How to cite this article: Martinez BP, Forgiarini Júnior LA, Camelier FRW, Alves GAA, Gomes Neto M, Soeiro SE. Viability and reliability among evaluators for vital capacity measured by ventilometer in healthy individuals. J. Physiother. Res. 2019;9(1):10-17. doi: 10.17267/2238-2704rpf.v9i1.2078

Original article

Journals BAHIANA
SCHOOL OF MEDICINE AND PUBLIC HEALTH

Viability and reliability among evaluators for vital capacity measured by ventilometer in healthy individuals

Viabilidade e confiabilidade entre avaliadores para capacidade vital medida por ventilômetro em indivíduos saudáveis

Bruno Prata Martinez[1], Luiz Alberto Forgiarini Júnior[2], Fernanda Rosa Warken Camelier[3], Giovani Assunção de Azevedo Alves[4], Mansueto Gomes Neto[5], Santuzza Evrágio Soeiro[6]

[1]Corresponding author. State University of Bahia, Federal University of Bahia. Salvador, Bahia, Brazil. ORCID: 0000-0002-4673-8698. brunopmartinez@hotmail.com
[2]Methodist University Center – IPA. Porto Alegre, Rio Grande do Sul, Brazil. ORCID: 0000-0002-6706-2703. forgiarini.luiz@gmail.com
[3]State University of Bahia. Salvador, Bahia, Brazil. ORCID: 0000-0003-2540-0142. fcamelier@uneb.br
[4]Catholic University of Salvador, Aliança Hospital. Salvador, Bahia, Brazil. ORCID: 0000-0002-5332-8574. giovani.alves@pro.ucsal.br
[5]Federal University of Bahia. Salvador, Bahia, Brazil. ORCID: 0000-0002-0717-9694. mansueto.neto@ufba.br
[6]Social Faculty of Bahia. Salvador, Bahia, Brazil. ORCID: 0000-0001-6650-8729. san_soeirog@hotmail.com

RESUMO | INTRODUÇÃO: A avaliação da capacidade vital (CV) é um método de avaliação respiratória relevante, visto que seus dados contribuem para o diagnóstico de doenças e alterações. **OBJETIVO:** Avaliar a viabilidade e confiabilidade relativa e absoluta da CV mensurada através do ventilômetro. **MATERIAIS E MÉTODOS:** Estudo transversal realizado entre setembro a outubro de 2010, onde foram incluídos indivíduos saudáveis em um hospital na cidade de Salvador/BA. As mensurações da CV foram realizadas em três momentos através do ventilômetro, sendo realizadas por dois avaliadores distintos (A e B). A primeira medida foi avaliada pelo examinador A (A1), a segunda pelo examinador B e a terceira novamente pelo examinador A (A2). Para confiabilidade inter-examinador utilizou-se as medidas A1 x B e B x A2 e para intra-examinador A1 x A2. Para análise estatística foi realizada o coeficiente de correlação intra-classe (CCI) para confiabilidade relativa e a análise de Bland-Altmann para confiabilidade absoluta. **RESULTADOS:** A amostra foi composta por 30 indivíduos, com idade média de 29,4 ± 6,0 anos. Não houve diferença na comparação das médias (A1 x B, p=0,55; B x A2, p=0,62 e A1 x A2, p=0,40). A confiabilidade relativa intra-examinador foi 0,97 (p=0,0001) e as inter-examinadores 0,87 (p=0,0001) e 0,97 (p=0,0001). A confiabilidade absoluta apresentou concordância, porém com viés variável (- 0,09; - 0,05 e - 0,03). **CONCLUSÕES:** Houve uma alta confiabilidade relativa e moderada confiabilidade absoluta da capacidade vital aferida através do ventilômetro.

PALAVRAS-CHAVE: Capacidade vital. Reprodutibilidade dos testes. Testes de função respiratória.

ABSTRACT | INTRODUCTION: Vital capacity (VC) assessment is a relevant respiratory assessment method, since its data contribute to the diagnosis of diseases and alterations of this system. **OBJECTIVES:** The aim of this study was to evaluate the feasibility and the relative and absolute reliability of CV measured by the ventilometer. **MATERIALS AND METHODS:** A cross-sectional study was conducted between September and October 2010, where healthy individuals were included in a hospital in the city of Salvador / BA. CV measurements were performed in the three moments through the ventilometer, being performed by two different raters (A and B). The first measure was reviewed by the observer A (A1), the second by observer B and the third again by observer A (A2). For inter-rater reliability, we used the measures A1 x B and B x A2 and intra-rater A1 x A2. **RESULTS:** The sample was composed of 30 subjects, mean age 29.4 ± 6.0 years. There was no difference in the comparison of means (A1 x B, p = 0.55, B x A2, p = 0.62 and A1 x A2, p = 0.40). The intra-rater reliability relative was 0.97 (p = 0.0001) and inter-rater 0.87 (p = 0.0001) and 0.97 (p = 0.0001). The absolute reliability showed agreement, but with variable bias (- 0.09, - 0.05 and - 0.03). **CONCLUSION:** There was a relative high reliability and moderate absolute reliability of vital capacity measured by the ventilometer.

KEYWORDS: Vital capacity. Reproducibility of results. Respiratory function tests.

Submitted 08/26/2018, Accepted 02/07/2019, Published 02/19/2019
J. Physiother. Res., Salvador, 2019 February;9(1):10-17
Doi: 10.17267/2238-2704rpf.v9i1.2078 | ISSN: 2238-2704
Responsible editor: Cristiane Dias

Introduction

Pulmonary function tests allow the identification of the alterations associated with diseases of the respiratory system, besides measuring the evolution of the disease and response to the treatment, in this way they collaborate for an adequate orientation in the clinical practice and allow the choice of the most effective physiotherapeutic intervention[1-3]. The vital capacity (VC) represents the largest volume of air mobilized between full inspiration and complete expiration[4]. The VC evaluation is used as an important test of respiratory function and its reduction is an abnormality present in people with respiratory muscle weakness and/or changes in respiratory function, such as restrictive or obstructive respiratory disorders[5].

Several studies have used the VC test to evaluate respiratory function with the spirometer in the pre- and postoperative period of abdominal surgeries[6-8], heart[3,9] and thoracic[5,10], in clinical patients with renal insufficiency, as well as application in ventilatory weaning in patients with neuromuscular disease[11]. However, there is a lack of studies on the inter-rater reliability and measurement of VC with ventilometer[12,13]. For continuous variables such as VC, relative and absolute reliability should be used. Relative reliability is associated with the linear relationship between measurements. The absolute measure of repeated measures between examiners is more effective in demonstrating the precision between two measures[12].

To know the reliability and accuracy of this instrument is fundamental for the respiratory evaluation in physiotherapeutic practice, since it is an evaluation tool widely used in clinical practice. Thus, the objective of the study was to evaluate the relative and absolute reliability and viability of the VC measured through the ventilometer.

Methods

An analytical study was performed with healthy individuals, through a random selection of employees of the Santo Antônio Hospital/Obras Sociais Irmã Dulce. We evaluated and included individuals with a body mass index between 18.9 and 24.9 kg/m^2, healthy and aged over 18 years. Those who were unable to perform the proper technique due to lack of understanding were excluded, as were those who were smokers. The study followed the regulations of the statement for observational studies (STROBE)[13].

All participants were informed about the objectives of the study and signed the informed consent form. The research project was approved by the Research Ethics Committee of the Santo Antônio Hospital/ Obras Sociais Irmã Dulce, under the identification number 45/09 (CAAE 0044.0.058.000-09). The VC measurement was performed with an analog ventilator Mark Wright 8 Ferraris (Louisville, CO, USA) and coupled to the expiratory branch of a one-way valve and silicone facial mask. All subjects were seated in a chair and instructed on the procedure to be performed. Subjects were instructed to perform maximal inspiration until total lung capacity (TLC), followed by expiration close to the residual volume (RV) to obtain VC values. This expiration was encouraged by the researcher so that the individual reached his maximum capacity[3,4,14]. A maximum of six measures were completed and finalized when the three consecutive measures were less than 5% of the difference between them. The largest value in liters (L) was considered for the analysis and between one and another measure the one-minute interval was used[4]. All the evaluators were trained to measure VC.

VC measurements were performed at three moments, and were performed by two different evaluators(A and B). The first measure was evaluated by the examiner A(A1), the second by the examiner B, and the third by the examiner A(A2). This interrater reliability was achieved between B and A1 x A2 x of B, and intra-examiner reliability was obtained between A1 x A2. There was no contact between the examiners in the interval between measurements, to avoid any bias in the data. The interval between these three moments was thirty minutes.

The descriptive analysis was performed using means and standard deviations, standard errors and confidence intervals. Because of the parametric distribution of the mean VC, the paired Student's t test was used to compare the means of the VC of the evaluator A (A1 and A2) and B. To determine the

relative reliability of the interviewers (VC measure by the assessor in the first moment) and the second time for the evaluator B) and the intra-evaluator (measure of the VC of the evaluator in the first and second moments), the Intraclass Correlation Coefficient (ICC) was used[15-18].

Reliability was classified as small (<0.25), low (0.26-0.49), moderate (0.50-0.69), high (0.70-0.89), and very high (> 0.90), according to the reference values described by Gross and Domholdt[15]. The absolute reliability was analyzed by the method of Bland-Altmann[19-21], which verifies the occurrence of systematic or random changes in the mean values of intra and interobserver VC. The data were analyzed in the software SPSS (Statistical Package for the Social Sciences) version 14.0, being considered a level of significance of 5% in the software.

Results

The sample consisted of 30 individuals, predominantly male (70%), with a mean age of 29.4 ± 6.0 years. The mean VC of the patients was 4.74 ± 1.27 L (A1 evaluator); 4.83 ± 1.15 L (evaluator B) and 4.86 ± 1.21 L (evaluator A2). Comparing the values of the inter-rater and intra-rater evaluation measures, no differences were found (p = 0.55 and p = 0.40, respectively).

The value obtained from the intraclass correlation coefficient for intra-examiner reliability was 0.97 (p = 0.0001). The inter-rater analysis between A1 and B was 0.87 (p = 0.0001) and between A2 and B was 0.97 (p = 0.0001); indicating a high reliability of VC in healthy individuals (Figure 1 and 2).

Figure 1. Intraobserver Relative Reliability (Evaluator A) for the measurement of vital capacity (CV). (n=30)

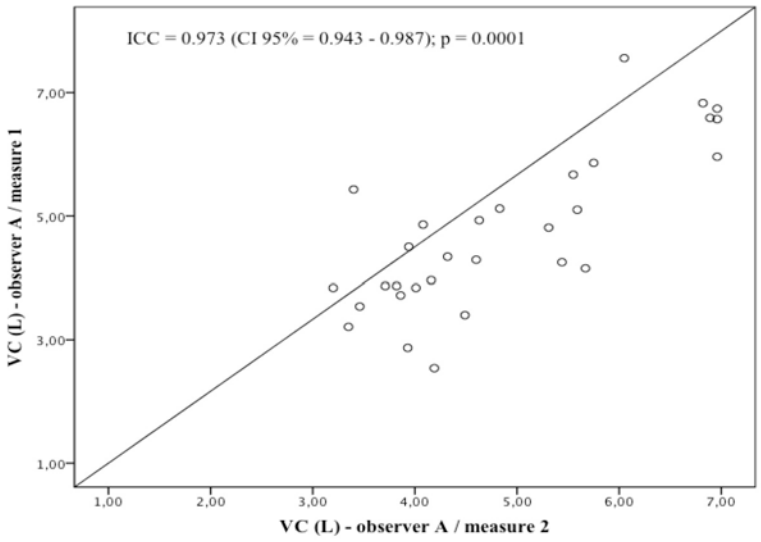
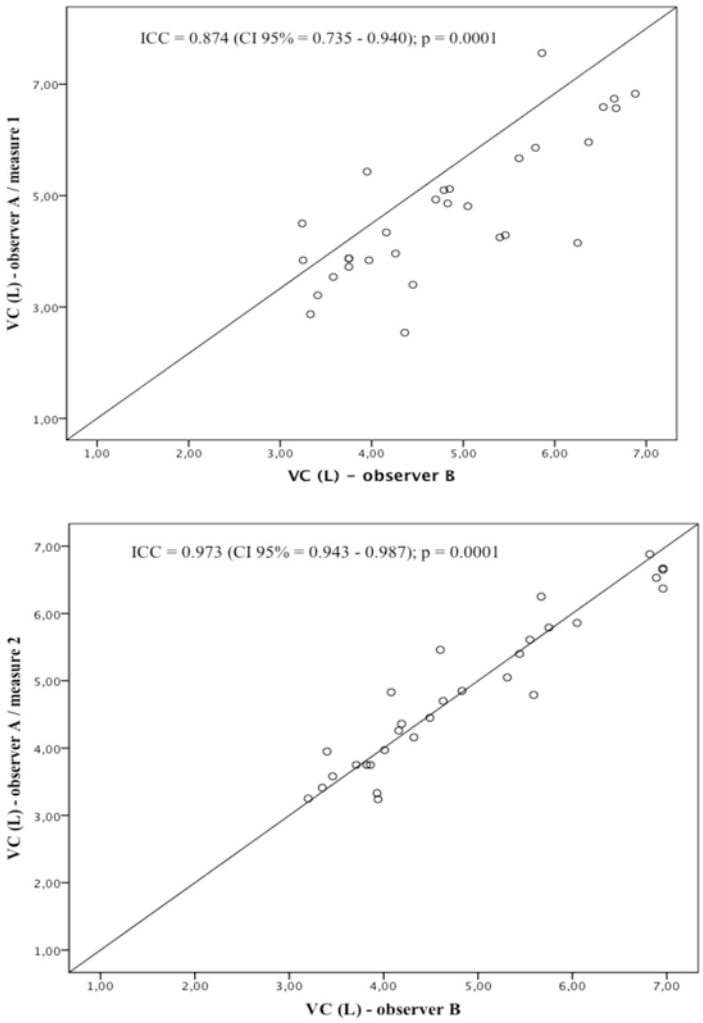


Figure 2. Interobserver relative reliability (Evaluator A, first measure x Evaluator B and Evaluator B x Evaluator A, second measure) for the measurement of vital capacity (CV). (n = 30)



The Bland-Altmann method showed that the inter- and intra-examiner absolute reliability of the VC measure showed agreement, but with a variable bias (A1 x B = -0.05, A1 x A2 = 0.09 and B x A2 = - 0.03, respectively) (Figures 3, 4 and 5). The limits of agreement were -1.71 to 1.50 for A1 x B; -1.61 to 1.50 for A1 x A2 and -0.79 to 0.72 for B x A2.

Figure 3. Individual variability of vital capacity as measured by the Evaluator (A) by Bland & Altman. The horizontal lines show the mean bias (- 0.05 ± 0.79 L) and the agreement limits (- 1.61 to 1.50 L). (n = 30)

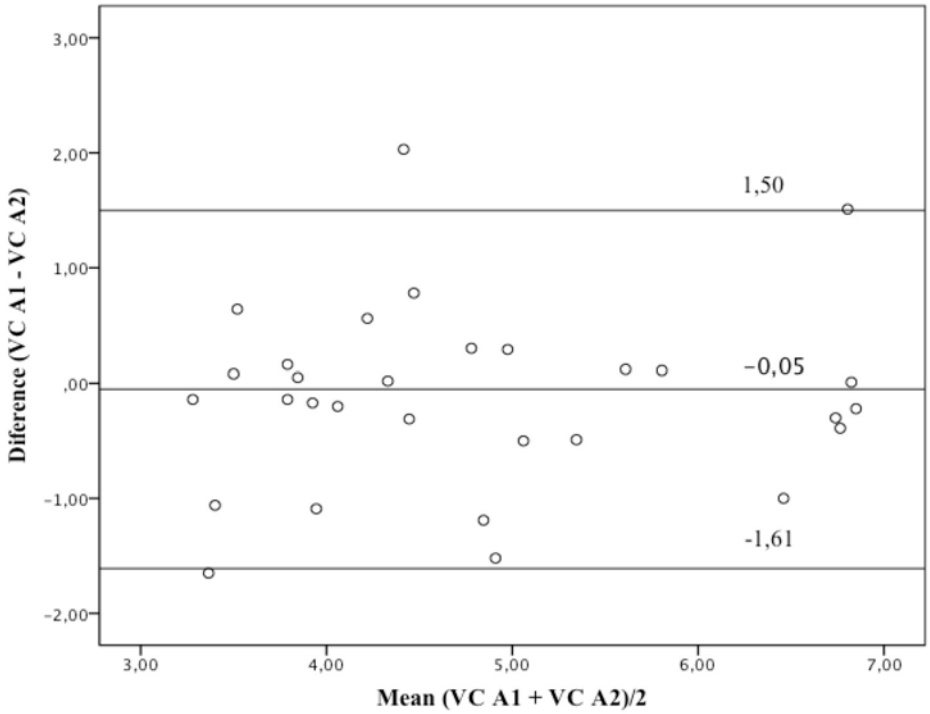


Figure 4. Individual variability of vital capacity in the inter-observer evaluation, by Bland & Altman. The horizontal lines show the mean bias (- 0.09 ± 0.81 L) and agreement limits (-1.70 to 1.50 L), (n = 30 individuals)

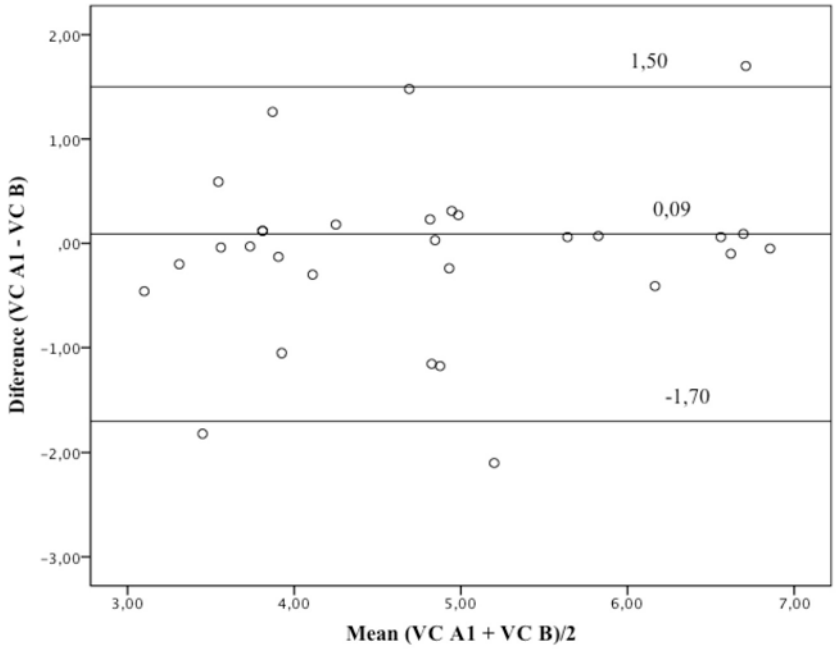
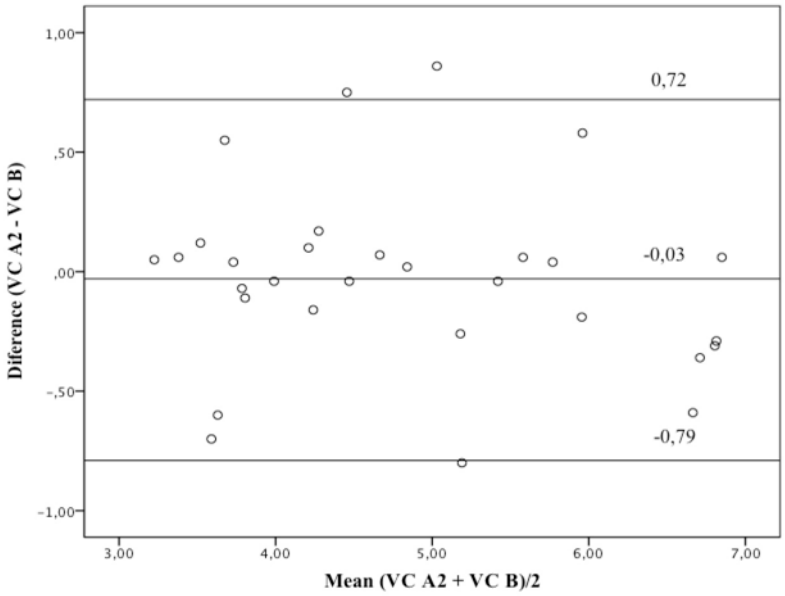


Figure 5. Individual variability measure vital capacity (VC) in assessing inter-observer, by Bland & Altman. The horizontal lines show the mean bias (- 0.03 ± 0.38 L) and limits of agreement (- 0.79 to 0.72 L). (n = 30 individuals)



Discussion

The study demonstrated high reliability in intra-examiner and inter-examiner measurement of VC. This is the first study to evaluate the relative and absolute reliability of VC measurement using the ventilometer. These results are relevant for clinical practice, since VC is an important parameter of pulmonary function, being useful to support the physiotherapeutic diagnosis, to direct the treatment, as well as to follow the development of diseases with respiratory function impairment.

Although there is a high linear association and there is no significant difference in the comparison of the mean values of the vital capacity, attention should be paid to the reliability of the VC. For a high correlation does not necessarily indicate a high convergence between the measures performed[22]. Thus, for an absolute and intraobserver reliability evaluation, the Bland-Altmann method was used, which evaluates the agreement and bias between the values obtained by the different evaluators and the same examiner at different moments[19-21].

The Bland-Altmann analysis[19-21] showed that there was concordance between intra- and inter-examiner measures, since most of the intersections between bias and mean values were within the limits of agreement. However, the limits of agreement (1.61 to 1.50, - 0.79 to 0.72 and - 1.70 to 1.50 liters) were clinically elevated, since they represented almost 30% of the mean VC values.

The mean bias was low among both assessors (A1 x B = - 90 ml x A2 and B = - 30 ml), and between the same evaluator (A1 x A2 = - 50 ml), since it was less than 1 % of mean values of vital capacity. However, this bias was relatively high in the interobserver (A1 x B and B x A2) and intraobserver analyzes, since 22% had differences greater than 500 ml. These differences can be explained by failures in standardization or factors related to learning during the measurement by the evaluators, even though they were previously trained evaluators.

This is because sixteen and eighteen measures among thirty obtained by different evaluators (A1 x B and B x A2, respectively), and twelve of the same evaluators were less than 5% of the mean value of the differences in vital capacity (240 ml), being considered reproducible according to the Brazilian Society of Pulmonology[4]. Learning is a factor that may have influenced the measurements, since the correlation level was higher and the lower mean bias in the second interobserver review (B x A2), for initial interobserver evaluation (A1 x B).

Several factors may influence VC measures, such as the evaluator's understanding and motivation, problems related to calibration of the instrument and inadequate implementation of the technical evaluation[23], directly influencing the reliability of the measure, not generating real information about the patient's respiratory condition and can lead to treatment. In order to have greater confidence in the data obtained by the VC measurement, all the evaluators were trained to develop the technique in a standardized way[24], however, it is possible that there may have been standardization failures.

This study presents some limitations, because although the evaluators had been training for two weeks, the reliability also depends on the evaluator and his experience with the technique[25], which may have influenced the main differences between the mean values of VC in 22% and the measurements.

Conclusion

Therefore, we conclude that there is high reliability in intra-examiner and inter-examiner reliability, as well as moderate absolute reliability of the VC measured by the ventilometer. Despite the low bias among the measures, a small number of acceptable measures did not present differences, suggesting the need for greater standardization and continuous training of the evaluators.

Author contribution

Martinez BP participated in the design and design of the study, statistical analysis of the research data, interpretation of the results and writing of the scientific article. Soeiro SE participated in the design and delineation of the study, collection of research data and writing of the scientific article. Gomes Neto MG, Camelier FWR, Alves GAA, Forgiarini Junior LA participated in the statistical analysis of the research data, interpretation of the results and writing of the scientific article.

Competing interests

No financial, legal or political competing interests with third parties (government, commercial, private foundation, etc.) were disclosed for any aspect of the submitted work (including but not limited to grants, data monitoring board, study design, manuscript preparation, statistical analysis, etc.).

References

1. Caldeira VS, Starling CCD, Britto RR, Martins JA, Sampaio RF, Parreira VF. Precisão e Acurácia da cirtometria em adulto saudáveis. J Bras Pneumol. 2007;33(5):519-26. doi: 10.1590/S1806-37132007000500006

2. Freitas CG, Pereira CAC, Viegas CAA. Capacidade inspiratória, limitação ao exercício, e preditores de gravidade e prognóstico, em doença pulmonar obstrutiva crônica. J Bras Pneumol. 2007;33(4):389-96. doi: 10.1590/S1806-37132007000400007

3. Paisani DM, Chivegato LD, Faresin SM. Volumes, capacidades pulmonares e força muscular respiratória no pós-operatório de gastroplastia. J Bras Pneumol. 2005;31(2):125-32. doi: 10.1590/S1806-37132005000200007

4. Sociedade Brasileira de Pneumologia e Tisiologia. Diretrizes para Testes de Função Pulmonar. J Pneumol. 2002;28(Suppl 3):S1-S238

5. Gibson J, Whitelaw W, Siafakas N. Tests of overall respiratory function. Am J Respir Crit Care Med. 2002;166(4):521-7.

6. Barbalho-Moulin MC, Miguel GPS, Forti EMP, Campos FA, Costa D. Effects of preoperative inspiratory muscle training in obese womem undergoing open bariatric surgery: respiratory muscle, lung volumes, and diaphragmatic excursion. Clinics. 2011;66(10):1721-27.

7. Ribeiro S, Gastaldi AC, Fernandes C. Efeito da cinesioterapia respiratória em pacientes submetidos à cirurgia abdominal alta. Einstein. 2008;6(2):166-9.

8. Pessoa KC, Araújo GF, Pinheiro AN, Ramos MRS, Maia SC. Ventilação não-invasiva no pós-operatório imediato de derivação gastrojejunal com bypass em Y de Roux. Rev Bras Fisiot. 2010;14(4):290-5. doi: 10.1590/S1413-35552010005000023

9. Matheus GB, Dragosavac D, Trevisan P, Costa CE, Lopes MM, Ribeiro GCA. Treinamento muscular melhora o volume corrente e a capacidade vital no pós-operatório de revascularização do miocárdio. Rev Bras Cir Cardiovascular. 2012;27(3):362-9. doi: 10.5935/1678-9741.20120063

10. Lima VP, Bomfim D, Risso TT, Paisani DM, Fiore Júnior JF, Chiavegato LD et al. Influência do dreno pleural sobre a dor. Capacidade vital e teste de caminhada de seis minutos em pacientes submetidos à ressecção pulmon. J Bras Pneumol. 2008;34(12):1003-07. doi: 10.1590/S1806-37132008001200004

11. Bach JR, Gonçalves MR, Hamdani I, Winck JC. Extubation of Patients With 12-Neuromuscular Weakness: A New Management Paradigm. Chest. 2010;137(5):1033-39. doi: 10.1378/chest.09-2144

12. Bruton A, Conway JH, Holgate ST. Reliability: what is it. And how is it measured? Physiotherapy. 2000;86(2):94-9. doi: 10.1016/S0031-9406(05)61211-4

13. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. PLoS Med. 2007;4(10):e297. doi: 10.1371/journal.pmed.0040297

14. American Thoracic Society/European Respiratory Society. ATS/ERS statement on respiratory muscle testing. Am J Respir Crit Care Med. 2002;166(4):518-624. doi: 10.1164/rccm.166.4.518

15. Jhonson SR, Gross MT. Intraexaminer Reability, Interexaminer Reability, and Mean Values for Nine Lower extremity Skeletal Measures in Healthy Naval Midshipmen. J Orthop Sports Phys Ther. 1997;(25):253-63. doi: 10.2519/jospt.1997.25.4.253

16. Weir JP. Quantifying test-retes realibility using the intraclass correlation coeficient and the SEM. J Strength Cond Res. 2005;19(1):231-40.

17. Rankin G, Stokes M. Realibility of assessment tools in rehabilitation:an illustration of appropriate statistical analyses. Clin Rehabil. 1998;12(3):187-99. doi: 10.1191/026921598672178340

18. Lexell JE, Downham DY. How to assess the reability of measurements in rehabilitation. Am J Phys Med Rehabil. 2005;84(9):719-23.

19. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. Statistician. 1983;32:307-17.

20. Altman DG, Bland JM. Measuring agreement in method comparison studies. Statistical Methods in Medical Research. 1999;8(2):135-60. doi: 10.1177/096228029900800204

21. Sedgwick P. Limits of agreement (Bland-Altman method). BMJ. 2013;346:f1630. doi: 10.1136/bmj.f1630

22. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. Phys Ther. 2006;86(5):735-43.

23. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A et al. Standardisation of spirometry. Eur Respir J. 2005;26(2):319-38. doi: 10.1183/09031936.05.00034805

24. Steier J, Kaul S, Seymour J, Jolley C, Rafferty GF, Man WD-C et al. The value of multiple tests of respiratory muscle strength. Thorax. 2007;62(11):975-80. doi: 10.1136/thx.2006.072884