ISSN 1809-9246
Rev. bras. fisioter., São Carlos, v. 11, n. 2, p. 119-123, Mar./Apr. 2007

INTRA AND INTEROBSERVER RELIABILITY OF THE ANGULAR KINEMATIC ANALYSIS OF THE HIP DURING THE SIT-AND-REACH TEST FOR MEASURING THE LENGTH OF HAMSTRING MUSCLES IN UNIVERSITY STUDENTS

CARDOSO JR [1], AZEVEDO NCT [2], CASSANO CS [2], KAWANO MM [1] & ÂMBAR G [2]

[1]Physical Therapy Department, Health Sciences Center, State University of Londrina, Londrina, PR – Brazil

[2]Physical Therapist

Correspondence to: Jefferson Rosa Cardoso, Hospital Universitário, Av. Robert Kock, 60, CEP 86038-440, Londrina, PR - Brazil, e-mail: jeffcar@hotmail.com

Received: 15/05/2006 - Revised: 26/01/2007 - Accepted: 22/03/2007

ABSTRACT

Background: The sit-and-reach test (SRT) is used to measure the flexibility of the lumbar spine and hamstring muscles and is better when the hip joint angle (HJA) is measured concomitantly. Objective: To assess the intra and interobserver reliability of the SRT for measuring hamstring muscle length through angular kinematic analysis. Method: Fifty university students (X= 21.5 years; SD= 1.5) without musculoskeletal abnormalities took part. A standard SRT bench with a door (to assess the action of the gastrocnemius muscles) and a digital photographic camera positioned on a tripod were used. Skin markers were placed on the anterosuperior iliac spine and greater trochanter. Two images were recorded: one with the door closed (CD) and another with the door open (OD). To test the intra and interobserver reliability, the intraclass correlation coefficient (ICC) (random effect with one factor) and the Bland and Altman concordance test ($\bar{d}$ = mean difference between measurements) were used, with the respective 95% confidence intervals (95% CI). Results: The intraobserver reliability of the HJA test was: ICCdc= 0.97 (95% CI [0.95;0.99]) and Bland and Altman $\bar{d}$ = -0.02 (95% CI [-0.11;0.07]); ICCdo = 0.98 (95% CI [0.96;0.99]), $\bar{d}$ = -0.01 (95% CI [-0.11;0.08]). For the interobserver reliability: ICCdc= 0.96 (95% CI [0.94;0.98]) and $\bar{d}$ = -0.07 (95% CI [-0.19;0.03]); ICCdo= 0.96 (95% CI [0.93;0.98] and $\bar{d}$ = -0.06 (95% CI [-0.19;0.52]. Conclusion: The intra and interobserver reliability tests (CD and OD) for HJA assessment using kinematic analysis showed high reliability. The technique is easy to apply (only requiring a standard SRT bench and a photographic camera) and provides a reliable method for measuring hamstring muscles using kinematic analysis in clinical practice.

Key words: biomechanics, hip joint, physical examination, test reproducibility.

INTRODUCTION

The sit-and-reach test (SRT), proposed initially by Wells and Dillon on the 50s, is commonly used for measuring the flexibility of the lumbar spine and hamstring muscles[1]. The interest in researching the hamstring muscles, and the measurement of its length is due to the possible dysfunctions and lesions caused by alteration in its flexibility. Not only are there injuries caused by the muscle itself, there are also biomechanical alterations that may lead to femoro-patellar dysfunction, pubic and lumbar pain, tendonitis and postural deviations[2-8].

It is believed that some factors may alter the results of the SRT: differences of proportions between the length of superior and inferior limbs, vertebral spine mobility and scapular abduction[1,9-12]. Different SRTs and their modifications were studied with the objective of eliminating theses factors[12-19]. They all take in account finger reaching of the most distant point, and not the angle formed by the hip joint angle (HJA), which may, supposedly, predict the hamstring muscles length. This length may be measured by sacrum and pelvic inclination angle in relation to the ground. This represents an indirect measuring of the HJA[9]. The same test may be used for measuring the flexibility of lumbar spine, but some studies demonstrate low correlation for both, which suggests ceasing of its realization with this objective[11,15].

For measuring HJA, kinemetry is used. This technique utilizes a set of methods that seeks to measure kinematic parameters of the movement; that is, position, orientation, velocity and acceleration. One of the possible systems to obtain kinematic measures is based on photographic or video cameras, that register the movement image and, then, using specific programs, variables of interest are calculated[20]. By kinemetry, it is possible to obtain photographs that provide

a record and, subsequently, may be compared to other pictures to facilitate interpretation of the findings related to the intervention process[21]. Therefore, the HJA analysis, by use of kinemetry, allows the evaluator to verify the hamstrings muscles length and, thus, to prevent and to treat the dysfunctions that muscle shortening may cause.

In clinical practice it is indispensable the use of reliable methods, since the utilization of visual and observational assessment methods are subjective and may compromise results of the programs that involve intervention. All instrument or method of assessment must be evaluated in terms of its reliability, that is, the capacity of the instrument in assessing, at different moments, the same measure, whether involving or not the examiner[22]. Thus, a reliable instrument assesses phenomena in a trustful, sequenced, accurate and predictable manner, without alterations. If a functional assessment, or any test is not reliable, the initial state of the patient, or the effect of the treatment may be modified. Reliability studies at the rehabilitation area are necessary to ensure that the error involved on the measurement may be reduced and to detect changes on what is being measured[23].

Some studies about SRT and its variations (modified assessments) have used the reliability test. These have demonstrated high reliability results, although they all have used only the intra-class correlation coefficient for statistical analysis[14,15,17]. Because there are problems such as the influence of the magnitude of inter-subjects variability, non-presentation of the obtained value, measuring error, and non-possibility of being clinically interpreted, this test should be used together with the Bland and Altman test[24], in which data are interpreted and visualized in objective manner together with the variation differences (average differences of the measurements and its reliability intervals of 95%)[23]. In this sense, the present study had as aim to evaluate intra- and inter-observers reliability of measurement of the hips joint angle during SRT by angular kinematic analysis.

METHODS

The convenience sample was composed of 50 individuals (18 men and 32 women), sedentary or not, all university students. As inclusion criteria, all should be healthy and aged between 19 and 25 years. As exclusion criteria, no volunteer should present pain, evident postural alteration and musculoskeletal disorders on the spine or upper and lower limbs. The project and the Free and Clarified Consent Term were submitted to the Ethics Committee in Research with Human Beings of the State University of Londrina (CEP 058/04) and approved by it.

For assessment, it was used the standard Box of the sit-and-reach test (Wells Box)[1], that consists in a wood box measuring 30.5 cm x 30.5 cm x 30.5 cm, with an extension of 23 cm for the support of the upper limbs of the subjects. Over the superior aspect of the box and the extension, there is a metric scale of 50 cm that allows determining how far the individual can reach. On the standard Box, during the test, participants' feet soles are at the 23º (twenty-third) cm of the metric tape. At the local for feet support, on the box, it was constructed an opening (door) of 27 cm high x 27 cm large. This door, when open, allows elimination of the influence of gastrocnemius muscle during the test, a fact that could underestimate the maximum reach of the volunteers when these muscles are found shortened. The anatomical references used, major throcanter of the femur, and anterior-superior iliac spine (ASIS), were located and demarcated with opaque skin markers to guide measuring on the images.

A digital photographic camera, with 5.1 megapixels resolution was placed and calibrated over a 72 cm high fixed tripod, distanced 202 cm from the background. The participants were photographed during the realization of the test under two conditions: (a) with the door closed (CD) and (b) with the door open (OD). A sequence of image recording regarding the type of test was randomized. The photographs were registered at the same local and time of day.

The registered tests counted with the placement of the subjects on the Box with adequately positioned markers over the surface. After this procedure, with flexed hips, knees extended and feet soles touching or not the box, at the mark of the 23º centimeter of the scale, it was asked the participant to take a deep breath, and during exhaling, to flex his trunk forward with the intention to reach the maximum distance possible. For a control over any possible lateral dislocation of the individuals during the test, the superior limbs were aligned with the metric tape of the Box, with a hand over the other (it was standardized the middle finger of the right hand over the middle one of the left hand). The photograph was recorded at this position with both the door open and closed, according to the randomization result. Maximum reach of the middle fingers were determined as SRT score, and the HJA was calculated using the software *Suite CorelDRAW*®, version 11. Hamstring muscles lengths were measured by analysis of the inclination angle of sacrum and pelvic bones in relation to the horizontal plane at the final position of SRT[9]. For that to be done, two straight lines that formed an angle were obtained: one between the ASIS markers and the major throcanter of the femur, and the other a horizontal line that passes through the major throcanter (Figures 1 and 2).

The two evaluators appraised the same 100 photos (two by participant – one with the CD and another with OD), simultaneously, on two independent computers (inter-observer). For the intra-observer assessment, the same pictures were evaluated with an interval of five days.

The variables (anthropometric and the HJA value – with closed and open door) were tested regarding normality distribution and are presented in a descriptive way (X= mean and SD= standard deviation) because they satisfy the necessary assumption. For the intra- and inter-observers reliability test were utilized: Intra-class Correlation Coefficient

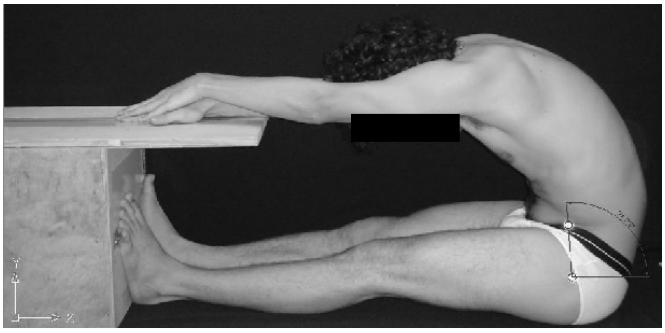

Figure 1. The sit-and-reach test with closed door.

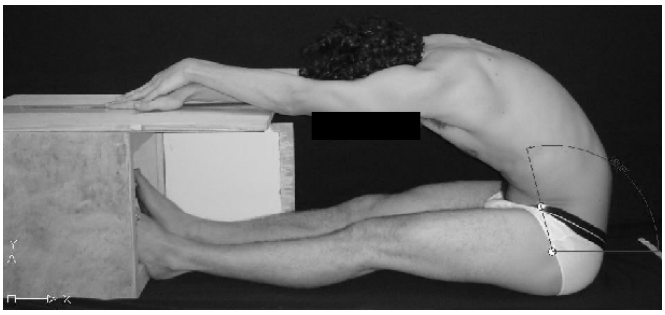

Figure 2. The sit-and-reach test with open door.

– ICC (random effect – one factor) and the agreement test of Bland and Altman[24]. An ICC of 1 indicates high reliability (i.e. errors are considered to be absent), while ICC = 0 indicates no reproducibility. The 95% confidence interval for each ICC value was calculated[21]. The Bland and Altman concordance test was also utilized[24]: the mean difference between the measures ($\bar{d}$) and its respective reliability intervals of 95% (IC 95% for $\bar{d}$), the standard deviation of the difference of the average (SD of $\bar{d}$) and the agreements limits. The closest to zero is the distribution of the values of the Bland and Altman test, the higher is the agreement[23]. The statistical programs SPSS® (version 11.5) and MedCalc® (version 8.1.0.0) were used in the study.

RESULTS

Anthropometric data from the sample (n=50) were: age between 19 and 25 years (X = 21.5 years; SD = 1.5), corporal mass between 41.5 and 92 kg (X = 60.8 kg; SD = 11.3), height between 1.48 and 1.99 (X = 1.67 m; SD = 0.10) and body mass index (CMI) between 13.1 and 26.9 (X = 21.3; SD = 2.9). Regarding physical activities, 18 (36%) participants practiced it regularly (at least three times a week). Among these, 11 (61.1 %) are male and seven (38.9%) are female.

Table 1 present the values for angle (HJA) and for reach (SRT) with closed and open door for all the sampling.

Table 1. HJA and SRT values with closed and opened door (n= 50).

	X (SD)	Min - Max
HJA - CD	94.7º (17.2º)	60.6º-144.8º
HJA - OD	99.1º (16,7º)	70.6º-145.4º
SRT - CD	28.2 cm (10)	6 – 49.5 cm
SRT - OD	32.5 cm (9.1)	10 - 51 cm

HJA= Hip joint angle; SRT= sit-and-reach test values; CD= closed door; OD= opened door; X= mean; SD= standard deviation; Min= minimum value; Max= maximum value.

At Table 2, are presented the values of the intra- and inter-observer reliability test of the analysis of the hip joint with the door open and closed. All the analyzed values showed high reliability (ICC between 0.96 – 0.98) and (Bland and Altman[24] between -0.01 and 0.07) both intra- and inter-observers.

DISCUSSION

At the present study it was made a door on the Well[1] so that, when open, the gastrocnemius muscle could be free.

Table 2. The reliability test: ICC and *Bland and Altman* (n= 50).

	ICC (1-way random)	Bland and Altman				
	ICC - CI 95%	$\bar{d}$	SE of $\bar{d}$	95%CI for $\bar{d}$	SD diff	95% LA
Intra-rater HJA closed door	0.97 – 0.95; 0.99	-0.02	0.89	-0.11; 0.07	-0.68	-0.84; 0.80
Inter-rater HJA closed door	0.96 – 0.94; 0.98	-0.07	0.40	-0.19; 0.03	-0.88	-1.08; 0.92
Intra-rater HJA opened door	0.98 – 0.96; 0.99	-0.01	0.33	-0.11; 0.08	-0.67	-0.84; 0.81
Inter-rater HJA opened door	0.96 – 0.93; 0.98	-0.06	0.43	-0.19; 0.52	-0.91	-1.12; 0.99

$\bar{d}$ = mean difference; SE of $\bar{d}$ = standard error of the mean difference; 95%CI for $\bar{d}$ = 95% confidence interval for the mean difference; SD diff = standard deviation of the differences; 95% LA = 95% limits of agreement.

The disposition of the same on a stretching position, while making the standard test (feet fixed against the closed door – trunk-flexed at 90º), could take to a false positive of shortened hamstring, since by flexing the trunk forward, the participant may refer tension to a posterior region of the leg, interfering with the final result of the test. Keeping the ankles free, the gastrocnemius muscle will be isolated, not interfering on the result due to a possible shortening. This confirms the assertion by Kendal et al.[10] about the bi-articulated muscles in which the same should be free over an articulation in a way to obtain the best possible range of movement over other articulation.

The intra-observer reliability test presented high reliability (5 days intervals), both on the tests with closed door (ICC= 0.97) and with the open door (ICC= 0.98). The same has occurred in the inter-observer reliability test, made in different environments by two trained evaluators (ICC= 0.96% and ICC= 0.96 respectively). The obtained results have proven good applicability of the technique and ease on the calculations of the values of the HJA by the evaluators. Skin markers of easy visualization and angle measurement made in simple way contribute significantly to the kinematic technique.

In a study that has used the Chair Sit-and-Reach Test (a SRT variation), it was proposed an analysis of the inter-observer reliability. Seventy-six volunteers participated in the study (34 men and 42 women), with mean age of 70.5 years, which were evaluated at different moments, with intervals from two to five days. The result found has showed high reliability (r= 0.92 for males and r = 0.96 for females)[17]. In another inter-observer reliability study, 46 girls and 42 boys were evaluated, aged between 11 and 15 years, with Back Saver Sit-and-Reach Test (on this test, each leg is evaluated separately)[15]. The results were: ICC= 0.97 for boys (both sides) and ICC= 0.97 for the right leg , and ICC= 0.96 for the left leg among the girls, which indicate high reliability. However, the authors did not mention if there was an interval between tests[15]. Hui & Yuen's[14] study, with the Modified Back-Saver Sit and Reach Test, has also showed a high inter-observer reliability, with a sample of 96 women and 62 men aged between 17 and 41 years. The result was: ICC = 0.96 for the left leg and ICC= 0.97 for the right leg. These studies have analyzed only inter-observer reliability and used only the intraclass correlation coefficient (ICC) or Pearson's correlation coefficient (r) for these analyses. At the present study, it was assessed intra- and inter-observer reliability of the angular kinematics analyses (using photography) and not from the repetition of the test, as in the studies above.

Cornbleet et al.[9] evaluated 410 children (211 girls and 199 boys) aged between five and twelve years. From the analyzed studies, this was the only one to consider the interference of the anthropometric factors, spine mobility and scapular abduction on the SRT results. The authors have discussed the idea that the final position of the SRT may be used for measuring the HJA and that reflects a real length of the hamstring muscles. For that to be done, they have used an inclinometer placed vertically in relation to the sacrum, aligned at the height of the posterior-superior iliac spines. Mean value found for the angle was 75º (SD= 10º) for boys, and 85º (SD= 10º) for the girls. This corroborates with the results found at the present study, in which the female gender presented higher flexibility, however, angle means were inferior. For the whole sample, it was found an mean of 81º (SD= 11º), which corresponds to a normal hamstring length according to Kendall et al.[10] (80º). The authors have analyzed inter-observer reliability with the first twenty volunteers and, as well as other authors, have used only the intra-class correlation coefficient (ICC), having as result a high reliability (ICC= 0.98).

In another study[25], with a different SRT application, it was also made intra- and inter-observer reliability of the analyses of a same picture. Three different techniques of measuring anterior trunk flexion (modified *Chaffin*, *Whistance*, and the perpendicular Marker) were evaluated in 20 volunteers, with mean age of 21.5 years (SD = 2.2). In this study it was demonstrated high intra-observer reliability for the three techniques (between 0.94 and 0.95). For the inter-observer reliability the modified *Chaffin* technique presented moderate reliability (ICC = 0.54), while the other techniques presented high reliability (ICC = between 0.97 and 0.99). Moderate reliability of the modified *Chaffin* technique is justified by the high mobility of one of the anatomical references (internal rotation of the shoulder covering the marker) and of the segments during realization of the technique.

Many studies that found high inter-observer reliability do not realize the importance of the intra-observer reliability analyses. However, it is important to register how much a technique is reliable for a same evaluator at another moment. Reliability analyses itself needs two statistical tests (ICC and Bland and Altman[24] on this study, for example), because only the ICC does not provide enough information about the measures reliability due to: influence of the magnitude of the between-individuals variation; non demonstration of presentation of the measured value or its variation; measurement error and impossibility of being clinically interpreted. On the other hand, the Bland and Altman[24] test provide a graphic in which the size and amplitude of the measurement differences may be easily interpreted, identifying mistakes or outliers. Furthermore, this method presents the reliability interval values for the difference of the mean and the agreement limits. These data, respectively, will indicate measuring error and may relate to the clinical acceptability[23].

The limitation found on the test was the skin marker movement, from its demarcation until final position of the SRT. Some participants, while moving during the test, allow that the marker migrate from the exact local of the anatomical reference, which obliges the making of a new test with the exact remarking. Another limitation consists on the anatomical differences of the participants, such as ASIS height and the

depth of the major throcanter, besides the anthropometric factors. This demands from the evaluators a good palpatory anatomy knowledge. Another difficulty for the making of the photographic record was the necessity of the volunteer to stay with the two anatomic reference points at a visible location, being, therefore, necessary the use of clothes with two pieces for women, for example.

The SRT with measurement of the HJA, by means of kinemetry, may be done at any place where it can be set a small evaluation studio. For that, it should be used a standard Wells[1] Box and a photographic camera for picture recording. This technique provides a reliable and simple hamstring length measurement when the evaluator uses the final SRT position (maximum trunk flexion) with the ankles free and registers the HJA.

CONCLUSION

The high intra- and inter-observer reliability found at this study, was due to the easy application of the technique and results analyses. Thus, the HJA assessment, by means of the angular kinematics analyses, demonstrate to be a reliable and reproducible modality of assessing the hamstring muscles, which permits, for the clinical practice, prevention and the establishment of an early and adequate treatment.

Acknowledgments: The Physical Therapist Ana Paula R. G. Cardoso, for the contribution on the elaboration of this project; to Professor Denise da Vinha Ricieri, for the technical support and to the students Beatriz I .R. de Oliveira, Marcela C. Bôer and Maria A. B. Pessan, for the help recruiting voluntaries and organization of the measuring place.

REFERENCES

1. Wells KF, Dillon EK. The sit and reach: a test of back and leg flexibility. Res Q Exerc Sport. 1952;23:115-8.
2. Lee D, Drez D, Miller MD. Orthopaedic sports medicine: principles and practice. 2ª ed. Philadelphia: Saunders; 2002.
3. Witvrouw E, Danneels L, Asselman P, D´Have T, Cambier D. Muscle flexibility as a risk factor for developing muscle injuries in male professional soccer players. Am J Sports Med. 2003; 31:41-5.
4. Witvrouw E, Bellemans J, Lysens R, Danneels L, Cambier D. Intrinsic risk factors for the development of patellar tendonitis in an athletic population – a two – year prospective study. Am J Sports Med. 2001;29:190-5.
5. Hartig DE, Henderson JM. Increasing hamstring flexibility decreases lower extremity overuse injuries in military basic trainees. Am J Sports Med. 1999;27:173-6.
6. Woods C, Hawkins RD, Maltby S, Hulse M, Thomas A, Hodson A. The football association medical research programme: an audit of injuries in professional football – analysis of hamstring injuries. Br J Sports Med. 2004;38:36-41.
7. Busquet L. As cadeias musculares. 1ª ed. Piracicaba: Skin Direct Store; 2001.
8. Barlow A, Clarke R, Johnson N, Seabourne B, Thomas D, Gal J. Effect of massage of the hamstring muscle group on performance of the sit-and-reach test. Br J Sports Med. 2004;38: 349-51.
9. Cornbleet S, Woolsey N. Assessment of hamstring muscle length in school-aged children using the sit-and-reach test and the inclinometer measure of hip joint angle. Phys Ther. 1996; 76:850-5.
10. Kendall FP, McCreary E, Provance PG. Músculos: provas e funções. 4ª ed. São Paulo: Manole; 1995.
11. Jackson A, Baker A. The relationship of the sit-and-reach test to criterion measures of hamstring and back flexibility in young females. Res Q Exerc Sport. 1986;57:183-6.
12. Hoeger WWK, Hopkins DR. A comparison of the sit-and-reach and the modified sit and reach in the measurement of flexibility in women. Res Q Exerc Sport. 1992;63:191-5.
13. Lemmink KAPM, Kemper HCG, de Greef MHG, Rispens P, Stevens M. The validity of the sit-and-reach test and the modified sit-and-reach test in middle-aged to older men and women. Res Q Exerc Sport. 2003;74:331-7.
14. Hui SSC, Yuen PY. Validity of the modified back-saver sit-and-reach test: a comparison with other protocols. Med Sci Sports Exerc. 2000;32:1655-9.
15. Patterson P, Wilksten DL, Ray L, Flanders C, Sanphy D. The validity and reliability of the back-saver sit-and-reach test in middle school girls and boys. Res Q Exerc Sport. 1996;67: 448-51.
16. Minkler S, Patterson P. The validity of the modified sit-and-reach test in college-age students. Res Q Exerc Sport. 1994;65: 189-92.
17. Jones CJ, Rikli RE, Max J, Noffal G. The reliability and validity of a chair sit-and-reach test as a measure of hamstring flexibility in older adults. Res Q Exerc Sport. 1998;69:338-43.
18. Hui SC, Yuen PY, Morrow JR, Jackson AW. Comparison of the criterion-related validity of sit-and-reach tests with and without limb length adjustment in Asian adults. Res Q Exerc Sport. 1999;70:401-8.
19. Baltaci G, Un N, Tunay V, Besler A, Gerçeker S. Comparison of three different sit-and-reach test for measurement of hamstring flexibility in female university students. Br J Sports Med. 2003;37:59-61.
20. Amadio AC, Costa PHL, Sacco ICN, Serrão JC, Araújo RC, Mochiziki L, et al. Introdução à análise do movimento humano – descrição e aplicação dos métodos biomecânicos de medição. Rev Bras Fisioter. 1999;3:41-54.
21. Chen YL. Accuracy and repeatability of the stick marker technique for external measurement of the sacral angle during trunk flexion. Int J Ind Ergon. 2000;26:101-7
22. Rosner B. Fundamentals of biostatistics. 5ª ed. Duxbury: Thonsom Learning; 2000.
23. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. Clin Rehabil. 1998;12:187-99.
24. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurements. Lancet. 1986;1:307-10.
25. Sato TO, Vieira ER, Gil Coury HJC. Análise da confiabilidade de técnicas fotométricas para medir a flexão anterior de tronco. Rev Bras Fisioter. 2003;7;53-9.