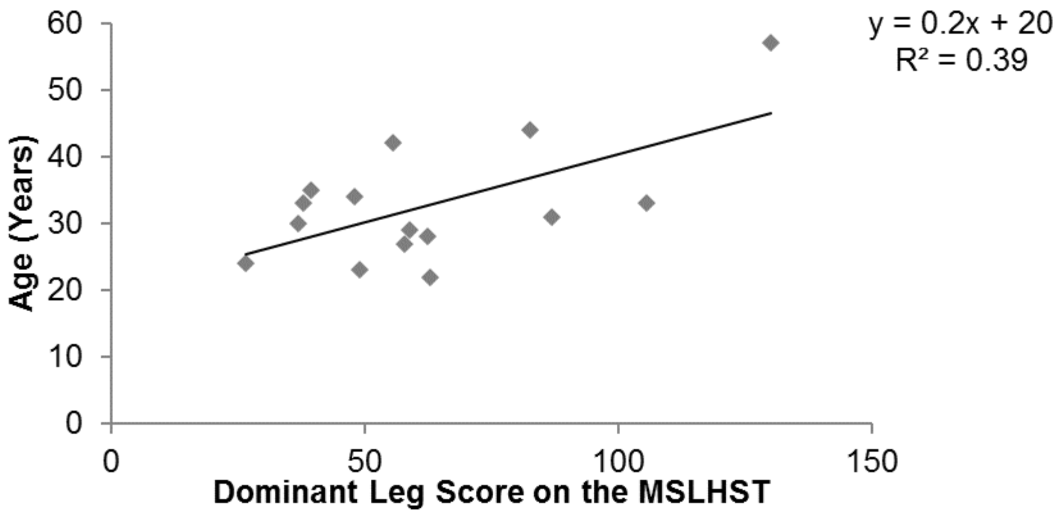

Figure 7. *A scatterplot showing the linear relationship between the average dominant leg scores on the multiple single-leg hop-stabilisation test and age*

Greater number of training hours per week were associated with lower scores on the MSLHST. This relationship, which was of moderate strength,[29] was significant for the non-dominant leg only (R^2=0.37 $p<0.05$).

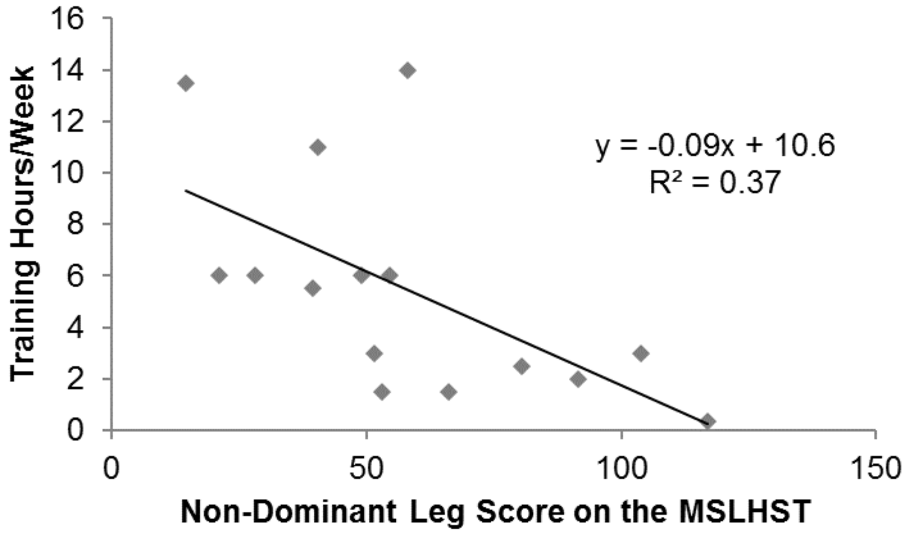

Figure 8. *A scatterplot showing the linear relationship between the average non-dominant leg scores on the and weekly multiple single-leg hop-stabilisation test training hours*

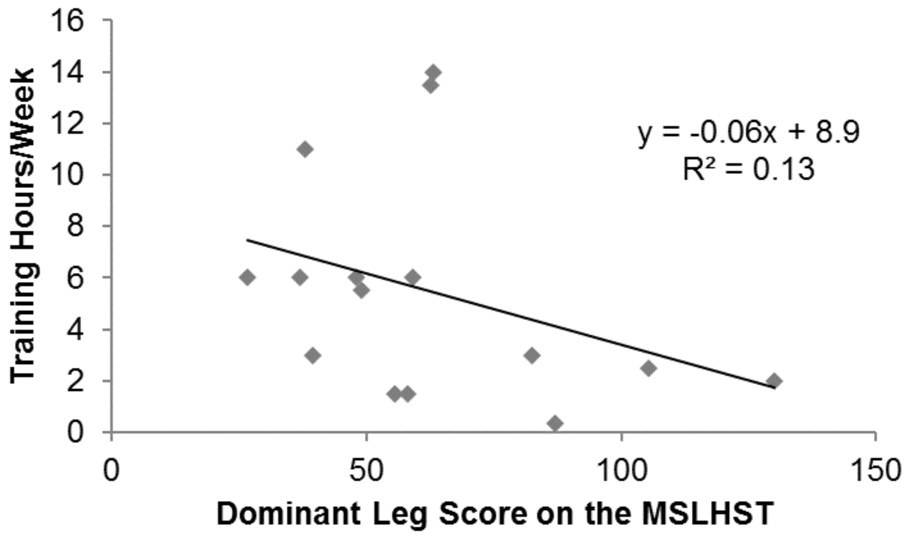

Figure 9. *A scatterplot showing the linear relationship between the average dominant leg scores on the MSLHST and weekly training hours*

Further analysis using t-tests showed a significant difference ($p = <0.05$) in overall scores between those training more and those training less than five hours per week. This was seen for both the average dominant and non-dominant leg scores.

DISCUSSION

ICC values can be interpreted as follows; 0.75 and above indicates excellent reliability, 0.4-0.75 is fair to good reliability and <0.4 is seen as poor reliability. [22] The ICC results for both the dominant and non-dominant leg both demonstrate a mean value of 0.85. Whereas this may be considered as demonstrating excellent intra-rater reliability, [22] examination of the 95% CI urges more caution. The intervals ranging from 0.62-0.95 for the dominant leg, and, 0.61-0.95 for the non-dominant leg, should be interpreted as showing that the MSLHST demonstrates good to excellent intra-rater reliability in a healthy, exercising population.

The varying degrees of reliability shown in Tables 4 and 5 allows a comparison with previous findings on the differences in the landing and balance score reliability. [18] The current findings show that ICCs range from 0.72-0.88; indicating good to excellent reliability. [22] The finding that reliability is greater with the balance scores than landing is in contrast to prior work. [18] While this may reflect the difference in the prescribed scores given for landing and balance errors, for the purpose of this work the focus upon intra-rater reliability is with the overall MSLHST score which is derived by totalling the balance and landing scores.

While ICCs were examined to provide a quantitative assessment of reliability in terms of consistency of agreement; Bland Altman plots were examined as a qualitative method of assessing reliability and determining degree of absolute agreement [30].

Inspection of these plots (Figures 2 and 3) show that the MSLHST intra-rater scores all lay within the 2 standard deviation limits. Considering these findings together with those of previous research, [18] it appears that the MSLHST could be a reliable functional outcome measure, and may be considered for inclusion in future clinical trials in a similar population.

Thorborg et al[19]suggested that one may expect to see a difference in balance ability between the dominant/ non-dominant legs. However, paired t-tests used to examine the current data demonstrated that there was no significant difference between the dominant and non-dominant limbs (p=>0.05). Furthermore a significant strong, positive correlation was observed between the MSLHST scores of the dominant and non-dominant leg. Those making less errors completing the test on their dominant leg, tend to perform similarly on ther non-dominant leg. This finding has also been observed in the sedentary population, [31] although future work is warranted to explore this in athletes.

A moderate and significant positive relationship was demonstrated between balance scores and age; higher error scores (indicative of worsening balance) occurred with increasing age when both the dominant and non-dominant legs were assessed. A deterioration of balance with age has been reported previously. [32] Changes include an increased amplitude and speed of postural sway, reduced dynamic balance and greater instability when sensory inputs controlling balance are perturbed or reduced. [33] Many of these studies compared balance ability in younger (<30 years) and older (>60 years) age groups. [32,33] It is of note that this measure of dynamic balance appeared able to detect variations in performance with age even within the relatively narrow age band of the current sample (22-57 years).

People who trained for longer periods each week had lower scores on the MSLHST (indicating better balance ability). This was only significant on the non-dominant leg. Interestingly, the task used to define the dominant leg was kicking a ball in which the opposite non-dominant leg is balancing, supporting the body weight. The moderate relationship seen between the hours spent training and better performance on the non-dominant leg balance scores might be because this leg is used more frequently for balancing activities; especially during asymmetric activities like football that involve phasic movements of the dominant leg.

Predicting performance scores through other variables can be useful in forecasting future performance outcomes. Led by the findings of earlier research [28] the number of training hours undertaken each week was explored as a predictor of subjects MSLHST scores; a significant difference ($p = <0.05$) was shown between participants when grouped in terms of the time spent engaged in exercise activities each week. More specifically the results show that it is possible to predict how well a participant will do on the MSLHST by looking at the number of hours that they spend training each week; more than five hours of training per week is a strong indicator that a participant will have a lower error score (indicative of better balance). This is supported by literature in other populations where engagement in sport and physical activities has been shown to be associated with better balance and postural control. [34]

CONCLUSION

The results of the current study demonstrate that the MSLHST demonstrates good to excellent intra-rater reliability in a healthy, active population. Furthermore simple regression analyses may suggest that predictions may be made as to participants' MSLHST error scores, based on known factors such as their age and training hours. The latter showing a significant difference (<0.05) in performance between those training more and less than five hours per week. However further work is required to confirm these findings.

In conclusion and concurring with previous work, [18] it appears that this test could be an appropriate functional measure of athletic balance to use in a future study with a young, healthy, active population.

REFERENCES

1. Peterka RJ. Sensorimotor integration in human postural control. *J Neurophysiol.* 2002;88(3):1097-1118.
2. Hahn T, Foldspang A, Vestergaard E, Ingemann-Hansen T. One-leg standing balance and sports activity. *Scand J Med Sci Spor.* 1999;9:15-18.
3. Bressel E, Yonker JC, Kras J, Heath EM. Comparison of static and dynamic balance in female collegiate soccer, basketball, and gymnastics athletes. *J Athl Train.* 2007;42(1):42-46.
4. Taube W, Gruber M, Gollhofer A. Spinal and supraspinal adaptations associated with balance training and their functional relevance. *Acta Physiologica.* 2008;193:101-116.
5. Swanenburg J, de Bruin ED, Stauffacher M, Mulder T, Uebelhart D. Effects of exercise and nutrition on postural balance and risk of falling in elderly people with decreased bone mineral density: randomized controlled trial pilot study. *Clin Rehabil.* 2007;21(6):523-534.
6. Barnett A, Smith B, Lord SR, Williams M, Baumand A. Community-based group exercise improves balance and reduces falls in at-risk older people: a randomised controlled trial. *Age Ageing.* 2003;32(4):407-414.
7. Wikstrom EA, Naik S, Lodha N, Cauraugh JH. Bilateral balance impairments after lateral ankle trauma: a systematic review and meta-analysis. *Gait Posture.* 2010;31(4):407-414.
8. Allet L, Armand S, De Bie R, et al. The gait and balance of patients with diabetes can be improved: a randomised controlled trial. *Diabetologia.* 2010;53(3):458-466.

9. Massion J. Postural control systems in developmental perspective. *Neurosci Biobehav Rev.* 1998;22(4):465-472.

10. Emery CA. Is there a clinical standing balance measurement appropriate for use in sports medicine? A review of the literature. *J Sci Med Sport.* 2003;6(4):492-504.

11. Malliou P, Gioftsidou A, Pafis G, Beneka A, Godolias G. Proprioceptive training (balance exercises) reduces lower extremity injuries in young soccer players. *J Back, Musculoskelet Rehabil.* 2004;17:101-104.

12. Mohammadi V, Alizadeh M, Gaieni A. The effects of six weeks strength exercises on static and dynamic balance of young males athletes. *Procedia Soc Behav Sci.* 2012;31:247-250.

13. Hrysomallis C, McLaughlin P, Goodman C. Relationship between static and dynamic balance tests among elite Australian Footballers. *J Sci Med Sport.* 2006;9:288-291.

14. Arumugam A, Milosavljevic S, Woodley S, Sole G. Effects of external pelvic compression on form closure, force closure, and neuromotor control of the lumbopelvic spine – a systematic review. *Manual Ther.* 2012;17(4):275-284.

15. Sawle L, Freeman J, Marsden J. The use of a dynamic elastomeric fabric orthosis (DEFO) in supporting the management of athletic pelvic and groin injury. *J Sports Rehabil.* 2016;25:101-110.

16. Risberg MA, Ekeland A. Assessment of functional tests after anterior cruciate ligament surgery. *J Orthop Sports Phys Ther.* 1994;19(4):212-217.

17. Ageberg E, Zatterstrom R, Moritz U. Stabilometry and one-leg hop test have high test-retest reliability. *Scandanavian Journal of Medicine & Science in Sports.* 1998;8(4):198-202.

18. Riemann B, Caggiano NA, Lephart SM. Examination of a clinical method of assessing postural control during a functional performance task. *J Sports Rehabil.* 1999;8:171-183.

19. Thorborg K, Couppe C, Petersen J, Magnusson SP, Holmich P. Eccentric hip adduction and abduction strength in elite soccer players and matched controls: a cross-sectional study. *Br J Sports Med.* 2011;45:10-13.

20. Hanna CM, Fulcher ML, Raina Elley C, Moyes SA. Normative values of hip strength in adult male association football players assessed by handheld dynamometry. *J Sci Med Sport.* 2010;13:299-303.

21. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess.* 1998;2(14).

22. Fleiss JL. *The Design and Analysis of Clinical Experiments.* New York City: John Wiley & Sons; 1986.

23. Rieman MP, Manske RC. *Functional Testing in Human Performance - 139 Tests for Sport, Fitness and Occupational Settings.* Champaign: Human Kinetics; 2009.

24. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clin Rehabil.* 1998;12:187-199.

25. Bruton A, Conway JH, Holgate ST. Reliability: what it is, and how is it measured? *Physiotherapy.* 2000;86(2):95-99.

26. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med.* 1998;26(4):217-238.

27. Dancey C, Reidy J. *Statistics without Maths for Psychology: Using SPSS for Windows.* 3rd ed. London: Prentice Hall; 2004.

28. Sundstrup E, Jakobsen MD, Andersen JL, et al. Muscle function and postural balance in lifelong trained male footballers compared with sedentary elderly men and youngsters. *Scand J Med Sci Spor.* 2010;20 Suppl 1:90-97.

29. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* 2nd ed. Hillsdale: Lawrence Erlbaum Associates; 1988.

30. Sampat MP, Whitman GJ, Stephens TW, et al. The reliability of measuring physical characteristics of spiculated masses on mammography. *Br J Radiol.* 2006;79:134-140.

31. Alonso AC, Brech GC, Bourquin AM, Greve JM. The influence of lower-limb dominance on postural balance. *Sao Paulo Med J.* 2011;129(6):410-413.

32. Woollacott MH, Shumway-Cook A, Nashner LM. Aging and posture control: changes in sensory organization and muscular coordination. *Int J Aging Hum Dev.* 1986;23(2):97-114.

33. Singh NB, Taylor WR, Madigan ML, Nussbaum MA. The spectral content of postural sway during quiet stance: influences of age, vision and somatosensory inputs. *J Electromyogr and Kinesiol.* 2012;22(1):131-136.

34. Perrin PP, Gauchard GC, Perrot C, Jeandel C. Effects of physical and sporting activities on balance control in elderly people. *Br J Sports Med.* 1999;33(2):121-126.