ORIGINAL RESEARCH

Reliability Index of inter- and intra-rater of manual goniometry and computerized biophotogrammetry to assess the range of motion of internal and external shoulder rotation

Leandro Antonietti,[I] Natalia Luna,[I] Gabriel Nogueira,[I] Ana Ito,[I] Marcelo Santos,[I] Angelica Alonso,[II] Moises Cohen[I]

[I] Sports Traumatology Center, Department of Orthopedics and Traumatology, Federal University of São Paulo, Brazil [II] Institute of Orthopedics and Traumatology, Hospital das Clínicas, Sao Paulo University Medical College

OBJECTIVES: Measurements of the joint angles of the shoulder complex are important for diagnosis, assessment and monitoring of the treatment progression of movement disorders, provided that they can be seen as valid and reliable. The object of this study was to determine inter- and intra-rater reliability of manual goniometry and computerized biophotogrammetry for the assessment of range of motion of the medial and lateral rotations of the shoulder.

METHODS: Four evaluators (two for goniometry and two for biophotogrammetry) assessed 11 males, 16 - 26 years old, right-handed and with no shoulder anomalies. A universal plastic goniometer was used for the goniometry assessment. The biophotogrammetry assessment involved the use of a digital camera Sony DSC-W1 (5.1Mp), with non-reflective markers placed on the subjects. Photographic frames were analyzed the though the SAPO software (version 0.67). Each evaluator was blinded to data from all other valuators; inter-rater datas were compared. Seven days after the first assessment, all measurements were repeated in order to complete the intra-rater comparison. The Wilcoxon test was used to check statistical significance, the Spearman correlation was calculated and inter-class correlations coefficients were determined.

RESULTS: Inter-class correlations for inter- and intra-rater goniometry results were 0.897 and 0.830 respectively; the corresponding biophotogrammetric values were 0.982 and 0.954, all representing excellent reliability levels.

CONCLUSION: Goniometry and biophotogrammetry are reliable methods for assessment of shoulder rotation; however, biophotogrammetry has been shown to be more reliable.

KEYWORDS: biophotogrammetry; goniometry; reliability; shoulder; joint angles.

Antonietti L, Luna N, Nogueira G, Ito A, Santos M, Alonso A, Cohen M. Reliability Index of inter- and intra-rater of manual goniometry and computerized biophotogrammetry to assess the range of motion of internal and external shoulder rotation. MEDICALEXPRESS. 2014;1(2):95-99.

Received for publication on January 20 2014; **First review completed on** February 11 2014; **Accepted for publication on** March 17 2014

E-mail: angelicacastilho@msn.com

INTRODUCTION

Measurement of articulation angles is important for diagnosis, assessment, and monitoring of treatment progression for motion/movement disorders; however, these measurements must be valid and reliable to enable the success of such procedures.[1]

Reliability of a measurement is the consistency between successive measurements of a variable from the same subject in identical conditions. Validity refers to how much the measurement represents the real value of the variable, assessed through its comparison to an alternative measurement, made through a different standardized instrument or by means of an established standard.[2]

For angular measurements in clinical practice and in research, the universal manual goniometer is the most commonly used device[3-5] because it is easy to handle, has low cost and yields a rapid evaluation. However, published data have not assessed validity or reliability of the goniometer in comparison to other measuring devices[1] for all articulations.

The reliability of goniometry has been demonstrated for some joints. Engh et al. investigated the intra- and inter-rater reliability of goniometry measurements for the evaluation of head posture in relation to the shoulders and found the method to be reliable. Most studies on the knee joint present high intra-rater reliability.[6,7] As for the validity of goniometry, only comparisons between this method and radiographic measurements have been reported.[2]

The majority of studies that assess medial and lateral rotation movements of the shoulder make use of goniometry.

DOI: 10.5935/MedicalExpress.2014.02.08

These studies usually involve the assessment of the medial rotation deficit on the dominant shoulder of overhead throwing athletes compared to the non-dominant side.[8-11] These reports do not describe the reliability of goniometry. But in some cases the inter- and intra-rater comparison was performed before the execution of the study proper, with the intention of testing the accuracy of the researchers. In some cases, in which the inter-rater variation was very significant, evaluators were led to perform it in pairs, one to immobilize the subject and the other to execute the measurements with the goniometer; this resulted in the single ''assessment system''.[12]

Measurements based on photographs can be employed to determine articulation angles, introducing it as an alternative to goniometry. Photogrammetry began with cartography and was adapted for medical use, renamed biophotogrammetry,[12-13] or computerized biophotogrammetry when performed through a computer program.[12-16]

The objective of this research was to verify the inter- and intra-rater reliability of manual goniometry versus computerized biophotogrammetry on the assessment of the range of motion of shoulder medial and lateral rotations.

METHODS

Participants

Eleven male subjects were recruited, all right-handed (for daily-life chores and sport activities), with no pathological alterations or history of shoulder injury or need of medical assistance during the previous 12 months. Subjects were informed about the content of the research and signed an informed consent term. The study was initiated after Research Ethics Committee approval from the institution where it was conducted – Federal University of São Paulo (approval n. 2039/07), according to the Helsinki Declaration.

The subjects filled out a questionnaire involving identification, clinical data (including shoulder instability history) and training history, if existent. The exclusion criteria of this study would have been subjects with episodes of shoulder instability, but no such cases were identified.

To guarantee the homogeneity of the group concerning the subjects' range of motion, four tests for glenohumeral joint instability were performed: sulcus test, apprehension test, posterior elevation test and anterior-posterior drawer test. The same clinician ministered the questionnaire and specific tests to all subjects. Moreover, they were also evaluated through the joint hypermobility test. If any subject had scored four or more out of the nine items,[24] he would have been counted as positive and excluded from the test, but no such cases were encountered.

Intervention

To assess shoulder joint rotation (with goniometry as well as with biophotogrammetry), the subjects were initially placed in neutral rotation with 90° of shoulder abduction (measured by the goniometer) and elbow flexion. They lay in dorsal decubitus with their lower limbs flexed and feet placed upon the bed 84 cm above the floor. The forearm was maintained in pronation and wrist flexion-extension was not permitted during the evaluation.

For the goniometric assessments, a single experienced clinician stabilized the shoulder joint, always at the anterior portion of the glenohumeral joint. The sequence of rotational movements was randomly defined for each subject. Active and passive medial and lateral rotation movements were assessed.

For goniometric evaluation, a 20 cm plastic goniometer of Carci® (São Paulo, Brazil) was used and the reference marks were: fixed arm positioned in horizontal, consequently parallel to the bed; support placed on the olecranon; movable/free arm placed in the medial line of the forearm directed to the ulnar styloid process. In this step, two evaluators, (A and B) collected the measurements; evaluator A always collected his measurements before evaluator B, who was blinded to the first assessment.

In between the assessments performed by evaluator A and evaluator B, a photo was registered for biophotogrammetric evaluation. This was taken at a 200 cm distance from the bed, with the camera placed on a tripod and adjusted at the same height as the bed. The bed was maintained leveled using a Stanley GP level gauge. For vertical reference in the photo, we used two green circular markers, 13 mm of diameter, and 50 cm apart from each other.

The two adhesive green circular markers were fixed on the skin surface of each subject's upper limbs: one placed on the ulnar styloid process, the other on the olecranon.

The sequence of rotational movements for these assessments followed the same random order determined. Images were obtained through a digital photo camera Sony®, model DSC-W, with 5.1Mp of resolution; all images were obtained at 2592 x 1944 pixels.

The ''free angle measuring'' modality of the SAPO (0.67 version) freeware software was used to measure angular rotation, a method based on vertical and distance calibrations through the software. Two experienced clinicians, blinded from each other, obtained the measurements (Evaluators C and D).

Outcome measurements

Measurements through the software were executed with a 100% zoom, on computers with 1280x1024 pixels of screen resolution. All assessments were repeated seven days later for intra-rater comparison; they were designated D1 and D2. Participants were instructed to not to modify their habits concerning physical activities during the interval.

For inter-rater goniometric reproducibility analysis measurements obtained by Evaluators A and Evaluators B, D1 and D2 were used. For biophotogrammetric reproducibility inter-rater analysis, the measurements obtained by Evaluator C and D were also compared in D1 and D2. Intra-rater goniometric and biophotogrammetric repeatability analysis was obtained by internally comparing D1 and D2, for each of the four evaluators.

Data analysis

Wilcoxon signed-ranks test was used to evaluate possible statistical differences between the data. The value $p \leq 0{,}05$ was assumed as statistically significant.

To assess inter- and intra-rater reliability of each procedure, the Spearman correlation test was applied on all analysis and it was possible to verify types 1,1 and 3,1 Intra-class correlation coefficients (ICC), according to Yaffe,[17] within the analyzed data. The value $\geq 0{,}80$ was assumed as an excellent reliability level, as suggested by Yaffe.[17] The other reliability intervals were as proposed by Fleiss[18]: values $< 0{,}40$ represent poor reliability, while values between 0,40 and 0,79 represent good reliability.

RESULTS

The subjects mean age was 22.2 ± 3.2 (range 16-26) years; the average body weight was 75.9 ± 6.6 Kg; the mean height was 1.76 ± 0.06 m, and the mean body mass index (BMI) was 23.6 ± 2.0 Kg/m^2. A total of 9 of the assessed subjects practiced some kind of physical activity.

None of the subjects presented shoulder instability for apprehension and posterior displacement, on either side; the AP Drawer test was negative in 91% of the cases on both sides; the sulcus test was 91% negative for the right side, and 82% negative for the left side.

The statistical values of the data of each movement, of all evaluators, with each measuring method and on each day of evaluation, are displayed in Table 1.

Table 2 delineates the values of the Spearman correlation and the statistical significance index among all measurements obtained by each measuring method, with each evaluator and on each day of assessment.

Values of the Spearman correlation, significance index and ICC, confidence interval of 95% of the ICC and inter- and intra-rater standard error of each one of the procedures used for measuring are described in Table 3.

DISCUSSION

Several studies have measured shoulder rotation range of motion and have invariably used manual goniometry as the measuring method; however, the reliability indexes of this technique were not reported. These papers report medial rotation deficit in the dominant upper limb for throwing (Glenohumeral Internal Rotation Deficit – GIRD), a context which indicates primary functional diagnostics for clinical practice.

The precision of these functional diagnoses is of fundamental importance because they apply to preventive treatment for subjects predisposed to develop symptoms of shoulder pain due to such rotational alterations. Symptomatic subjects usually present medial rotation deficits above 25° compared to the non-dominant upper limb.[11] Therefore, it is important that, apart from reliability assessments, the validation of the methodology should be obtained. The procedures used for validation were described by Amorim et al.[15] and Ellenbecker et al.[16] Another fundamental issue is the standardization of the method, given that, without it, the results may be imprecise. Thus, the procedures used in this study were described in detail.

Three interesting points were observed in this study when the statistical differences between each one of the

Table 1 - Statistical values (p value) of the data of each one of the movements, of all evaluators, with each measuring method and on each day of evaluation

			Evaluator C		Evaluator D		Evaluator A		Evaluator B	
			p		p		p		p	
			D1	D2	D1	D2	D1	D2	D1	D2
Dominant	**Evaluator C**	**Active lateral rotation**		.929	.312	.953	.333	.109	.722	.374
		Passive lateral rotation		.929	.385	.047*	.016*	.041*	.110	.505
		Active medial rotation		.625	.066	.281	.004*	.010*	.003*	.010*
		Passive medial rotation		.859	> .99	.475	.013*	.004*	.006*	.003*
Non dominant		**Active lateral rotation**		.790	.045*	.754	.130	.929	.306	.212
		Passive lateral rotation		.790	.138	.241	.006*	.657	.021*	.859
		Active medial rotation		.424	.721	.538	.050*	.008*	.003*	.003*
		Passive medial rotation		.328	.261	.106	.004*	.003*	.004*	.003*
Dominant	**Evaluator D**	**Active lateral rotation**	.312	.953		.722	.328	.142	.657	.790
		Passive lateral rotation	.385	.047*		.722	.016*	.010*	.110	.534
		Active medial rotation	.066	.281		.859	.004*	.026*	.003*	.021*
		Passive medial rotation	> .99	.475		.286	.013*	.004*	.006*	.003*
Non dominant		**Active lateral rotation**	.045*	.754		.477	.110	.790	.286	.594
		Passive lateral rotation	.138	.241		.859	.006*	.374	.016*	.657
		Active medial rotation	.721	.538		.657	.062	.006*	.003*	.003*
		Passive medial rotation	.261	.106		.534	.004*	.003*	.004*	.003*
Dominant	**Evaluator A**	**Active lateral rotation**	.333	.109	.328	.142	.099	.099	.798	.036*
		Passive lateral rotation	.016*	.041*	.016*	.010*	.075	.075	.050*	.035*
		Active medial rotation	.004*	.010*	.004*	.026*	.894	.894	.563	> .99
		Passive medial rotation	.013*	.004*	.013*	.004*	.155	.155	.119	.006*
Non dominant		**Active lateral rotation**	.130	.929	.110	.790		.102	.330	.533
		Passive lateral rotation	.006*	.657	.006*	.374		.021*	.032*	.539
		Active medial rotation	.050*	.008*	.062	.006*		.824	.119	.056
		Passive medial rotation	.004*	.003*	.004*	.003*		.041*	.374	.635
Dominant	**Evaluator B**	**Active lateral rotation**	.722	.374	.657	.790	.798	.036*		.959
		Passive lateral rotation	.110	.505	.110	.534	.050*	.035*		.008*
		Active medial rotation	.003*	.010*	.003*	.021*	.563	1.00		.561
		Passive medial rotation	.006*	.003*	.006*	.003*	.119	.006*		.114
Non dominant		**Active lateral rotation**	.306	.212	.286	.594	.330	.533		.064
		Passive lateral rotation	.021*	.859	.016*	.657	.032*	.539		.010*
		Active medial rotation	.003*	.003*	.003*	.003*	.119	.056		.092
		Passive medial rotation	.004*	.003*	.004*	.003*	.374	.635		.041*

D1 = First day of evaluation; D2 = Second day of assessment.
* = $p \leq 0.05$.

Table 2 - The values of the Spearman correlation and the statistical significance index among all measures obtained with each measuring method, with each evaluator and on each day of assessment

	Evaluator C		Evaluator D		Evaluator A		Evaluator B	
	D1	D2	D1	D2	D1	D2	D1	D2
	r(p)	r(p)	r(p)	r(p)	r(p)	r(p)	r(p)	r(p)
Evaluator C		.93(.47)*	1.0(.22)*	.96(<.001)*	.87(.12)*	.85(<.001)*	.86(<.001)*	.81(<.001)*
Evaluator D	1.0(.22)*	.96(<.001)*		.96(.78)*	.87(.12)*	.87(<.001)*	.86(<.001)*	.85(<.001)*
Evaluator A	.87(.12)*	.85(<.001)*	.87(.12)*	.87(<.001)*		.85(<.001)*	.90(<.001)*	.89(<.001)*
Evaluator B	.86(<.001)*	.81(<.001)*	.86(<.001)*	.85(<.001)*	.90(<.001)*	.89(.003)*		.82(<.001)*

r = correlation Spearman. p = p value.
* = significant correlation at the level of 0.01.
Legend: D1 = First day of evaluation; D2 = Second day of assessment.

Table 3 - The values of the Spearman correlation, significance index and ICC, confidence interval of 95% of the ICC and inter- and intra-rater standard error of each one of the procedures used for measuring

	Goniometry		Biophotogrammetry	
	Inter-Examiner	Intra-Examiner	Inter-Examiner	Intra-Examiner
C. Spearman	.905**	.848**	.982**	.947**
Value P	<.001*	<.001*	.007*	.735
ICC	.897	.830	.982	.954
IC 95% do ICC	.844 - .930	.737 - .885	.976 - .987	.938 - .965
Standart Error	7.396	9.502	4.064	6.498

C = correlation; IC = confidence interval; ICC = confidence interval index.
* = $p \leq 0.05$; ** = correlation significance.

movements were assessed: (1) when the goniometric inter- and intra-rater comparisons where completed, three out of the eight assessed movements presented statistical differences: Passive lateral rotation; Active medial rotation; Passive medial rotation; (2) In contrast, the inter- and intra-rater biophotogrammetric comparison presented only one out of eight movements with statistically significant differences; (3) the method with fewer differences between assessments (D1 versus D2), was the inter- and intra-rater findings for biophotogrammetry.

When analyzing the statistical significance index and the Spearman correlation of all the measurements obtained with both measuring methods, five crucial points can be observed: (1) the goniometric measurement comparisons presented statistically significant differences in all inter- and intra-rater measurements; (2) the biophotogrammetric inter- and intra-rater comparisons presented statistically significant differences in only one case; (3) there were no statistically significant differences on the intra-rater biophotogrammetric comparison; (4) when the biophotogrammetric and goniometric measurements were compared, there were no statistically significant differences in six out of eight cases; (5) all Spearman correlation values were superior to 0,80.

Hence, the measurements that presented fewer differences between them were obtained through biophotogrammetry. The largest difference was observed when both techniques were compared, confirming the hypothesis that comparisons between them would not be indicated. For example, it would not be statistically reliable to compare the measurements obtained with goniometry in a determined day with the values obtained by biophotogrammetry in another day.

Inter- and intra-rater comparisons for both techniques achieved mostly excellent or otherwise strong values of ICC and of the Spearman correlation. Consequently, the results of this study demonstrated excellent reliability values for intra- and inter-rater evaluations, manual goniometry and computerized biophotogrammetry.

Nonetheless, biophotogrammetry proved to be more reliable. This is demonstrated by the superior values of the ICC and Spearman correlation on inter- and intra-rater biophotogrammetric comparisons as opposed to the goniometric data.

To the best of our knowledge, no studies comparing the reliability indexes between goniometry and biophotogrametry have been reported. Only one study determined whether there were statistically significant differences between values measured by these two techniques. Though the image caption conditions were not entirely satisfactory and another software was used for angular assessment (AutoCad®), the authors concluded that there were no statistically significant differences between the data and that the procedure was probably reliable.[15]

A study by Hayes et al.[19] assessed intra- and inter-rater reliability of five methods for shoulder range of motion evaluation. The intra- and inter-rater ICC rates for shoulder movements assessed by goniometry varied between 0.53-0.65 and 0.64-0.69 respectively, fulfilling the hypothesis that goniometry is not a reliable method for this evaluation, since reliability may only be considered as excellent if over 0.75. There were discrepancies in the ICC because it was calculated for each one of the movements assessed by this technique. These data disagree with the present study, where the intra- and inter-rater ICC values were 0.90 e 0.83 respectively, thus, with excellent reliability.

It is worth emphasizing that none of the five assessment methods for shoulder range of motion in the study conducted by Hayes et al.[19] presented reliability rates better than the goniometric and biophotogrammetric methods used in this research. Interestingly, in this same study, one of the methods analyzed was instant photography followed by manual angular measuring with goniometry, and this produced one of the highest reliability rates among the tested techniques.

Moreover, assessments conducted with biophotogrammetry have some advantages. Documentation of the assessment and visual feedback for evolution follow-up become available and, in case of doubt, measurements can be easily redone on the software. Displayed angles are measured to the first decimal digit, whereas goniometry only allows integer or approximate values, because of the precision limit of its scale.

There are disadvantages as well. If the standardization precautions are not followed when registering the photographs, reliability can be compromised. This also applies to goniometry. Photographic and computer competence are necessary, while goniometric assessments only require anatomical and positioning competence. Biophotogrammetry is more expensive in that it requires digital cameras and computers, even though this equipment is often already available.

The Amorim et al.[15] study also computed expenses with the software used (AutoCad®), while this research used the SAPO software freeware, which is highly reliable and, consequently, costs less.

The establishment of the standard error is very important to perform follow-ups and detect variations on the subjects' posture in response to treatment.[20-21] Moreover, in this study, both standard error values (inter- and intra-rater) were lower for biophotogrammetry compared to goniometry, a fact partially explained because biophogrammetric evaluators executed their assessments on the same photograph.

REFERENCES

1. Dunlevy C, Cooney M, Gormely J. Procedural considerations for photographic-based joint angle measurements. Physiother Res Int. 2005;10(4):190-200.
2. Bower E, Ashburn A. Princípios de conduta fisioterapêutica e medidas de resultado final. In: Stokes M, editor. Neurologia para fisioterapeutas, 4th ed. São Paulo; 2000. p 49-63.
3. Norkin CC, White DJ. Introdução à Goniometria. In: Norkin CC, White DJ, editor. Medida do Movimento Articular: Manual de Goniometria, 2nd ed. Porto Alegre; 1997. p 21-63.
4. Coxford P, Jones K, Barker K. Inter-tester comparison between visual estimation and goniometric measurement of ankle dorsiflexion. Physiother Theory Practice. 1998;14:107-13.
5. Engh L, Fall M, Hennig M, Soderlund A. Intra- and inter-rater reliability of goniometric method of measuring head posture. Physiother Theory Practice. 2003;9:175-82.
6. Brosseau L, Tousignant M, Budd J, Chartier N, Duciaume L, Plamondon S, et al. Intratester and intertester reliability and criterion validity of the parallelogram and the universal goniometers for active knee flexion in healthy subjects. Physiother Res Int. 1997;2(3):150-66.
7. Watkins MA, Riddle DL, Lamb RL, Personius WJ. Reliability of goniometric measurements and visual estimates of knee range of motion obtained in a clinical setting. Phys Ther. 1991;71(2):90-7.
8. Borsa PA, Dover GC, Wilk KE, Reinold MM. Glenohumeral range of motion and stiffness in professional baseball pitchers. Med Sci Sports Exerc. 2006;38(1):21-6.
9. Crockett HC, Gross LD, Wilk KE, Schwartz ML, Reed J, O'Mara J, et al. Pitchers Osseous Adaptation and Range of Motion at the Glenohumeral Joint in Professional Baseball. Am J Sports Med. 2002;30(1):20-6.
10. Lintner D, Mayol M, Uzodinma O, Jones R, Labossiere D. Glenohumeral internal rotation deficits in professional pitchers enrolled in an internal rotation stretching program. Am J Sports Med. 2007;35(4):617-21.
11. Myers JB, Laudner KG, Pasquale MR, Bradley JP, Lephart SM. Glenohumeral range of motion deficits and posterior shoulder tightness in throwers with pathologic internal impingement. Am J Sports Med. 2006;34(3):385-91.
12. O COFFITO. São Paulo: Conselho Federal de Fisioterapia e Terapia Ocupacional, n. 17, dez. 2002. 42 p.
13. Magazoni VS. Estudo correlacional entre a expansibilidade da caixa torácica e a capacidade vital pulmonar nos indivíduos portadores e não de espondilite anquilosante [dissertação]. Uberlândia (MG): Centro Universitário do Triângulo da UNITMG; 2000. 129 f.
14. Raine S, Twomey LT. Head and shoulder posture variations in 160 asymptomatic women and men. Arch Phys Med Rehabil. 1997;78(11): 1215-23.
15. Amorim DA, Ribeiro EA, Cordeiro GG, Silva MAS, Bustamante JCP. O programa AutoCAD 2000 (r) como forma de medida angular para articulações. Fisioweb [periódico na internet]. 2007 [acesso em: 05 05 2007] [9 páginas]. Disponível em: URL: <http://www.wgate.com.br/ fisioweb>.
16. Ellenbecker TS, Ellenbecker GA, Roetert EP, Silva RT, Keuter G, Sperling F. Descriptive profile of hip rotation range of motion in elite tennis players and professional baseball pitchers. Am J Sports Med. 2007; 35(8):1371-6.
17. Yafee RA. Enhancement of reliability analysis: application of intraclass correlations with SPSS/Windows v.8; available: www.nyu.edu/acf/ socsci/Docs/intracls.html. 1998.
18. Fleiss JL, 1986 apud Delitto A, Strube MJ. Reliability in the clinical setting. American Physical Therapy Association Research Section Newsletter. 1991;24:2-8.
19. Hayes K, Walton JR, Szomor ZL, Murrell GAC. Reliability of five methods for assessing shoulder range of motion. Aust J Physiotherapy. 2001;47(4):289-94.
20. Burkhart SS, Morgan CD, Kibler WB. The Disabled Throwing Shoulder: Spectrum of Pathology Part I: Pathoanatomy and Biomechanics. Arthroscopy. 2003;19(4):404-20.
21. Guaratini MI. Confiabilidade e precisão da media para teste-reteste no dinamômetro isocinético biodex. [Dissertação]. São Carlos (SP): Universidade Federal de São Carlos; 1999.