Reliability and accuracy of cirtometry in healthy adults*

Valéria da Silva Caldeira¹, Célia Cristina Duarte Starling², Raquel Rodrigues Britto¹, Jocimar Avelar Martins¹, Rosana Ferreira Sampaio⁴, Verônica Franco Parreira⁵

Abstract

Objective: To determine the intrarater and interrater reliability of cirtometry (measurements of the circumference of the chest and abdomen taken during respiratory movements) as well as its correlation with pulmonary volumes measured by respiratory inductive plethysmography.

Methods: A total of 40 healthy individuals were evaluated. The mean age was 28 years. The measurements were taken in the supine position at three different time points: at rest, at maximal inspiration, and at maximal expiration. Two trained investigators, each of whom was blinded as to the results obtained by the other, performed the measurements. The Friedman test was used to determine intrarater reliability, and the Wilcoxon test, together with the intraclass correlation coefficient, were used to determine interrater reliability. The correlation between the cirtometry measurements and the plethysmography results was obtained using Spearman's correlation coefficient. The level of significance was set at 0.05 for all tests.

Results: Intrarater reliability was satisfactory. Regarding interrater reliability, statistically significant differences (2.8 cm at the most) were found in all sets of measurements. However, through the analysis of the intraclass correlation coefficient, the investigators were found to be responsible only for a small portion of the variability (1.2–5.08%) found among the measurements. When the cirtometry measurements were compared to the volumes measured by respiratory inductive plethysmography, low correlations (range, \( r = 0.170–0.343 \)) were found.

Conclusions: The findings of this study suggest that, although cirtometry is a reliable measurement, it does not accurately measure pulmonary volumes.

Keywords: Reproducibility of results; Lung volume measurements; Plethysmography; Statistics, nonparametric.
Introduction

In recent decades, physical therapy has sought to establish a scientific basis for guiding clinical practice and supporting the choice of interventions. The first step is to develop a clinical evaluation, through quality tests and measurements, that will make it possible to identify the problem, plan the treatment, document its efficacy, and reclaim the scientific credibility of the procedures. In this sense, it is necessary to adopt precise and accurate measurements, thereby improving the informative content and the validity of the inferences.

The terms precision, reproducibility, and reliability are synonymous and indicate the extent to which the measurements of a stable phenomenon – repeated by different individuals, with different instruments, at different times and places – produce results yielding similar and consistent scores.

Accuracy or validity is the degree to which the measurement represents the phenomenon of interest, and it is objectively better evaluated through the comparison with another measurement considered the gold standard.

Cirtometry, or thoraco-abdominal perimeter measurement, consists of a set of measurements of the circumference of the chest and abdomen taken during respiratory movements. Its objective is to evaluate chest expansion in a simple and accessible way, and, to that end, only a tape measure is needed. More recently, cirtometry has had great applicability, and it has been referred to as a parameter of ‘lung expansion’ measurement, in addition to being used with the aim of evaluating other parameters such as ‘chest width’, ‘lung volumes and capacities’, ‘pulmonary compliance’, ‘thoraco-abdominal mechanics’, ‘diaphragmatic function’, ‘muscle work’, and ‘dyspnea’. In the literature, a series of properties have been found to be attributed to cirtometry, although without scientific evidence, several protocols have adopted different postures/anatomical reference points, and formulas/equations have been calculated based on cirtometry.

In 1999, one group of authors studied the effect that a program of stretching of chest muscles has on thoracic mobility, pulmonary function, and dyspnea in patients with chronic obstructive pulmonary disease. The measurements were performed by placing a tape measure around the chest, at the axillary and xiphoid levels, at maximal inspiration and maximal expiration, with the arms of the subject hanging loosely. Recently, other authors have evaluated chest expansion in healthy adults and in patients with Parkinson’s disease using a steel tape measure around the chest. The data were registered as the total perimeter (inspiratory and expiratory).

Since there is as yet no scientific evidence qualifying cirtometry as a precise and accurate measurement, capable of supporting the inferences reported, the objective of the present study was to determine the intrarater reliability of cirtometry through repeated measurements, the interrater reliability of cirtometry through measurements performed by two different investigators, and the concurrent validity of cirtometry for measuring lung volumes through its correlation with respiratory inductive plethysmography.

Methods

The sample comprised individuals who were selected in the community after advertisement of the study by word of mouth, explanatory posters, and appointment by other volunteers. The inclusion criteria used were as follows: being a volunteer, regardless of gender; being between the ages of 21 and 50; having a body mass index (BMI) within the age-appropriate range of reference values; being a nonsmoker; and presenting no signs or symptoms of lung disease. Spirometry was performed using a Vitalograph spirometer (2120; Vitalograph, Buckingham, England) and in accordance with the guidelines established by the Brazilian Thoracic Society. The subjects presented spirometric values within normal parameters. The exclusion criteria were as follows: being unable to perform the measurements proposed, having an osteoarticular or neuromuscular disease that influences respiratory mechanics, and presenting spirometric alterations.

This research project was approved by the Ethics in Research Committee of the Federal University of Minas Gerais, and all subjects gave written informed consent.

After a pilot study involving 24 subjects, a sample of 31 individuals was calculated using a power of 0.95 and a level of significance of 0.01. The calculation was performed using a clinical correlation of 0.80 as a reference.
Respiratory inductive plethysmography, hereinafter referred to only as plethysmography, is an instrument used to monitor the volume and time components of respiratory patterns, as well as the thoraco-abdominal configuration. It is based on the cross-sectional area changes that occur in the compartments of the chest cavity and abdomen.\(^{15-17}\) The method is minimally invasive, does not require a mouthpiece or a nose clip (except for calibration), and there is therefore no need for direct contact with the airway. In the present study, this method was used as the gold standard, and the variable measured was tidal volume, at rest and at maximal inspiration. The accuracy of plethysmography in providing tidal volume values is satisfactory and depends on an appropriate initial calibration as well as on the subject maintaining the same body position.\(^{17,20}\) The same equipment (Respitrace\(^{®}\) 204, NIMS, Miami, FL, USA), which was calibrated using the qualitative diagnostic calibration method, was used in all subjects.\(^{20}\) A detailed description of this procedure has recently been published.\(^{21}\)

In order to perform cirtometry, we used three common tape measures, which were adapted with a strap made of cotton shoelace to serve as a guide for their sliding movement during the respiratory movements and to facilitate the reading. The values were recorded in centimeters.

At the beginning of the collection, a pulse oximeter (J.G. Moriya Indústria Ltda, São Paulo, Brazil) was used to measure peripheral oxygen saturation in order to characterize the sample.

Due to the diversity of procedures found in the literature for performing cirtometry, a protocol procedure was instituted for this study. After the subjects were scheduled to undergo spirometry, they were informed and instructed about the procedures. They then underwent clinical evaluation, which consisted of investigation of demographics, clinical history, and smoking habits. All subjects were weighed and measured, barefoot but with clothing, and their BMI was calculated. The protocol was performed after the subjects had given written informed consent.

Three anatomical reference points, which were marked using a ballpoint pen on the individual body surface, were considered – axillary fold, xiphoïd process, and umbilical line. The three tape measures were placed on the points mentioned above, below the pen mark on the axilla and above the other two pen marks, in order to perform the cirtometry.

The elastic bands of the plethysmograph were placed on the axilla and on the navel area, and were connected to the equipment by wires. The bands and tape measures were positioned so as to avoid folds and the distortions such folds might create in the measurements.

In order to begin taking the measurements, the subject was placed in the supine position with 0° of dorsal inclination using two pillows, one under the head and one under the knees. Peripheral oxygen saturation was measured, and the plethysmograph was calibrated.

The cirtometry measurements were taken at three different time points: at rest; after a deep, slow maximal inspiration up to total lung capacity; and after a slow maximal expiration down to residual volume. The investigator asked the subject to perform the maximal inspiration and maximal expiration maneuvers after having reported the measurement at rest to the computer operator, who recorded the information and identified the measurements by time on the register form, thus allowing their subsequent analysis concomitantly with the plethysmographic tracing. For each anatomical reference point, three measurements were taken, at the three different time points, with one-minute intervals between each. The computer operator informed the investigator, after each minute and according to the stability of the tracing registered and observed by the former, as to whether the next measurement could be taken. The order in which the anatomical reference points were to be measured was randomized at each measurement time point.

Each investigator performed a total of 27 cirtometries per subject. After Investigator 1 took the measurements, Investigator 2 performed the same procedure, again in random order. The investigators were previously trained for the collection. Neither was present during the measurements taken by the other, and both were blinded as to those values.

The statistical analysis of the data was based on the measurements of central tendency and of variability (mean and standard deviation). Since the variation did not present a normal distribution, nonparametric tests were employed. Intrarater reliability was determined using the Friedman test.\(^{22}\) Interrater reliability was determined using the Wilcoxon test and the intraclass correlation coefficient,\(^{22-25}\) the latter of which quantifies the
cycles were transformed into a single measurement (mean), and 720 respiratory cycles at maximal inspiration. Table 1 shows the results related to intrarater reliability. Regarding Investigator 1, a significant difference was found among the three measurements taken in the axillary region (at maximal inspiration) and among the three measurements taken in the xiphoid and umbilical regions (at rest), and there were no differences among the measurements taken in the other situations. Regarding Investigator 2, significant differences were found among the three measurements taken in the axillary region (at rest and at maximal expiration) and among the three measurements taken in the xiphoid region (at maximal expiration), and there were no differences among the three measurements taken in the other situations. For both investigators, it was observed that, where there was a significant difference, the mean maximal difference was always less than 0.5 cm.

Table 2 shows the results related to interrater reliability. A significant difference was detected between the measurements performed by the two investigators in all the regions and at all of the time points tested. The differences among the measurement means ranged from 0.75 cm (umbilical region at maximal inspiration) to 2.8 cm (axillary region at maximal expiration). However, through the analysis of the intraclass correlation coefficient, the investigators were found to be responsible only for a small portion of the variability (1.20-5.08%) found among the measurements.

For the evaluation of the concurrent validity, that is, the correlation between cirtometry and plethysmography, only the measurements performed at rest and at maximal inspiration were used, since plethysmography does not provide lung volume data at maximal expiration.

Initially, 58 subjects were evaluated. Of those, 18 were excluded: 2 for having a BMI above the reference values, 8 for presenting spirometric alterations, and 8 for dropping out. The data presented are related to 40 individuals and are expressed as mean ± standard deviation. The mean age was 28 ± 7 years (range, 21-50 years). A total of 31 subjects (77.5%) were female, and 9 (22.5%) were male. The mean BMI of the group was 21 ± 2 kg/m². Regarding the spirometric data, the following percentage means were observed in comparison with the predicted values for the Brazilian population: forced expiratory volume in one second (FEV₁) = 99 ± 8.45%; forced vital capacity (FVC) = 98 ± 8.61%; and forced expiratory flow between 25 and 75% of FVC = 92 ± 23.53%. The FEV₁/FVC ratio was 87.41 ± 5.87. These results characterize the subjects as normal. Peripheral oxygen saturation ranged from 96 to 99%, and the mean heart rate was 71 ± 10 bpm, both results being within the range considered normal.

The at-rest data represent the mean of three respiratory cycles immediately prior to maximal inspiration. We analyzed 2160 respiratory cycles at rest, totaling 720 measurements, since every three measurements were synthesized into a single value (the sample mean).
spreading their hands on the chest of the patient during respiratory movements. This measurement is part of the physical examination of the chest and makes it possible to assess chest mobility, as well as to detect, in particular, asymmetry between the two hemithoraces. In clinical practice, this measurement is also used as a parameter that provides an idea of lung volumes and, when used in combination with pulmonary auscultation, guides the use of physical therapy techniques and resources. It is important to emphasize the qualitative characteristic of this measurement, and that, therefore, it is

Table 3 shows the correlations between cirtometry and plethysmography based on these means. All of the correlations were found to be low (range, r = 0.170-0.343), and the correlations of the results related to the xiphoid region at rest and to the umbilical region at maximal inspiration proved to be significant.

Discussion

The evaluation of chest expansion has been cited in the literature as a qualitative and subjective measurement that investigators perform by spreading their hands on the chest of the patient during respiratory movements. This measurement is part of the physical examination of the chest and makes it possible to assess chest mobility, as well as to detect, in particular, asymmetry between the two hemithoraces. In clinical practice, this measurement is also used as a parameter that provides an idea of lung volumes and, when used in combination with pulmonary auscultation, guides the use of physical therapy techniques and resources. It is important to emphasize the qualitative characteristic of this measurement, and that, therefore, it is

Table 1 - Intrarater reliability of the measurements of the axillary, xiphoid, and umbilical regions performed at rest, at maximal inspiration, and at maximal expiration.

<table>
<thead>
<tr>
<th>Time point</th>
<th>Measurement</th>
<th>Inv. 1: Mean ± SD (cm)</th>
<th>p</th>
<th>Inv. 2: Mean ± SD (cm)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Axillary region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At rest</td>
<td>1</td>
<td>89.95 ± 9.04</td>
<td>0.187</td>
<td>91.96 ± 9.15</td>
<td>0.034*</td>
</tr>
</tbody>
</table>
not necessarily appropriate for the measurement of lung volumes. The main results of this study were as follows: intrarater and interrater reliability were observed, characterizing cirtometry as an instrument of some precision; and cirtometry was not found to be accurate in measuring lung volumes.

In the analysis of intrarater reliability, in order to minimize the possibility of bias caused by investigator recollection of the first score, the measurements were recorded by a third party, the order was randomized so that the chance of measurements being subsequently taken at the same point was smaller, and the interval between the measurements was used by the investigator to perform the next randomization, which minimized the probability of memorizing the previous measurement. It should be highlighted that the two investigators reported, a posteriori, that they did not remember the values measured.

For the measurements to be considered precise, it is necessary that there be no significant differences among them. In the present study, the results revealed some statistically significant differences, which does not necessarily imply clinical relevance. The differences found were of less than 0.5 cm. Regarding the measurement of chest and abdomen expansion, the variation of 0.5 cm in the cirtometry measurements cannot be considered clinically relevant. Therefore, cirtometry is considered reproducible.

In concurrent validity, the degree to which a measurement correlates with a simultaneous criterion, usually the gold standard, is measured. The

### Table 2 – Interrater reliability of the measurements of the axillary, xiphoid, and umbilical regions performed at rest, at maximal inspiration, and at maximal expiration.

<table>
<thead>
<tr>
<th>Time point</th>
<th>Investigator</th>
<th>Mean ± SD (cm)</th>
<th>p</th>
<th>Intraclass correlation coefficient (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Axillary region</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At rest</td>
<td>1</td>
<td>89.93 ± 8.97</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.02</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>92.01 ± 9.18</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.02</td>
</tr>
<tr>
<td>Maximal inspiration</td>
<td>1</td>
<td>93.36 ± 9.09</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.75</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>94.90 ± 9.42</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.75</td>
</tr>
<tr>
<td>Maximal expiration</td>
<td>1</td>
<td>87.75 ± 8.88</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5.08</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>90.55 ± 9.20</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5.08</td>
</tr>
<tr>
<td><strong>Xiphoid region</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At rest</td>
<td>1</td>
<td>79.97 ± 6.61</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.19</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>81.28 ± 6.54</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.19</td>
</tr>
<tr>
<td>Maximal inspiration</td>
<td>1</td>
<td>83.01 ± 6.25</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.36</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>83.88 ± 6.25</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.36</td>
</tr>
<tr>
<td>Maximal expiration</td>
<td>1</td>
<td>77.95 ± 6.61</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4.26</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>79.88 ± 6.69</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4.26</td>
</tr>
<tr>
<td><strong>Umbilical region</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At rest</td>
<td>1</td>
<td>75.75 ± 6.75</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.86</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>76.85 ± 6.80</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.86</td>
</tr>
<tr>
<td>Maximal inspiration</td>
<td>1</td>
<td>78.54 ± 6.75</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.20</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>79.29 ± 6.81</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.20</td>
</tr>
<tr>
<td>Maximal expiration</td>
<td>1</td>
<td>72.60 ± 7.03</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.48</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>74.22 ± 7.04</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.48</td>
</tr>
</tbody>
</table>

SD: standard deviation; *p < 0.05 using the Wilcoxon test.
Appropriate precautions were taken in an attempt to minimize measurement errors that might bias the results: the measurements were taken after the signal calibration, with the subject in the supine position, and this position was maintained throughout the collection, thus ensuring measurement accuracy; 

The sample, comprised of volunteers, did not have an equivalent number of subjects in terms of gender. There was a predominance of females (77.5%). Since the objective of the study was to evaluate measurements related to the respiratory pattern using two noninvasive methods, this gender imbalance did not constitute a bias and therefore did not compromise the internal validity of the study. In a study conducted in Brazil,

The correlations were found to be low (r = 0.343 at the most), which characterizes cirtometry as an inaccurate method for measuring lung volumes in healthy adults. It could be argued that this result is related to a less than sufficient sample size. However, the number of subjects studied was 29% higher than that established in the sample size calculation as the minimum number required to ensure that significant correlations would be detected by the statistical tests.

In the present study, the supine position was adopted due to the fact that it facilitated the handling of the instruments, as well as to other factors. This posture allowed the positioning and the use of the tape measures alternately and without the need to mobilize the subject during the collection, thereby reducing the use of technical artifacts in the plethysmographic tracing. Since the posture adopted was the same for the two measurement methods, which were used simultaneously, we can rule out the possibility that posture influenced the results.

Table 3 - Correlation between the cirtometry measurements (cm) and the tidal volume measured by respiratory inductive plethysmography (mL) at rest and at maximal inspiration.

<table>
<thead>
<tr>
<th>Region</th>
<th>Time point</th>
<th>Spearman’s correlation (r)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Axillary</td>
<td>At rest</td>
<td>0.211</td>
<td>0.060</td>
</tr>
<tr>
<td></td>
<td>Maximal inspiration</td>
<td>0.201</td>
<td>0.085</td>
</tr>
<tr>
<td>Xiphoid</td>
<td>At rest</td>
<td>0.343</td>
<td>0.002*</td>
</tr>
<tr>
<td></td>
<td>Maximal inspiration</td>
<td>0.170</td>
<td>0.149</td>
</tr>
<tr>
<td>Umbilical</td>
<td>At rest</td>
<td>0.176</td>
<td>0.117</td>
</tr>
<tr>
<td></td>
<td>Maximal inspiration</td>
<td>0.230</td>
<td>0.049*</td>
</tr>
</tbody>
</table>

*p < 0.05 using Spearman’s correlation.

instrument, since it is considered accurate, serves as a parameter for determining the accuracy of a new instrument. Plethysmography was used because it constitutes a reference standard for lung volume measurements. The similarity of the methods was considered, since plethysmography uses bands and cirtometry uses tape measures around the chest and abdomen. In addition, neither of the methods is invasive or uses mouthpieces/nose clips, which minimizes errors in the comparison.

The author observed that males and females maintain the same behavior in terms of respiratory pattern and thoraco-abdominal configuration in the dorsal and sitting positions. One group of authors evaluated the respiratory pattern of 120 healthy subjects using plethysmography and found no significant gender-related differences that would interfere with the respiratory pattern at rest.

However, in the present study, the correlational analyses of the measurements, in terms of intrarater and interrater reliability, were made individually for each of the subjects. The analysis of the correlation between cirtometry and plethysmography was performed by comparing the measurements obtained by the first instrument with the measurements obtained by the second instrument, although always in relation to the same subject. The gender of the subject, as any other individual characteristic in this case, would not interfere with the results.

The results of the present study regarding the psychometric properties of cirtometry demonstrate that intrarater reliability, analyzed using repeated measurements, and interrater reliability, analyzed using measurements performed by two investigators who were trained and blinded, are satisfactory. Nevertheless, no concurrent validity was found.
between cirtometry and plethysmography, which demonstrates that cirtometry is not accurate for measuring lung volumes, despite having proved to be a precise instrument, suggesting that cirtometry should not be used for such a purpose in clinical practice or in scientific research.

Acknowledgments

The present study received financial support from the Brazilian Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, National Council for Scientific and Technological Development). We would like to extend our sincere thanks to Daniela Silveira, Juliana Rodrigues, and Letícia Dias, all of whom are recipients of grants from the CNPq Programa Institucional de Bolsas de Iniciação Científica (PIBIC, Institutional Program for Scientific Initiation Scholarships), for their contributions to the completion of this study.

References

25. Carrasco J, Jover J. The concordance correlation coefficient estimated through variance components. IX Conferencia Española de Biometría; 2003 May 28-30; La Coruña, Spain.