ABSTRACT | The objective of this study was to analyze the reliability of the intra and inter-evaluators measures, with universal goniometer and fleximeter. Eighty university students (45 women and 35 men [average±SD]; age: 20.81±2.63 years, body mass: 68.36±16.31 kg, height: 1.69±0.09 m, BMI: 23.88±4.15 kg/m²) were submitted to goniometry and fleximetry of elbow and knee flexion, ankle dorsiflexion and wrist extension. It was calculated the intraclass correlation coefficient (ICC), considering a significance level of 5%. Overall, the results showed very strong (ICC: 0.91≤0.99; P<0.001) and strong (0.75≤0.90; P<0.001) correlations in all movements, intra and inter-evaluators. On inter-correlation instruments, there are strong coefficients on all measures (0.75≤0.90; P<0.001) for the two instruments. In the studied movements, the evaluator experience did not influence the measurements and both the universal goniometer and the fleximeter showed high reliability.

Keywords | Movement; Measures; Students.

RESUMO | O objetivo deste estudo foi analisar a confiabilidade das medidas intra e inter-avaliadores, com goniômetro universal e flexímetro. Oitenta universitários (45 mulheres e 35 homens [média±DE]; idade: 20,81±2,63 anos, massa corporal: 68,36±16,31 kg; estatura: 1,69±0,09 m e IMC: 23,88±4,15 kg/m²) foram submetidos à goniometria e fleximetry de flexão do cotovelo e joelho, dorsiflexão e extensão de punho. Foi calculado o coeficiente de correlação intraclass (CCI), considerando um nível de significância de 5%. De modo geral, os resultados mostraram correlações muito fortes (CCI: 0,91≤0,99; P<0,001) e fortes (0,75≤0,90; P<0,001), em todos os movimentos, tanto intra quanto inter-avaliadores. Na correlação inter-instrumentos, constatou-se fortes coeficientes em todas as medidas (0,75≤0,90; P<0,001), para os dois instrumentos. Nos movimentos estudados, a experiência do avaliador não influenciou nas medições e tanto o goniômetro universal quanto o flexímetro mostraram alta confiabilidade.

Descritores | Movimento, Medidas, Estudantes.

RESUMEN | El objetivo de este estudio fue analizar la fiabilidad de las medidas inter y intra-evaluadores, con goniómetro universal y flexímetro. Ochenta estudiantes universitarios (45 mujeres y 35 hombres [media±DE], edad: 20,81±2,63 años, masa corporal: 68,36±16,31 kg, estatura: 1,69±0,09m, IMC. 23,88±4,15 kg/m²) fueron sometidos a goniometría y fleximetry de la flexión del codo y extensión de la rodilla y la flexión dorsal de la muñeca. Se calculó el coeficiente de correlación intraclass (ICC), considerando un nivel de significación del 5%. En general, los resultados mostraron correlaciones muy fuertes (CCI: 0,91≤0,99; P<0,001) y fuertes (0,75≤0,90; P<0,001) en todos los movimientos, tanto inter como intra-evaluadores. En los instrumentos intercorrelación, encontramos fuertes coeficientes en todas las medidas (0,75≤0,90; P<0,001) para ambos los instrumentos. En los movimientos estudiados, la experiencia del evaluador no influye en las mediciones y tanto el goniómetro como flexímetro mostró una alta fiabilidad.

Palabras clave | Movimiento, Medidas, Estudiantes.
INTRODUCTION

Measurement of the joint range of motion (ROM) is an important component in the physiotherapeutic evaluation since it not only identifies joint limitations but it also allows following quantitatively the efficiency of therapeutic interventions during rehabilitation. The ROM measurement depends on three factors: • instrument being used; • examiner’s skill and experience; and • characteristics of the assessed subjects.

It should be performed using reliable instruments, preferably noninvasive ones, which may be easily reproduced and should be based on scientific evidence. Reliability of measures demonstrates its consistence obtained through an instrument or by an examiner in the same evaluation conditions. Intra-evaluator reliability is the consistence of measures performed in the same evaluation conditions in two different moments. On the other hand, the inter-evaluator reliability is associated with consistence of the measures performed by two or more different evaluators.

Among these instruments, both the universal goniometer (UG) and the fleximeter (electromagnetic goniometer) have showed reliability indices classified from “regular” to “excellent” depending on the joint and movement being analyzed.

Although literature shows that variations between the second and seventh in the ROM are acceptable considering the tested joint, and there has already been evidence of the reliability of UG and fleximeter alone, studies in general restrict the number of instruments when two or more joints are being compared. Or they limit the number of movements to be analyzed when two or more instruments are being correlated.

Thus, the hypotheses of this study are:
• the measure would present lower reliability when done by one not very experienced evaluator;
• manual goniometry is an evaluator-dependent measure, therefore it would have lower reliability than fleximetry; and
• there would be a strong reproducibility of inter-instrument measures.

Thus, this study aimed at analyzing the reliability of intra and inter-evaluators measures for elbow and knee flexion movements, ankle dorsiflexion and fist extension, which were performed with the UG and fleximeter.

METHODOLOGY

Sample

In this study, the sample was chosen by convenience and 80 healthy subjects from both sexes were investigated (45 women and 35 men; 20.81±2.63 years old, 68.36±16.31 kg, 1.69±0.09 m and BMI: 23.88±4.15 kg/m²). They had no musculoskeletal alterations, pain or discomfort in the higher and lower limbs or any other disease or dysfunction that could compromise the elbow, fist, knee and ankle ROM.

The study was approved by the Research Ethics Committee in Human Beings of the University Hospital Lauro Wanderley (UFPB), according to protocol CEP/HULW number 436/2010, with the Certificate of Presentation for Ethical Appreciation (CAAE) number 0361.0.126.000-10. Subjects who agreed to participate signed the free informed consent (FIC), according to resolution 196/96 of the Brazilian National Health Counsel (NHC).

Procedures

The following were used as instruments: UG (Carci, Brazil) and fleximeter (Instituto Code de Pesquisa, Brazil). In the UG, the part set apart to read the angles was covered with a poster paper in order to not influence the subsequent measures (Figures 1A, 1C, 1E and 1G), while in the fleximeter, reading was always direct since the evaluator could not have an influence on the results.

The study included two evaluators: one deemed experienced (three years of practice) and another considered inexperienced (only one month of practice), with low familiarization with the goniometer and fleximeter. Each measure was performed three times in the dominant limb, in only one session, by calculating their mean. Subjects were oriented to wear light clothes in order that the ROMs were not limited, and the evaluator demonstrated all movements before the maximum active ROM measuring. One of the evaluators did measurement, while the second one performed its reading and recording in order to avoid induction of results.

The order to measure movements (elbow and knee flexion; ankle dorsiflexion and fist extension) and use instruments (UG and fleximeter) was randomly performed (http://www.lee.dante.br/). After each measure, the subject was requested to return to his/her initial position keeping a 1-minute interval between the three
measures, for each movement (intra-evaluator), and a 5-minute one between evaluators (inter-evaluators) and instruments (inter-instruments).

In order to identify dominance, the subject was questioned on what leg he/she uses to kick a ball and with what hand he/she writes, with the purpose of confirming the dominance of the lower (LL) and higher limb (HL) respectively.

When measuring elbow flexion (Figure 1A and 1B), the subject was positioned in the supine position with the HL in supination and extended beside the body with something to support him/her in the distal portion of the arm dorsal face in order to check a possible elbow hyperextension.

Measures of knee flexion (Figure 1C and 1D) and ankle dorsiflexion (Figure 1E and 1F) were performed in the prone position. In the first movement, the subject was positioned with his/her LL extended with something to support it in the distal region of the femur in order to check the existence of a knee hyperextension. For the dorsiflexion, the same position was kept with the addition of flexion (90º) of the knee that was going to be evaluated.

For measuring the fist extension (Figure 1G and 1H), subject sat in a chair with his/her shoulder adducted,
inflected elbow (≥90º), forearm at pronation and fist in neutral position supported on a board19.

In all measures done with the fleximeter, the same positions of goniometry were adopted. The fleximeter was positioned in the lateral face of the fist (radius styloid process), of the ankle (lateral malleolus), of the foot (head of the fifth metatarsus) and medial face of the hand (fifth metacarpus) to measure elbow and knee flexion, ankle dorsiflexion and fist extension respectively20.

**Data analysis**

Statistical procedures were carried out in the software Statistical Package for the Social Sciences (SPSS – 15.0). Initially, data normality (Kolmogorov-Smirnov’s test) and variance homogeneity (Levene’s test) were observed followed by the intraclass correlation coefficient (ICC – reproducibility) in the comparison of ROMs of all intra and inter-evaluators (Av1 X Av2) joints (knee, ankle, elbow and fist) done with UG and fleximeter, using a 5% level of significance.

For the ICC analysis, the following classification was considered: null=0.00; weak=0.01 to 0.30; regular=0.31 to 0.60; strong=0.61 to 0.90; very strong=0.91 to 0.99; and complete=1.0021.

### RESULTS

ICC tests between the three measures done with UG for each evaluator demonstrated that there was a very strong correlation (0.91≤0.99; P<0.001) in all assessed movements, both for the evaluator 1 and evaluator 2, with the exception of the knee flexion movement performed by evaluator 2 (0.83; P<0.001) that had a strong correlation (Table 1).

With regard to comparisons between the three measures of ROMs performed using the fleximeter (Table 2), results also showed a very strong ICC (0.91≤0.99; P<0.001) in all analyzed movements, both for evaluator 1 and evaluator 2.

As to comparisons of inter-evaluators measures (Table 3), it was also seen a very strong ICC for elbow and knee flexion movements (0.95 and 0.93; P<0.001, respectively) done using UG, and for movements of knee flexion and fist extension (0.92 and 0.93; P<0.001, respectively) with the fleximeter. All the other results, regardless of the instrument used (UG or fleximeter), achieved a strong ICC (0.8≤0.9; P<0.001).

The results of ICC between instruments (UG x fleximeter) found strong correlations (0.76 to 0.90; P<0.001) for all the analyzed movements (Table 4), both for evaluators 1 and 2.

### Table 1. Values of the intraclass correlation coefficient (ICC), intra-evaluators, of elbow and knee flexion goniometry, fist dorsiflexion and extension

<table>
<thead>
<tr>
<th>Movements</th>
<th>Evaluator 1</th>
<th>P-value</th>
<th>Evaluator 2</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elbow flexion</td>
<td>0.9823</td>
<td>&lt;0.0001</td>
<td>0.9792</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Knee flexion</td>
<td>0.9836</td>
<td>&lt;0.0001</td>
<td>0.8282</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Dorsiflexion</td>
<td>0.9522</td>
<td>&lt;0.0001</td>
<td>0.9508</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Fist extension</td>
<td>0.9740</td>
<td>&lt;0.0001</td>
<td>0.9622</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

### Table 2. Values of the intraclass correlation coefficient (ICC), intra-evaluators, of elbow and knee flexion fleximetry, fist dorsiflexion and extension

<table>
<thead>
<tr>
<th>Movements</th>
<th>Evaluator 1</th>
<th>P-value</th>
<th>Evaluator 2</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elbow flexion</td>
<td>0.9819</td>
<td>&lt;0.0001</td>
<td>0.9743</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Knee flexion</td>
<td>0.9776</td>
<td>&lt;0.0001</td>
<td>0.9672</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Dorsiflexion</td>
<td>0.9560</td>
<td>&lt;0.0001</td>
<td>0.9428</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Fist extension</td>
<td>0.9898</td>
<td>&lt;0.0001</td>
<td>0.9777</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
DISCUSSION

In the physiotherapeutic practice, a measure cannot be deemed significant if it is not valid and reliable. During a physiotherapeutic treatment, the patient usually is assessed more than once to check his/her clinical evolution and sometimes even by different physical therapists. Therefore, reliability of the examiner and of the ROM evaluation instrument is very important to allow the consistence of measures during the clinical evolution and in scientific studies. According to Bonagamba et al. and Côté et al., the most proper statistical test for the reliability analysis is ICC because it analyses data association in a certain time interval.

When comparing data from the present study with literature, it was seen that the ICC indices generally were similar or even higher than in other studies, both concerning evaluators and instruments, in most studied movements.

Under et al. study analyzed the knee flexion of 52 subjects who had undergone total arthroplasty, in three different positions (prone, supine and sat), and they also observed very strong correlations (ICC≥0.91) between inter and intra-evaluators positions.

Similarly to our results, Lustosa et al. and Lessen achieved high reliability rates in the measures for both instruments (UG and fleximeter) and between evaluators (0.87 and 0.99) when there was an analysis of knee and elbow flexion movements, therefore despite UG be an instrument of measure considered evaluator-dependent, it also reached similar correlation levels to the fleximeter.

Piriyaprasarth and Morris, in a systematic revision article, stated that ROM measures are reliable using both the UG and the electromagnetic goniometer (fleximeter), which is in agreement with the findings of this study.

On the other hand, Sacco et al. found moderate to weak correlations for the knee and ankle joints, respectively. However, it is believed that these lower correlation levels, if compared to our study, happen because the researchers used a different methodology from the present study.

An investigation developed by Chaves et al. including 106 healthy children, also correlating UG and fleximeter, found contrarily to the results from this study, regular and weak correlations. Nevertheless, although they have used the same instruments, they assessed different movements (cervical spine) that may justify the

<table>
<thead>
<tr>
<th>Movements</th>
<th>Universal goniometer</th>
<th>Fleximeter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC (r)</td>
<td>P-value</td>
</tr>
<tr>
<td>Elbow flexion</td>
<td>0.9496</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Knee flexion</td>
<td>0.9342</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Dorsiflexion</td>
<td>0.8825</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fist extension</td>
<td>0.8599</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Movements</th>
<th>Evaluator 1</th>
<th>Evaluator 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC (r)</td>
<td>P-value</td>
</tr>
<tr>
<td>Elbow flexion</td>
<td>0.7651</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Knee flexion</td>
<td>0.7971</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Dorsiflexion</td>
<td>0.8287</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fist extension</td>
<td>0.9012</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
results, since some instruments present higher or lower ICC depending on the assessed movement.

With regard to the measure of active ROM of ankle dorsiflexion, Venturini et al.\(^8\), assessing 28 healthy and young (18 to 30 years old) subjects, found a very strong ICC, using both the UG (0.91 and 0.97) and the inclinometer (0.91 and 0.83). In the measures between the two evaluators, the ICC was moderate for the UG (0.72) and strong for the inclinometer (0.83).

When Konor et al.\(^30\) were analyzing ROM of the elbow dorsiflexion in closed kinetic chain (CKC) of 20 healthy young subjects, through three different instruments (UG, digital inclinometer and metrical tape), with unexperienced evaluators (without previous experience), they also found a very strong ICC, both inter-evaluators (0.85) and inter-instruments (0.99) on both sides (left and right).

On the other hand, when Santos et al.\(^31\) analyzed the dorsiflexion movement in 42 subjects, using a similar methodology to the present study, they found moderate correlations both for inter-evaluators (ICC=0.72) and intra-evaluators (ICC=0.55) measures.

Differently from the results of the present study, Van Trijffel et al.\(^32\), using UG, reported lower reliability of the LL ROM measures compared to the HL. They stated that it was harder to find bone marks accurately, making it more difficult to get the perfect alignment of the goniometer. However, some studies using UG showed high reliability for the knee flexion (Currier et al.\(^33\); ICC=0.87) and ankle dorsiflexion (Diamond et al.\(^34\); ICC=0.74 to 0.87).

As to the movement of fist extension, only the study performed by LaStayo and Wheeler\(^25\) was found, which assessed the UG in three measurement ROM types and, similarly to our results, it was also seen strong reliability indices (ICC=0.80), both for intra and inter-tests.

Another study carried out by Carter et al.\(^35\) to analyze the ROM of fist flexion and extension, using UG by two different evaluators, also found similar results to the present study, with an ICC between 0.80 (very strong) and 1.00 (complete), both intra and inter-evaluators.

Like other authors\(^1,2,5,11,25,31,36\), the present study performed consecutive measures in only one session for all the movements. The first examiner did three consecutive measures, with the aim of eliminating effects of viscoelastic adaptation (warming, extensibility, etc.) of the soft tissues, possibly produced by movement repetition, may be indicated as a limitation of the study. However, like other authors\(^1,2,5,11,31,36\), the present study performed consecutive measures, in only session, for all movements, which presented high reliability among them.

**CONCLUSIONS**

Results from this study show that: evaluator's experience did not influence ROM measures of the joints studied; the instruments used (UG and fleximeter) generally present high reliability when performed in the same session of measurements; and use of fleximeter in the physiotherapeutic clinic may be expanded without damage to reliability of measures.

**REFERENCES**