ORIGINAL ARTICLE

Analysis of the interexaminer reliability of two clinical tests to measure the flexion range of motion of the lumbar spine

Análise da confiabilidade entre-examinadores de dois instrumentos que mensuram a amplitude de movimento de flexão da coluna lombar

José Carlos Taveira de Souza Filho[1], Ana Carolina do Vale Abras[1], Maíra Teixeira Carvalho[1], Miliany Graziele Fonseca Souza[1], André Taveira de Souza[1], Leonardo Oliveira Pena Costa[2]

ABSTRACT

Background: The measurement of the flexion range of motion of the lumbar spine is a common practice in clinical settings. Many methods are used to carry out these measurements, such as the Back Range of Motion Inclinometer (BROMII) and the Flexible Ruler methods. Objective and Methods: The aim of this study was to analyze the interexaminer reliability for the measurement of the lumbar spine flexion by using the BROMII and the Flexible Ruler methods. Thirty-seven physical therapy students were recruited as volunteers and a double-blind test-retest study design was carried out. Results: The results showed moderate reliability of the BROMII measurements [ICC (2.1) 0.71 (95%CI 0.49-0.84) p<0.000]; however, the reliability of the Flexible Ruler measurements was poor [ICC(2.1) 0.37 (95%CI 0.06-0.62), p<0.012]. Conclusion: The BROM II presented sufficient reliability to be performed in clinical practice.

KEYWORDS

range of motion articular, lumbosacral region, validity of tests/instrumentation

RESUMO

Introdução: A mensuração da amplitude de movimento de flexão da coluna lombar é uma prática clínica comum. Vários são os métodos para tais medidas, destacando-se entre elas duas ferramentas clínicas: o Inclinômetro Back Range of Motion II (BROM II) e a Régua Flexível. Métodos: O objetivo desse estudo foi analisar a confiabilidade entre-examinadores para as medidas de flexão da coluna lombar utilizando o BROM II e a Régua Flexível. Trinta e sete estudantes de Fisioterapia foram examinados num design teste-reteste duplo-cego Resultados: Os resultados mostraram confiabilidade moderada para as medidas com o BROM II [CCI (2,1) 0,71 (IC 95% 0,49-0,84) p<0,000] e pobre para a Régua Flexível [CCI(2,1) 0,37 (IC 95% 0,06-0,62) p<0,012]. Conclusão: Conclui-se que o BROM II apresentou confiabilidade suficiente para sua utilização na prática clínica.

PALAVRAS-CHAVE

amplitude de movimento articular, região lombossacral, validade dos testes /instrumentação

Souza Filho JCT, Abras ACV, Carvalho MT, Souza MGF, Souza AT, Costa LOP.
Analysis of the interexaminer reliability of two clinical tests to measure the
flexion range of motion of the lumbar spine

INTRODUCTION

The assessment of the range of motion (ROM) of the lumbar column in the clinical setting represents an important biomechanical factor in the functional diagnosis and treatment of many disorders of the vertebral column.[1] The ROM measurement is commonly used as a routine method in clinical practice and research projects with the objective of establishing functional limits, in addition to being useful in the follow-up of the treatment responses. This is an important measurement in the diagnosis and rehabilitation process in situations of musculoskeletal system dysfunction.

The gold standard to measure ROM of the lumbar column is the use of x-rays.[1] However, its use in the clinical setting is restricted, as it is an expensive, invasive method that can result in side effects. Among the resources the measure the ROM of the lumbar column are Schober's test,[2] the fingertip-to-floor method,[3] the use of inclinometers and the flexible ruler. One of the main problems of clinical measurements in rehabilitation is due to the fact that many of these tests are used by health professionals when the necessary repetitiveness (reliability) and accuracy (validity) studies have not been performed. In addition, the values expressed in these tests vary not only due to the different degrees of aptitude and experience of the examiners, but also due to the lack of test standardization.[1] The ideal situation would be for the professionals to have a number of simple, standardized and reliable available tests to assess their patients and that different examiners attained similar results when evaluating the same patient. In the case of trunk flexion ROM measurements, there are currently two good tests that are available to health professionals and do not pose any risk to the patient's and/or volunteer's health: the Back Range of Motion II (BROM II) instrument [4] and the flexible ruler.[5]

The BROM II (Performance Attainment Associates, St. Paul – MN) is an instrument used to obtain three types of ROM measurements: flexion/extension, rotation and lateral inclination of the lumbar and thoracic column (Figure 1-A). The flexion/extension unit, used in this study, is a modified inclinometer that consists of a basis with a scale positioned through points of contact on the first sacral vertebra (S1) and an acrylic shaft that is fitted at the basis with its extremity positioned on the last thoracic vertebra (T12). The instrument is attached to a Velcro® strap that must be crossed over the lower abdomen, which allows the maintenance of the points of contact fixed at S1 during the flexion and extension of the lumbar column.
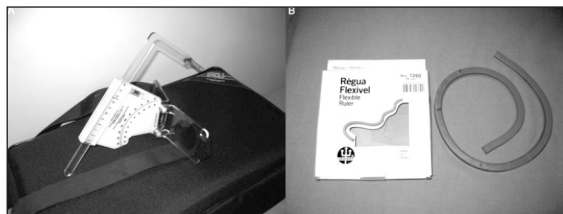

Figure 1
A: BROM II / B: Flexible Ruler

The flexible ruler (FR) - (Trident® Indústria de Precisão, Brazil) – is a flexible metal shaft covered by flexible plastic that has the capacity of deforming and that can be molded to any surface (Fig. 1B). The spinal processes L1 and S2 are used as points of reference and the curvature of the lumbar column in the neutral position as well as in flexion obtained by the FR can be reproduced through a drawing on paper.[5]

The ROM of the lumbar column can be influenced by individual factors such as age, sex, weight, height, anatomical variations and time of the day when the measurement was performed. These influences are still considered debatable and no correlation between ROM and age, sex and weight of the patients has been found, but some authors point out the importance of investigating the lumbar column ROM at the same time of the day, as significant differences have been found in measurements made at different times of the day.[6] The mean ROM values increased significantly along the day due to important factors such as the increase in muscular temperature and changes in the structure of the intervertebral discs.[6]

All measurement methods present some type of error. The professionals and researchers must be aware of the extent of this error, to confirm that the intervention was in fact effective or to decide which measurement method must be used for an assessment.[7]

OBJECTIVE

The objective of the study was to evaluate the Interexaminer reliability of BROM-II and the Flexible Ruler in the measurement of the trunk flexion in normal individuals.

METHODS

Sample

The study was carried out at the Clinical Physical Therapy Center of PUC-MINAS in the Belo Horizonte campus. The sample calculation was based so that the tests were sensitive to the point of detecting a reliability rate between 0.3 and 0.9 with a confidence interval of 95% and it was concluded that 30 individuals would be enough for the analysis.

All the volunteers were informed about the procedures and objectives of this study and the authorization was obtained from an informed consent form approved by the Ethical Review Board of the institution. The demographic data (name, age and gender) of each volunteer were obtained before the tests.

To be included in the study, the individuals had to be 18 to 25 years of age and a current student of the Physical Therapy course at PUC-Minas. Individuals with a history of lumbar pain in the previous 12 months or any other severe pathology of the vertebral column (fractures, tumors, cauda equina syndrome, etc), were excluded from the study. The flexion ROM of the lumbar column was evaluated using the BROM II inclinometer and the FR.

Procedures

The study was carried out with a test-retest design and had a mean duration of 10 minutes for each examiner. Both testers per-

ACTA FISIATR 2007; 14(4): 214 - 218

Souza Filho JCT, Abras ACV, Carvalho MT, Souza MGF, Souza AT, Costa LOP.
Analysis of the interexaminer reliability of two clinical tests to measure the
flexion range of motion of the lumbar spine

formed the two tests (BROM II and FR) on the same day with the volunteers, with no interval between the measurements performed by the two testers (one examiner started his or her tests right after the data collection carried out by the preceding examiner).

The testers were blinded for the measurements obtained by the preceding examiner; the sequence of the testers and the tests was randomized so that their order would not influence the results; additionally, the anatomical references were erased after each measurement, so that they would not interfere with the measurements obtained by the following examiner. At the analysis by the BROM II, the mean was calculated after three repetitions. As for the measurement carried out by the FR, due to practical factors and availability of the volunteers, only one measurement was obtained (value of the neutral position added to the value of the flexion position).

For both tests, the volunteers were placed in a standing position, on a line that was previously fixed on the floor, so that they formed a right angle, keeping feet and knees aligned with the hip. During the measurement, the volunteers were advised to maintain the eyes focused on the horizon. The volunteers remained standing in front of and with the back turned to the examiner, who, in the sitting position, performed the palpation and marked the anatomical references related to the instrument with a marker pen.

The examiner positioned the BROM II over the spinal process S1 and the volunteer was asked to fix the straps crossing them over the lower abdominal region. Then, the examiner verified whether the inclinometer was fixed and positioned on the reference and placed him or herself to the right side of the volunteer, looking at the right side of the volunteer's body. The shaft of the BROM II was placed on T12, so that the shaft line was positioned in the middle of the markings made by the marking pen. The examiner carried out the reading for the assistant so that his or her eyes were fixed on the straight line marking.

Subsequently, the volunteer was asked to perform a trunk flexion, sliding his or her hands along the legs and letting the arms hang down at the end of the movement. Once more, the examiner read the angle registered at the BROM II and asked the volunteer to return to the initial position (Figure 2).

Then, the examiner performed the palpation and anatomical reference markings (S2 and L1) and positioned the FR on the volunteer's lumbar column with the brand TRIDENT to the right (the FR has a brand marking that says TRIDENT®, which was used as a point of reference for the positioning of the FR during the tests; this was necessary because the FR format is not homogeneous throughout all its dimensions). Using a marking pen, the examiner drew the anatomical references (S2 and L1) on the ruler with a straight line. Then, the FR was removed carefully and using both hands, the ruler was placed on sheet of paper (with the TRIDENT® brand marking facing upwards) and reproduced the curvature molded by the ruler. The upper margin of the straight line of L1 and the lower margin of S2 made by the ruler were considered upon the drawing of the tracing (Figure 3).

Figure 3
Transference of the curvature measurement onto paper.

Next, the examiner asked the volunteer to perform an anterior trunk inclination by sliding the hands over the legs and letting the arms hang down at the end of the movement. Once again the examiner made the anatomical reference markings on the ruler and reproduced the drawing on paper.

To calculate the angle measured by the flexible ruler the following equation was used: $\hat{a} = 4 \times [\arctan (2\,h\,/L)]$.[5] The distance of the two extreme points of the curve is called "l", and "h" is the distance of the center of "l" to the farthest point of the curve, perpendicularly (Figure 4). The values of "l" and "h" are given in centimeters.
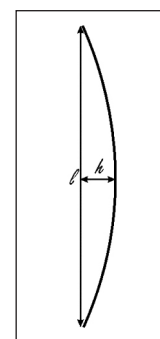

Figure 4
Schematic representation on paper of the drawing obtained from the FR.

Figure 2
Test procedures using the BROM II.

The measurements of the BROM II and the FR were tabulated using the software Microsoft Excel 2002. The measurements were standardized as follows: the mean value of three measurements of the ROM of each flexion of the volunteer was considered. At the

ACTA FISIATR 2007; 14(4): 214 - 218

Souza Filho JCT, Abras ACV, Carvalho MT, Souza MGF, Souza AT, Costa LOP.
Analysis of the interexaminer reliability of two clinical tests to measure the
flexion range of motion of the lumbar spine

end of each anterior inclination, the value obtained at the end of the flexion, subtracted by the value obtained at the neutral position, was called Actual Flexion. To consider that ROM of the actual flexion of each volunteer, a simple arithmetic mean was calculated. In the case of the FR, the sum of values obtained for the neutral position and for the flexion was considered as the flexion ROM of each volunteer.

The interexaminer reliability was calculated by the Intraclass Coefficient of Correlation (ICC) type (2,1) and the respective 95% Confidence Intervals (CI), through the statistical package SPSS 14.0 for Windows.

The interexaminer ICC was also calculated, in an exploratory way, of the initial and final positions of each test. The rationale is that the initial and the final positions can significantly influence the results of trunk ROM, and thus, possible improvements in the measurement protocol can be elucidated by this analysis. Descriptive calculations (means and standard deviations) were also performed by each examiner and these mean values were compared by the Student's t test (paired) with $\alpha=0.05$.

RESULTS

The data of 34 of the 37 recruited volunteers were analyzed. The mean age of the volunteers was 22 years (± 3.90) and 12 of them were males (22 females). Only one volunteer did not attend the re-test and two volunteers were excluded from the study, as they presented an important lateral deviation, which impairs the measurement with the FR.

The mean ROM values observed by examiner 1 were 66.9 ± 22.8 (BROM II) and 57.6 ± 10.9 (FR); the values observed by examiner 2 were 63.7 ± 21.1 (BROM II) and 53.8 ± 12.8 (FR). No statistically significant differences were observed at the direct comparison between the means of each test [(BROM II – t = 1.12; p=0.27) (FR – t = 1.66 p= 0.10).

The interexaminer ICC for the values obtained with the BROM II was 0.71 (95% CI: 0.49-0.84) p<0.000; for the FR it was 0.37 (95% CI: 0.06-0.62) p<0.012. The descriptive data regarding the interexaminer measurements in the neutral position and in flexion for each tool are described in Table 1.

Table 1
ICC values for measurements in the neutral position and in flexion for BROM II and FR.

	ICC value	IC 95%	p
BROM II			
Neutral	0.74	0.41 – 0.88	p<0.000
Flexion	0.65	0.40 – 0.80	p<0.000
FR			
Neutral	0.31	0.00 – 0.58	p<0.02
Flexion	0.81	0.59 – 0.91	p<0.000

DISCUSSION

Studies of reliability take into account the capacity of an ex-

aminer in performing a test and repeat it, trying to obtain similar results or results that are as close as possible to the previous ones.[8] The present study was a double-blind test-retest study, i.e., two examiners performed the measurements without any knowledge of the previously collected data and performed the palpation of the anatomical references without following the anatomical markings made by the previous examiner, which increases the possibility that the reliability levels are lower; on the other hand, the procedures carried out in this experiment better reproduce the clinical practice.

Previous studies have shown good reliability regarding the BROM II[4] and FR,[5] however, the literature search did not show any "blind" study carried out with the FR. In a previous study, the flexion measurements were carried out with the BROM II, in a blind study design and the authors obtained an inter-examiner ICC of 0.74 in a sample 91 volunteers.[4] The biggest difference of this study was that the researchers used an independent reader, who only read the information to an assistant. We believe that reliability studies must be carried out in situations that are similar to those of the clinical practice and therefore, the present study was carried out with only one examiner throughout the test. When comparing the results of the two studies, it is observed that there were no differences between the ICC values and thus, we suggest that the presence of one more examiner is not necessary.

The BROM II and the FR are tools that allow the obtaining of actual measurements of lumbar flexion with no influence of the thoracic column or the hip. Several steps were taken to ensure that all the procedures of all measurements were carried out using the best standardized form available (previous training of the examiners, carrying out of pilot studies and thorough reading on the subject). However, it can be observed by the results that there was no consensus regarding the neutral position during the test with FR, which negatively influenced the reliability of this procedure. Thus, special care should be given to the neutral positioning in future studies with this instrument.

Another possible factor that influences the measurement with the FR would be that the deformation the FR undergoes in the neutral position is higher due to the fact that it encounters a barrier of soft tissue, which makes the modeling on the spinal processes more difficult. The FR in the neutral position is modeled over muscles and ligaments that are loose, on adipose tissue and the skin, which is under less tension. In this case, the force exercised on the spinal processes must have been different between the examiners. The force exercised by the examiner so that the FR was modeled on the spinal processes was not controlled in this study and no finding on this type of standardization (quantification of the applied force) was found in previous studies that concern the use of the FR.

The opposite happens in the flexion position, where the soft tissue (muscles, tendons, adipose tissue and skin) are under a high degree of tension, which does not allow much deformation by the FR and does not vary with the applied force, as there is resistance to this force caused by the fact that the spinal processes are more evident. The ICC of each measurement can hypothesize this finding (Table 1). As reliability studies evaluate the capacity of reproducibility of values, the standardization in data collection is mandatory

ACTA FISIATR 2007; 14(4): 214 - 218

Souza Filho JCT, Abras ACV, Carvalho MT, Souza MGF, Souza AT, Costa LOP.
Analysis of the interexaminer reliability of two clinical tests to measure the
flexion range of motion of the lumbar spine

and decisive for the results.

One important limitation of this study that could probably explain the considerable difference between the ICC of the two tools can be explained by the fact that only one measurement was performed with the FR and not a mean of three repetitions, as carried out for the BROM II, which might certainly have increased the examiners' mean error. This choice was due to the fact that most previous studies involving the FR[5,9] also standardized only one measurement for each position (neutral and flexion). These differences in the number of repetitions of the same test were also found in other studies that compared the FR with radiological examinations and again, controversial results were observed, although isolated tests showed a high degree of intra-and interexaminer reliability.[5,9] It is necessary to perform a new study to clarify whether the use of the FR is really little reliable or if these contradictory results occurred due to methodological limitations in the different studies.

Although the FR allows the measurement of the lumbar flexion with no interference of the hip and the thoracic column, its use is complicated by the systematic mathematical calculations that demand time and many times require the use of more specific calculation instruments. One advantage of the use of the BROM II over the FR is to allow the analysis of movement at other planes (inclinations and extension) in a clinical setting; however, its use must still remain limited to the clinical scenarios, considering that the validity of this tool compared to radiographic examinations has yet to be determined.[1,4] In a study carried out with the CROM (Cervical Range of Motion), a device that is similar to the BROM II that measures the ROM of the cervical column, the authors point out to possible errors that can occur related to the reading difficulty and inaccuracy, incorrect perceptions of the end of the movement by the volunteers and the different levels of effort presented by the volunteers during the measurements.[10]

It is necessary to investigate the possible alterations in the protocols of these tests that could lead to an increase in the respective reliability rates, as well as perform more validation studies (correlating them with "gold-standard" tests), so that these tests can be performed by professionals as accurately and precisely as possible.

CONCLUSION

The results of this study demonstrated that the BROM II presents enough reliability to justify its use in clinical practice, a fact that cannot be attributed to the FR. In research practices, in which accuracy is essential, the use of imaging tests to measure the ROM of trunk flexion is still the best option.

ACKNOWLEDGEMENTS

REFERENCES

1. Cakir B, Richter M, Kafer W, Wieser M, Puhl W, Schmidt R. Evaluation of lumbar spine motion with dynamic X-ray--a reliability analysis. Spine. 2006;31(11):1258-64.
2. Williams R, Binkley J, Bloch R, Goldsmith CH, Minuk T. Reliability of the modified-modified Schöber and double inclinometer methods for measuring lumbar flexion and extension. Phys Ther. 1993;73(1):33-44.
3. Gauvin MG, Riddle DL, Rothstein JM. Reliability of clinical measurements of forward bending using the modified fingertip-to-floor method. Phys Ther. 1990;70(7):443-7.
4. Kachingwe AF, Phillips BJ. Inter- and intrarater reliability of a back range of motion instrument. 1: Arch Phys Med Rehabil. 2005;86(12):2347-53.
5. Hart DL, Rose SJ. Reliability of a noninvasive method for measuring the lumbar curve. J Ortho Sport Phys Ther.1986;8(4):I80-4.
6. Ensink FB, Saur PM, Frese K, Seeger D, Hildebrandt J. Lumbar range of motion: influence of time of day and individual factors on measurements. Spine. 1996;21(11):1339-43.
7. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979;86(2):420-8.
8. Nitschke JE, Nattrass CL, Disler PB, Chou MJ, Ooi KT. Reliability of the American Medical Association guides' model for measuring spinal range of motion. Its implication for whole-person impairment rating. Spine. 1999;24(3):262-8.
9. Mosner EA, Bryan JM, Stull MA, Shippee R. Investigation of the flexible ruler as a noninvasive measure of lumbar lordosis in black and white adults female sample populations. J Orthop Sports Phys Ther. 1989;11:3-7.
10. Carvalho CO, Magalhães DAS, Silva Junior JAA, Bicalho LFH, Costa APB, Costa LOP, et al. Estudo comparativo das amplitudes de movimento da coluna cervical em idosos com diferentes níveis de aptidão física. Acta Fisiatr. 2006;13(3):147-51.