ARTIGO ORIGINAL | ORIGINAL ARTICLE

Concordância Intra- e Inter-Observador entre Diferentes Métodos de Medição da Circunferência da Coxa

Intra and Interrater Agreement Between Different Measurement Methods of Thigh Circumference

Rui Cadilha[(1)] | Hugo Amorim[(1)] | José Santoalha[(1)] | Afonso Rocha[(1)] | Fernando Parada[(1)]

Resumo

Introdução: Avaliar a reprodutibilidade intra e inter-observador na medição da circunferência da coxa utilizando diferentes referências anatómicas.

Material e Métodos: Vinte e cinco voluntários sem antecedentes de patologia ou cirurgia no membro inferior dominante entraram no estudo. As medições foram realizadas por dois avaliadores independentes, em duas ocasiões, com um intervalo de uma semana. A ordem das medições e dos participantes foram randomizadas. Os resultados das medidas intercalares foram ocultados, sendo analisados por um terceiro investigador. O protocolo de avaliação foi definido previamente. A correlação intra e inter-examinadores foi determinada pelo *intraclass correlation coefficient* (ICC). Os limites de concordância foram estabelecidos de acordo com o método de Bland *and* Altman.

Resultados: O coeficiente de correlação na reprodutibilidade intra-observador foi elevada (SPP_{ICC} 0,96, KJL_{ICC} 0,95, $ASIS_{ICC}$ 0,96). Nos resultados inter-observador os limites de concordância foram: SPP_{ICC} 0,91 (IC 95%: 0,79–0,96), KJL_{ICC} 0,94 (IC 95%: 0,86–0,97), $ASIS_{ICC}$ 0,90 (IC 95%: 0,77–0,95).

Conclusão: Todos os métodos apresentaram alta reprodutibilidade intra e inter-observador, que pela simplicidade do método de medição poderá favorecer a escolha do pólo superior da rótula na ausência de patologia no segmento anatómico avaliado.

Palavras-chave: Antropometria; Coxa.

Abstract

Introduction: *Evaluate the intra- and interrater reliability in the thigh circumference measurement using different anatomical references.*

Material and Methods: *Twenty five volunteers without history of pathology or surgery in the dominant leg entered in the study. The measurements were performed by two independent evaluators, on two occasions with an interval of one week. The measurements and participants order were randomized. The results of the interim measures were concealed, being analyzed by a third investigator. The assessment protocol was previously defined. The intra- and inter-rater correlation was measured through the intraclass correlation coefficient (ICC). The limits of agreement were established in accordance with the method of Bland and Altman.*

Results: *The intraclass agreement in intrarater reproducibility was high (SPP_{ICC} 0.96, KJL_{ICC} 0.95, $ASIS_{ICC}$ 0.96). In the interrater results the limits of agreement were: SPP_{ICC} 0.91 (IC 95%: 0.79–0.96), KJL_{ICC} 0.94 (IC 95%: 0.86–0.97), $ASIS_{ICC}$ 0,90 (IC 95%: 0.77–0.95).*

Conclusion: *All methods presented high intra- and interrater reliability, which by the simplicity of the measurement method may favor the choice for upper pole of the patella in the absence of pathology in anatomical segment evaluated.*

Keywords: Anthropometry; Thigh.

(1) Serviço de Medicina Física e de Reabilitação, Hospital de S. João, Porto, Portugal
Autor correspondente: Rui Cadilha. rpcadilha@hotmail.com. Alameda Prof. Hernâni Monteiro, 4200-319 Porto
Data de submissão: setembro de 2016
Data de aceitação: janeiro de 2017

Introdution

Thigh muscle mass estimation is commonly used as a surrogate for muscle structural and functional characteristics in both clinical and non-clinical (sports) settings, amongst other common anthropometric features such as thickness of skin folds, length of limbs, body fat and body mass index.[1-3] Methods for thigh muscle mass estimation range from most accurate costly options such as magnetic resonance, computed tomography and DEXA ("*dual-energy X-ray absorptiometry*") to simpler readily available methods such as manual measurement using different anatomical reference points.[4,5]

Thigh circumference measurement is an essential component of musculoskeletal examination both to assist diagnosis and to monitor success of rehabilitation treatments,[6] although correlation to both structural (muscle cross sectional area) and functional (muscle power and strength) could be surprisingly low.[5,7-9] The most common reference points used are the medial knee joint line, the anterior tibial tuberosity or the superior pole of the patella.[7,10] which are used interchangeably according to the examiner's preference and experience. Like any other measurement method, the intra- and interrater reliability are crucial to allow comparisons overtime and between different observers.

The authors sought to assess and compare the intra- and interrater reliability in thigh circumference measurements, using different anatomical references according to three methods: 10 cm above the upper pole of the patella (SPP) *versus* 15 cm above the medial knee joint line (KJL) vs junction point of the distal third to the proximal two thirds of the distance between the anterior superior iliac spine and anterior tibial tuberosity (ASIS).

Material and Methods

Participants: A consecutive sample of 25 healthy volunteers were recruited amongst patients and coworkers not directly involved in the study. A structured questionnaire regarding socio-demographic and clinical data, including musculoskeletal complaints involving the lower limbs, was applied by a trained interviewer. Eligibility criteria for the study included: age ≥ 18 years, adequate cognitive skills to answer the questionnaire, and, no history of recent pain (last 6 months), major trauma, fracture or surgery in the dominant leg.

Procedures: measurements were performed on two different occasions, one week apart, by two independent examiners blinded to interim measurements. Both order of participants and measurements were randomly assigned according to a computer generated sequence. The average of three measurements was recorded for each location. Each examiner received a three-hour training period according to a predefined protocol to standardize procedures among evaluators. Measurements were made on the dominant leg using non-extensible plastic tape marked to the millimeter, applied in full contact with the skin, uniform tension and perpendicular to the long axis of the member. Subjects laid supine with the knee relaxed in full extension, upper and lower limbs in contact with the examination table and all reported this position as comfortable. Three reference surface landmarks for thigh girth measurement were identified and marked using a water-soluble pen and removed between each assessment (Fig.1):

- 10 cm above the upper pole of the patella (SPP) identified through manual palpation;
- 15 cm above medial knee joint line (KJL): the medial knee joint line was identified with the hip flexed to 45° and externally rotated and knee in 90° flexion;
- transition point between the lower and the upper two-thirds of the distance between the anterior superior iliac spine and the anterior tibial tuberosity (ASIS), identified through manual palpation.

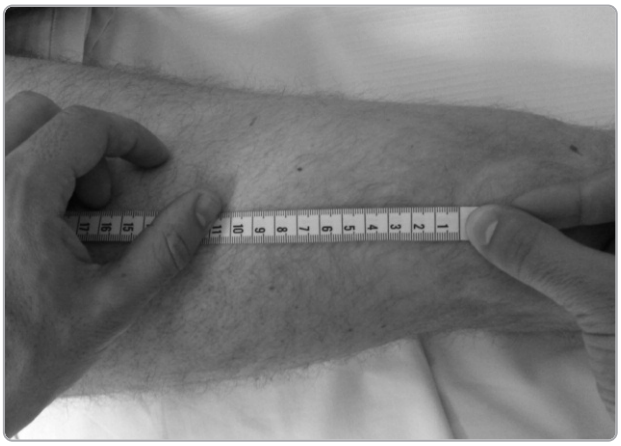

Figure 1 A - 10 cm above the upper pole of the patella (SPP)

Figure 1 B1 - Determination of medial knee joint line (KJL)

Figure 1 B2 - 15 cm above medial knee joint line (KJL)

Figure 1 C - Union point of the distal third to the proximal two thirds of the distance between anterior superior iliac spine and anterior tibial tuberosity (ASIS)

Statistical analysis was made using *Statistical Package for the Social Sciences* (SPSSv23®). According to standard recommendations for intra and interrater reliability studies we used the intraclass correlation *coefficient* (ICC) and the Bland-Altman method.[11,12]

Since our aim was to use the study information for general application in clinical practice the ICC equation was chosen.[12] ICC values range from 0 (no agreement) to 1 (total agreement) and values above 0.90 suggest excellent agreement, 0.75-0.89 moderate agreement and below 0.75 poor agreement.[13] We used the Bland Altman method for plotting the difference against the mean of two measurements to allow visual judgment of any systematic error. Limits of agreement were calculated and presented as mean 1.96 +/- x standard deviation of the difference between measurements.[11] According to the method suggested by Zou GY,[14] using a previous pilot study to determine the ICC 95% confidence intervals, we estimated that 25 participants would be required to ensure that the half width of a 95% two-sided confidence interval for the ICC would be no greater than 0.15 with an 80% assurance probability.

Results

Study sample subjects were mostly male and showed a wide distribution regarding age and anthropometric characteristics (Table 1). We found a good to excellent intrarater agreement in all 3 measurements methods (Table 2) in both raters (ICC ≥ 90). ICC values within each rater were remarkably similar, independently of anatomical location of measurement. On the other hand, the interrater reliability was also good to excellent, with ICC ranging from 0.90 to 0.94 for ASIS and KJL, respectively (Table 3). However, absolute differences between raters were noticeable with 95% of values falling within -3.37 cm and +2.59 cm, -1.99 cm and 2.08 cm for ASIS and KJL, respectively.

Visual inspection of the Bland-Altman plots showed differences did not vary in any systematic way over the range of measurement, hence there no systematic bias was seen even for the most extreme observations (Fig. 2).

Table 1: Demographic data of 25 subjects.

		Range (min.-max.)
Gender (M/F), n	19/6	
Age (years), average (SD)	48.3 (17.27)	26 - 70
Height (cm), average (SD)	168.59 (9.10)	155 -183
Weight (kg), average (SD)	74.17 (13.15)	51 - 96
BMI (kg/m2), average (SD)	26.02 (3.63)	20.69 - 34.01

SD - standard deviation; BMI - body mass index.

Table 2: Intrarater correlation coefficients for the different locations

	Evaluator 1 (T1)	Evaluator 1 (T2)	Evaluator 2 (T1)	Evaluator 2 (T2)
SPP				
Average (SD)	44.62(3.23)	44.79(3.29)	44.84(3.01)	44.55(2.97)
ICC, (95%CI)	**0.96 (0.92-0.98)**		**0.92 (0.82-0.96)**	
KJL				
Average (SD)	44.64(2.98)	44.87(3.08)	44.60(2.85)	44.40(2.85)
ICC, (95%CI)	**0.95 (0.90-0.98)**		**0.91 (0.80-0.96)**	
ASIS				
Average (SD)	44.24(3.49)	44.09(3.34)	44.64(3.20)	44.26(3.45)
ICC, (95%CI)	**0.96 (0.92-0.98)**		**0.90 (0.76-0.95)**	

ICC- intraclass correlation coefficient, SPP - 10 cm above the upper pole of the patella; KJL- 15 cm above medial knee joint line; ASIS - union point of the 1/3 distal to the proximal two thirds of the distance between anterior superior iliac spine and anterior tibial tuberosity; T1 - first assessment; T2 - second assessment.

Table 3: Interrater correlation coefficients and limits of agreement for the different locations

	ICC	*ICC-95%CI*	*Mean diff (d)*	*SDdiff (SD)*	*LOA _(d+/-1.96*SD*
SPP	*0.91*	*0.79-0.96*	*-0.23*	*1.35*	*-2.87; 2.42*
KJL	*0.94*	*0.86-0.97*	*0.04*	*1.04*	*-1.99; 2.08*
ASIS	*0.90*	*0.77-0.95*	*-0.39*	*1.52*	*-3.37; 2.59*

ICC - intraclass correlation coefficient; LOA - limits of agreement; SPP - 10 cm above the upper pole of the patella; KJL- 15 cm above medial knee joint line; ASIS - union point of the 1/3 distal to the proximal two thirds of the distance between anterior superior iliac spine and anterior tibial tuberosity.

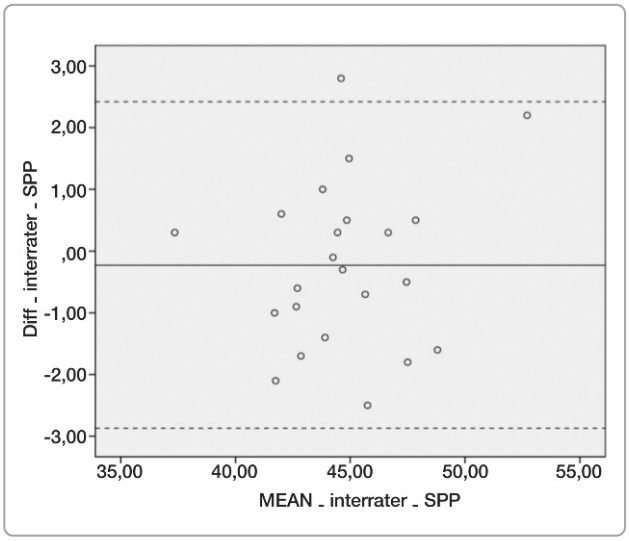

Figure 2 A - Analysis with Bland and Altman method (average measurements versus the average of the differences). A- superior pole of patella (SPP)

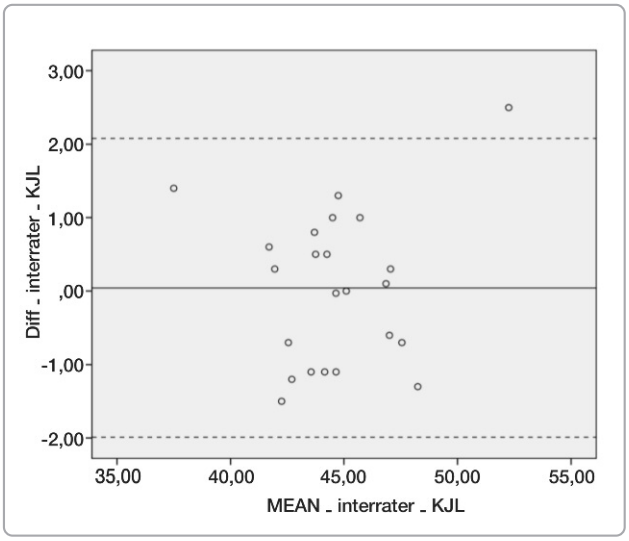

Figure 2 B - Analysis with Bland and Altman method (average measurements versus the average of the differences). B- medial knee joint line (KJL)

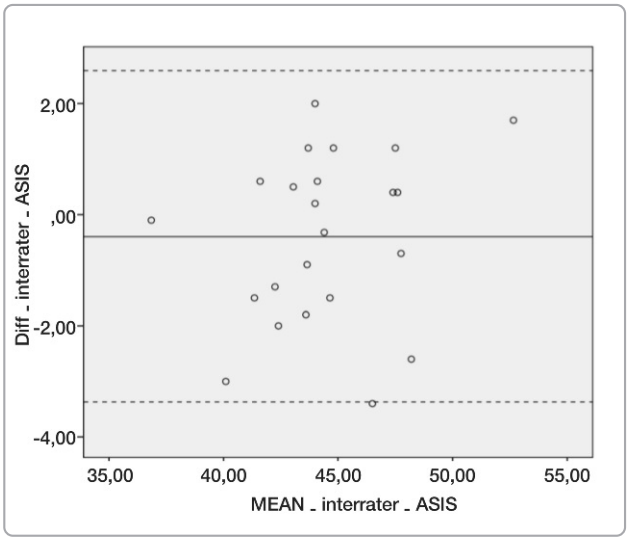

Figure 2 C - Analysis with Bland and Altman method (average measurements versus the average of the differences). C- anterior superior iliac spine (ASIS)

Discussion

The authors founded a high intra and interrater agreement for measuring thigh circumference, irrespective of measurement method, although absolute differences are considerable and must be taken into account for clinical and investigational purposes.

High agreements do not necessarily reflect clinically acceptable absolute differences between measurements. The study from Maylia *et al,*[15] using several raters and comparing thigh measurements 10 cm above the top of the patella, found average differences between measurements of 4.0 cm and 3.5 cm for inter and intrarrater comparisons, respectively. The author s study showed mean differences between raters of 0.04 to 0.39 cm with 95% measurements lying within 3.37 cm range of variation. This higher agreement might be explained by lower variation due to a more restricted number of raters and standardization of measurement protocols and intensive training in measurement techniques which might have reduced measurement bias.[10]

Limitations of manual limb circumference measurements interpretation could be patient-dependent (weight and body fat changes, hydration state, collaboration) and examiner-dependent factors (measuring technique and protocol, clinical experience). This study has limitations to be considered when interpreting study results: study sample subjects were mostly male (76%), eligibility criteria restricted the study sample to healthy volunteers affecting external validity and extrapolation of results for other populations, including clinical use in patients. Large confidence intervals for the ICC possibly indicate suboptimal sample size for the purpose intended. On the other hand, randomization of patients and evaluation sequence, concealment of raters for measurement results, the standardization of measurement procedures and the heterogeneity of participants concerning age and gender are methodological strengths which assure sufficient internal validity of study results.

Conclusion

All methods presented high intra- and interrater reproducibility, whereby the simplicity of the measurement method may favor the choice for SPP in the absence of pathology in anatomical segment evaluated. Further research is needed to assess agreement of different manual measurement methods with structural and functional status of muscle, namely through imaging techniques.

Conflitos de interesse: Os autores declaram não possuir conflitos de interesse. Suporte financeiro: O presente trabalho não foi suportado por nenhum subsídio ou bolsa. Confidencialidade dos dados: Os autores declaram ter seguido os protocolos do seu centro de trabalho acerca da publicação dos dados de doentes. Protecção de pessoas e animais: Os autores declaram que os procedimentos seguidos estavam de acordo com os regulamentos estabelecidos pelos responsáveis da Comissão de Investigação Clínica e Ética e de acordo com a Declaração de Helsínquia da Associação Médica Mundial.

Conflicts of interest: The authors have no conflicts of interest to declare. Financing Support: This work has not received any contribution, grant or scholarship. Confidentiality of data: The authors declare that they have followed the protocols of their work center on the publication of data from patients. Protection of human and animal subjects: The authors declare that the procedures followed were in accordance with the regulations of the relevant clinical research ethics committee and with those of the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Referências / *References*:

1. Brunkhorst L, Kielstein H. Comparison of anthropometric characteristics between professional triathletes and cyclists. Biol Sport. 2013;30:269-73.
2. Knechtle B. Relationship of anthropometric and training characteristics with race performance in endurance and ultra-endurance athletes. Asian J Sports Med. 2014;5:73-90.
3. Friedrich M, Rust CA, Rosemann T, Knechtle P, Barandun U, Lepers R, et al. A Comparison of Anthropometric and Training Characteristics between Female and Male Half-Marathoners and the Relationship to Race Time. Asian J Sports Med. 2014;5:10-20.
4. Al-Gindan YY, Hankey CR, Leslie W, Govan L, Lean ME. Predicting muscle mass from anthropometry using magnetic resonance imaging as reference: a systematic review. Nutr Rev. 2014;72:113-126.
5. Mathur S, Takai KP, Macintyre DL, Reid D. Estimation of thigh muscle mass with magnetic resonance imaging in older adults and people with chronic obstructive pulmonary disease. Phys Ther. 2008;88:219-30.
6. Ross M, Worrell TW. Thigh and calf girth following knee injury and surgery. J Orthop Sports Phys Ther. 1998;27:9-15.
7. Laupattarakasem W, Paholpak P, Kosuwon W. The relevant level to estimate girth difference between thighs after anterior cruciate ligament deficiency. J Med Assoc Thai. 2012;95 Suppl 10:S178-83.
8. Allison SC, Westphal KA, Finstuen K. Knee extension and flexion torque as a function of thigh asymmetry. J Orthop Sports Phys Ther. 1993;18:661-6.
9. Nicolakis P, Nicolakis M, Dorotka R, Ebenbichler G, Uher E. Beurteilung des Rehabilitationsfortschritts mittels Oberschenkelumfangmessung. Z Orthop Ihre Grenzgeb. 2000;138:526-9.
10. Soderberg GL, Ballantyne BT, Kestel LL. Reliabiity of lower extremity girth measurements after anterior cruciate ligament reconstruction. Physiother Res Int. 1996;1:7-16.
11. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1:307-310.
12. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. Clin Rehabil. 1998;12:187-99.
13. Bartko JJ. Measures of agreement: a single procedure. Stat Med 1994;13:737-45.
14. Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. Stat Med. 2012;31:3972-81.
15. Maylia E FJ, Nokes LD, Jones MD. Can thigh girth be measured accurately? a preliminary investigation. J Sport Rehabil. 1999:43-9.