DOI: 10.1590/1809-2950/18007825042018

ORIGINAL RESEARCH

Reliability in the measurement of maximum inspiratory pressure and inspiratory capacity of a physiotherapist in training

Confiabilidad en la medición de la presión inspiratoria máxima y de la capacidad inspiratoria de un fisioterapeuta en entrenamiento

Confiabilidade na mensuração da pressão inspiratória máxima e da capacidade inspiratória de um fisioterapeuta em treinamento

Rodrigo Muñoz Cofré[1,2], Mariano del Sol Calderón[1], Paul Medina González[3], Nicolás Martínez Saavedra[4], Máximo Escobar Cabello[2,3,5]

ABSTRACT | The objective of this study was to explore the impact of clinical experience on the reliability and concordance of maximal inspiratory pressure (MIP) and inspiratory capacity (IC) measurements in a period of clinical training. For convenience, 37 participants in a body plethysmograph were evaluated by an experienced physiotherapist (EF) and a novice physiotherapist (NF). Intra-Class Correlation Coefficient (ICC) was used to analyze the reliability of the MIP and IC tests; to explore the individual differences, the Bland-Altman (gB/A) graphs were used. ICC analysis in three trials showed excellent inter-rater reliability (ICC 1st: 0.914; ICC 2nd: 0.915; ICC 3rd: 0.925) for the MIP test and (ICC 1st: 0.955; ICC 2nd: 0.965; ICC 3rd: 0.970) for the IC test. However, concordance according to gB/A among the evaluators showed a systematic trend with higher absolute scores for EF of 9.2 cmH2O in MIP, and of 0.06 L in IC, respectively. The results suggest that NF acquired reliable technical and discriminative skills for the MIP and IC test, but patients tended to improve performance with an experienced assessor. The evaluator's experience influences the results obtained from the measurement of the MIP in the subjects; the formation of a NF requires incorporating more skills to recognize a sincere and maximum effort.

Keywords | Maximal Respiratory Pressures; Data Accuracy; Physical Therapists.

RESUMEN | Este estudio pretende explorar el impacto de la experiencia clínica en la fiabilidad y consistencia de la medición de la presión inspiratoria máxima (PIM) y la capacidad inspiratoria (CI) durante el período de la entrenamiento clínico. Los 37 participantes fueron evaluados por un fisioterapeuta especializado (FE) y un fisioterapeuta novato (FN), por medio de un pletismógrafo corporal. Se utilizó el coeficiente de correlación intraclase (ICC, en inglés) para analizar la fiabilidad de las pruebas PIM y CI, mientas que para explorar las diferencias individuales se utilizaron los gráficos de Bland-Altman (gB/A). El análisis ICC en tres estudios clínicos demostró excelente fiabilidad interevaluadores (ICC 1°: 0,914; ICC 2°: 0,915; ICC 3°: 0,925) para la prueba PIM y (ICC 1 °: 0,955; ICC 2°: 0,965; ICC 3°: 0,970) para la prueba CI. Sin embargo, la correlación según gB/A entre los evaluadores reveló una tendencia sistemática con resultados absolutos más elevados para FE de 9,2 cmH2O en PIM y 0,06 L en CI, respectivamente. Los resultados demostraron que el FN tuvo habilidades técnicas y de discernimiento fiables en la prueba PIM y CI, pero los pacientes suelen mejorar el rendimiento con un evaluador experimentado. La experiencia del evaluador influye en los resultados obtenidos de la medición de PIM en los pacientes, la formación de un FN requiere la

Study performed on the Universidad Católica del Maule, Talca, Chile.
[1]Programa de Doctorado en Ciencias Morfológicas, Universidad de La Frontera, Temuco, Chile.
[2]Laboratorio de Función Ventilatoria, Departamento de Kinesiología, Universidad Católica del Maule, Talca, Chile.
[3]Departamento de Kinesiología. Facultad de Ciencias de la Salud, Universidad Católica del Maule, Talca, Chile.
[4]Estudiante, Escuela de Kinesiología, Facultad de Ciencias de la Salud, Universidad Católica del Maule, Talca, Chile.
[5]Programa de Doctorado en Educación en Red, Universidad del Bío-Bío, Chillán, Chile.

Corresponding address: Máximo Escobar Cabello, Universidad Católica del Maule, Avda. San Miguel 3605, Talca, Chile – Zip Code: 3460000 – Phone: +56992787603 – E-mail: maxfescobar@gmail.com.cl – Financing source: Nothing to declare – Conflict of interests: Nothing to declare – Presentation: Mar. 30th, 2018 – Accepted for publication: Sept. 13th, 2018 – Approved by the Human Research Ethics Committee of the Universidad Católica del Maule under Protocol no. 23/2016.

incorporación de más habilidades para que se reconozca su verdadero esfuerzo.
Palabras clave| Presiones Respiratorias Máximas; Exactitud de los Datos; Fisioterapeutas.

RESUMO | O objetivo deste estudo foi explorar o impacto da experiência clínica na confiabilidade e concordância da medição da pressão inspiratória máxima (PIM) e da capacidade inspiratória (CI) em um período de treinamento clínico. Por conveniência, 37 participantes foram avaliados em um pletismógrafo corporal por um fisioterapeuta especializado (FE) e um fisioterapeuta novato (FN). O Coeficiente de Correlação Intraclasse (CCI) foi utilizado para analisar a confiabilidade dos testes PIM e CI; enquanto para explorar as diferenças individuais foram usados os gráficos de Bland-Altman (gB/A). A análise CCI em três ensaios mostrou excelente confiabilidade inter-avaliadores (CCI 1º: 0,914; CCI 2º: 0,915; CCI 3º: 0,925) para o teste PIM e (CCI 1º: 0,955; CCI 2º: 0,965; CCI 3º: 0,970) para o teste de CI. No entanto, a concordância de acordo com gB/A entre os avaliadores, mostrou uma tendência sistemática com resultados absolutos mais altos para FE de 9,2 cmH2O em PIM e 0,06 L em CI, respectivamente. Os resultados sugerem que a FN adquiriu habilidades técnicas e discriminativas confiáveis para o teste PIM e CI, mas os pacientes tendem a melhorar o desempenho com um avaliador experiente. A experiência do avaliador influencia os resultados obtidos a partir da medição do PIM nos sujeitos, a formação de um FN exige a incorporação de mais habilidades para reconhecer um verdadeiro esforço.
Descritores | Pressões Respiratórias Máximas; Acurácia dos Dados; Fisioterapeutas.

INTRODUCTION

Measurement is part of the daily exercise of professionals and such results are used to make decisions that determine important considerations concerning diagnosis, prognosis, and intervention[1]. Sometimes, the actual capacity of the instrument to evaluate a study condition is unknown, and values obtained do not necessarily represent the event to be measured or the expression in all its complexity[2]. Although technological development has managed to decrease the uncertainty of results, measurement as a phenomenon of learning requires carefully refinement of the reliability of its actions. Thus, reliability is the degree to which data is free of error, to provide stability and precision to observations[3], and it can be measured through internal consistency, the application of a test-retest, or the application of the same test to a person, but having two different operators (inter-evaluator reliability). Correlating scores between two evaluators, then, enables observing to what extent it is possible to correct the randomness factor[4]; and, according to Bunogamba et al.[5] and Coté et al.[6], the most appropriate statistic test for reliability analysis is Intra-Class Correlation Coefficient (ICC), because this data analysis is associated with certain time interval. Part of the error reduction can be induced with delimitation or standardization of the procedure to ensure low variability in its measurements[7].

The first approaches in learning standardized procedures are observational and conducted under controlled conditions of much training. Subsequently, there is training on responses that are selected or built so the apprentice measure and record the response in quality and quantity to then increase its reliability. Thus, not always declared, the acquisition of clinical habilitation is the most common of the learning objectives in the education of the physiotherapist[8].

There are few previous studies on the impact of evaluator experience on results of the test of maximum inspiratory pressure (MIP) and inspiratory capacity (IC). In the first case, it is known that one of the reasons for the high variability of MIP demonstrated by Black and Hyatt[9] is the effect of learning[10,11]. Thus, if this important source of systematic error is disregarded[12], there is greater possibility of reporting non-existent increase or decrease[13]. The same situation can be observed in the absence of warm-up of inspiratory muscles[14]. With greater difficulty, to this context is also added that MIP results are influenced by the subjects' sex, age, and cognitive abilities[15,16]. Despite protocol compliance aided by quality criterion, the actual maximum value for the subject is not necessarily obtained[17]. As well as the maximum ventilatory pressures that are especially sensitive to the level of effort exerted[18]. While the IC is defined as the measurement that allows knowing the maximum volume of air that can be inspired from expiratory rest position and includes the tidal volume plus the inspiratory reserve volume[19]. It represents a procedure that shows excellent clinimetrics[20], but with limited contribution[21], although it is considered a better indicator of chronic obstructive disorders. There is not enough evidence about the inter-evaluator applicability of this measurement and only some studies report small and specific populations[22]. In this context, a novice physiotherapist (NP) must not only demonstrate an optimal operation of the equipment and obtain quality results in patients, but also – and on a framework of competency-based training – reach high levels of reproducibility in regulated times, which depend on the structure of the curricular matrix[23].

Considering that in ideal contexts the evaluators should be experienced clinicians who are familiar with the measurements[25], it is frequent, more practical, and less costly to incorporate novice evaluators[23]. Thus, competent performance of a NP will require dedicating time, as well as having a model of individual clinical education so as to integrate basic skills, testing protocols, and analysis based on results obtained[26]. However, the specificity of the conditions that can ensure efficient training for a NP is unknown. Nevertheless, some researchers point out that they remain enhancing skills, such as optimizing the relationship with the patient and the commitment with the specific test, so the results represent the actual values of the patient[29].

Thus, the hypothesis of this research is that the acquisition of reliability in measurements of MIP and IC in a period of clinical training is insufficient. In this respect its confirmation or rejection would establish initial guidelines for the training of physiotherapists and creation of professional practice standards in the area of diagnosis of ventilatory function and dysfunction[25]. In this context, this research aimed to determine the reliability and consistency level of the MIP and IC measurement of a NP versus an expert physiotherapist (EP) in a clinical training period.

METHODOLOGY

Cross-sectional exploratory study carried out between July and August 2016 in the Laboratory of Ventilatory Function-Dysfunction of the Catholic University of Maule (UCM), Chile.

Participants

We selected by convenience through non-probability sampling 37 subjects of the Maule Region, given that the dispersion of the behavior of the variables in studies with similar design is not known[27]. Participants aged over 18 years, with no clinical evidence of acute respiratory disease, and with normal spirometry values qualified for the study[21]. We excluded subjects with smoking habits, morphological alterations of the thorax or spinal column, or Body Mass Index (BMI) greater than 30 Kg/m^2. All participants signed an informed consent form previously endorsed by the Committee of Ethics of the UCM (23/2016).

Design

Induction period: The NP was supervised during a clinical training period of 8 weeks, 44 hours each with exclusive dedication (March, April 2016). In the first week, ten patients were evaluated by the NP in order to assimilate the management of the technical protocols before starting the study. In parallel, they participated in the activities during the development of the module Model of Practices and Decision-Making (LKI-312) of the matrix of the UCM School of Kinesiology.

Assignment of evaluators: The participants were registered in a database and randomly assigned to one of the two possible evaluators, using the method of the coin at random, or "heads or tails": "heads" corresponded to the EP, and "tails" corresponded to the NP. First, the two groups were evaluated by the corresponding evaluator randomly, and then the participant was instructed to return 10 days later to be re-evaluated by the remaining evaluator.

Products of the period: The NP carried out a total of 161 MIP measurements; 135 IC measurements; written report of 10 clinical cases of interest; conduct of the weekly meetings of the laboratory; control of field instruments used by undergraduate students; oral presentation of the results of the pilots of MIP and IC measurement; and formal care of subsidiary public of the laboratory.

Expert Physiotherapist: Accredited by the Ministry of Health (MINSAL), with 15 years of experience, in charge of the Laboratory of Ventilatory Function of the UCM Department of Kinesiology.

Procedures

First, spirometry was evaluated in accordance with ATS standards[16]. All patients received similar instructions at the beginning of each test session. MIP and IC were measured in a quiet room with the participant sitting on the Plethysmograph (MedGraphics Platinum Elite serie corporal®), with a time of 10 minutes between each test. Both for the MIP and IC protocols, participants were instructed to apply the nose clips and insert a mouthpiece between the teeth. After 5 cycles at tidal volume for MIP, subjects were asked to exhale completely "until your lungs are empty," and inhale "with as much strength as possible" against the occluded airway, whereas for IC they were instructed to conduct maximum inspiration until total lung capacity. For both tests we used the three highest scores with a minimum variability (as maximum 5% variability between the minimum and maximum values), in accordance with the American Thoracic Society. In addition, the number of attempts to reach the percentage of variability was monitored, in order to observe the behavior between the evaluators.

Statistical analysis: Descriptive statistics are presented as mean ± standard deviation. The distribution of the variables was evaluated using the Shapiro-Wilk test, confirming the normality of all of them. Reliability for both tests, MIP and IC, was evaluated through the intra-class correlation coefficient (ICC) with a 95% confidence interval (95%CI). The significance level was established at $p \leq 0.05$. The consistency analysis was obtained from individual differences that were evaluated using the plot method of Bland and Altman (gB/A). All statistics were calculated using the software SPSS statistics v15.0 and plots were generated with the software Graph Pad Prism 5.0v.

RESULTS

The sample consisted of 37 healthy adults aged 19–33 years (17 women, 20 men). Table 1 presents the description. All participants were evaluated by EP and NP. Table 2 presents inter-evaluator data for the three attempts in the MIP and IC tests. Analysis of the ICC shows inter-evaluator reliability (mean 0.767–0.964 of the 3 attempts) for the MIP test; as well as for the IC test (mean 0.928–0.981 of the 3 attempts). However, the *gB/A* of the difference (Figure 1 and 2) between the results obtained by evaluators A (EP) and B (NP), against the mean performance of the 37 participants in the study, showed a systematic tendency to higher results for the EP, confirming the differences in averages or means of the three attempts at 9.2 cmH_2O and 0.06 L in favor of EP, significant in the case of the MIP (Figure 1 and 2; Table 2, respectively). In addition, Table 3 presents the number of MIP and IC attempts reached by the patients according to the evaluators (NP/EP).

Table 1. Characteristics of the participants.

Variable	Mean±SD
Age (years)	23±2,8
Mass (Kg)	66,94±12,67
Height (m)	1,66±0,10
BMI (Kg/m^2)	24,11±2,93
FVC (L)	4,68±0,98

SD: standard deviation; BMI: body mass index; FVC: forced vital capacity.

Table 2. MIP and IC reliability between evaluators EP/NP.

Variable	Reference value (Mean±SD)	EP (Evaluator A) (Mean±SD)	NP (Evaluator B) (Mean±SD)	P-value	ICC	95%CI	SEM
MIP (-cmH_2O) 1st attempt	109,68+17,74	140,7±31,7	130,7±30,1	0,0002	0,914	0,731-0,964	5,08
MIP (-cmH_2O) 2nd attempt	-	139,7±31,5	130,9±30,7	0,0015	0,915	0,782-0,962	5,11
MIP (-cmH_2O) 3rd attempt	-	140,4±31,8	131,5±30,8	0,0007	0,925	0,790-0,967	5,15
IC (Liters) 1st attempt	3,32±0,76	3,30±0,70	3,26±0,71	0,393	0,955	0,912-0,977	0,11
IC (Liters) 2nd attempt	-	3,30±0,68	3,23±0,66	0,116	0,965	0,931-0,982	0,10
IC (Liters) 3rd attempt	-	3,29±0,68	3,22±0,67	0,096	0,970	0,942-0,985	0,11

MIP: maximum inspiratory pressure; cmH2O: centimeters of water; IC: inspiratory capacity; EP: expert physiotherapist; NP: novice physiotherapist; SD: standard deviation; ICC: intra-class correlation coefficient; SEM: standard error of measurement, in cmH2O for MIP and in Liters for IC. Highest MIP value, EP: 140.7±31.7; highest NP value: 131.5±30.8; highest IC value, EP: 3.30±0.70; highest value NP: 3.26±0.71.

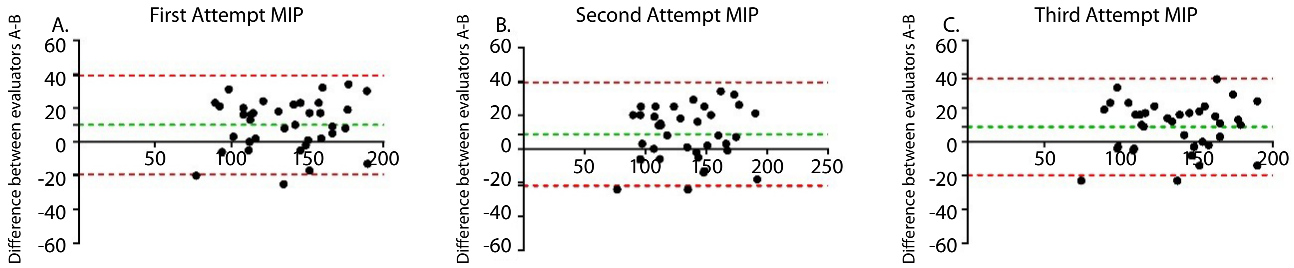

Figure 1. Bland-Altman plot of individual differences of the EP in relation to the NP, corresponding to the first, second, and third attempts (A, B, and C) of the MIP test. The green line represents the trend of the mean, while the red lines represent the limits of the confidence intervals in cmH_2O. It is observed the systematic behavior of greater magnitude of the result in favor of evaluator A (EP), in the three attempts (points on line 0, absolute agreement).

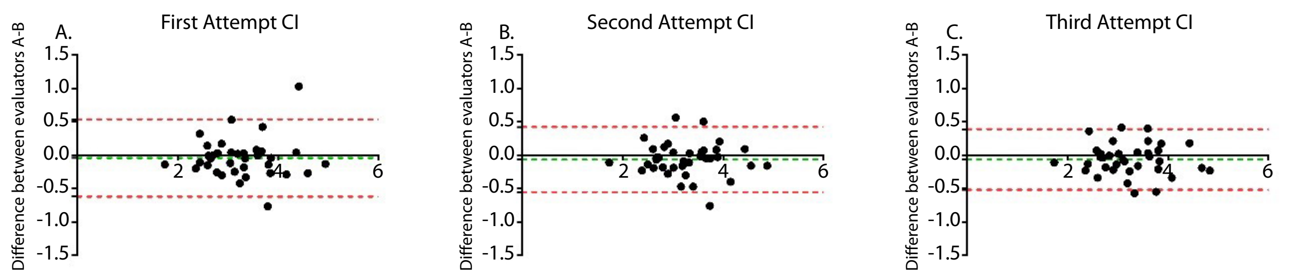

Figure 2. Bland-Altman plot of individual differences of the EP in relation to the NP, corresponding to the first, second, and third attempts (A, B, and C) of the IC test. The green line represents the mean of the trend, while the red lines represent the limits of the confidence intervals in liters. It is observed the systematic behavior of greater magnitude of the result in favor of evaluator A (EP), in the three attempts (points on line 0, absolute agreement).

Table 3. Number of attempts and percentage of minimum variability of MIP and IC according to evaluators EP/NP.

	EP (evaluator A)			NP (evaluator B)		
	Median attempts	Median Range	% minimum variability	Median Attempts	Median Range	% minimum variability
MIP	3	6	22,57	3	9	23,29
IC	3	2	20,97	3	5	21,05

No. Attempts: median of the number of attempts; Range: maximum value minus minimum value of the number of attempts; % minimum variability: standard deviation divided by the mean number of attempts, multiplied by 100.

DISCUSSION

At the end of the training period, the first finding of this study is the high reliability (Table 2) observed between the two evaluators (EP – NP) in the three attempts of the IC test. Although reports on the specific area are scarce, the results do not differ from known investigations on other domains, where it is indicated that evaluator experience has no influence on: a) goniometry and fleximetry[23]; (b) kinematic analysis[5]; or (c) gravity scores for decision-making[33]. In this regard, such studies are highlighted in their results as tools to document training programs, ensuring the development of competent procedures[28].

However, in this study, when comparing the individual difference between both evaluators, both MIP and IC presented high dispersion and standard error of measurement, observed in the gB/A (Figure 1 and Figure 2). Thus, gB/A showed that participants consistently yielded better results with the EP, in relation to the NP. In addition, we also observed significant differences in favor of the EP in the mean of the absolute value of MIP (Table 2); with a difference in the range of attempts for MIP and IC of 3, and with a minimum variability of

MIP equivalent to 0.72% and of 0.08% in IC (Table 3). Although previous studies have observed nonsignificant variabilities of these values between test and retest[29], these remain more important for MIP[30]. In this context, while the excellent ICC values attribute high reliability, they may not be sufficient to detect errors such as those observed in the *gB/A* plots for the MIP and IC tests.

Differently from Garrido and Muñoz[30], who managed in their two instances of measurement of Peak Expiratory Flow (PEF), IC and MIP a predetermined order fixing first evaluator A and then evaluator B, in this study the learning effect was controlled by randomizing the order of evaluators. In addition, we analyzed the difference between the first and the last attempt of the patient. However, subjects tended to have better performance with the EP, regardless of the random order. In other words, the order of evaluation did not influence performance. This makes it clear that the NP acquired technical skills, since they developed sufficient sensitivity to detect interindividual variations in IC performance (Table 2). These results could be interpreted as an effective acquisition of technique and discriminatory skills by the NP; however, the incidence of other variables that could be influencing the results of those evaluated cannot be disregarded (Figure 1 and Figure 2) .

On the other hand, MIP showed significant difference between EP and NP, which is consistent with previous experiences[30], to correct and improve the methodology used. In this context, it is known that expert evaluators of ventilatory function provide unambiguous instructions and manage patients allowing them to concentrate on generating maximum inspiratory effort[25], situation that is reinforced by the lower number of attempts and higher values obtained by the EP in this test (Table 2 and 3). However, it was difficult to ensure in this study the control of the motivation of the subject, despite minimizing the effects by using the same instructions for each test. Such conditions can be critical in interpreting a better capture or registration that allows increasing the quality of the learning process to obtain greater approach to the skills of the expert. However, this condition may be much more scarce with respect to other similar research, which recognize a potential learning phenomenon in the development of evaluation tests for assessment of ventilatory function given its greater variations in MIP (10.2 cmH_2O), and lower ICC (0.70) intra-observer[29].

Another aspect that can show the expertise of the EP is the lower number of attempts to achieve maximum values; in this regard, Table 3 provides records of the number of attempts (range:) 6 v/s 9 for MIP and range: 2 v/s 5 for IC). This history can be reinforced by Adaos et al. (2017), since in their comparative study it was observed that the warm-up group incorporating 30 repetitions to 40% of the MIP, and the learning group conducting 20 Mueller maneuvers could register higher mean values than those achieved with ATS/ERS regulations in men[31].

A limitation of this study is that there is no record of a learning protocol for MIP and IC tests. This may be a good future option to complement the process through the application of an evaluation with a checklist to be applied to NPs[32], in order to corroborate the optimal acquisition of technical and interpretative skills. Such evaluation tool should emphasize variables that influence patient commitment to the test, such as the ability to truly detect efforts, distraction, or lack of motivation in patients. Given that, if in physiotherapy the intention is to achieve an autonomous practice, these tendencies require that in professionals training they progressively show greater degree of certainty in clinical decision-making, independent and specific to the management of the patient/client[33].

Another limitation is that the evaluation was not blind. Although the sample was random, participants may have conditioned their performance to the experience, with the evaluator detecting slight differences with respect to their experience. An excellent research report on the influence of experience on clinical decision-making describes how the process of gaining experience evolves. Briefly, it begins with the basic knowledge and technical skills, and improves even more; as the examiner reflects on previous actions, they better detect the real efforts and is less likely to be surprised by new situations[33]. In fact, other studies show that the therapeutic alliance can influence the outcomes of patients, because experienced evaluators can be more efficient to improve the commitment of the patient with the test[34,35], a factor that could have influenced the differences observed in this study. Finally, the incorporation of preceding corrections in conjunction with a more heterogeneous sample, of elderly subjects and/or with respiratory dysfunctions, would enable that future research concerning the subject could elucidate more specifically the minimum times of training of a NP.

Based on this experience, further studies can be formulated to explore thisphenomenon in order to generate alternatives educational models that consider the transversal and reflective skills about and in the action of the NP. It has been reported previously that therapeutic and contextual factors may influence the outcomes of patients with different symptomatologic perceptions, as well as the quality of the clinician-patient alliance in the multiple clinical problems, but clearly on the context of the learning process, which has been less studied, there are specific complexities in the teaching of physical therapy, still unknown. In particular, such studies should consider the control of the learning effect, the blind effect on patients[34], and, especially, the delimitation of the strategies of the periods for the induction, transmission, and interpretation of the information relevant for the NPs, in addition to conducting this experiment with a larger number of NPs, which would provide better knowledge of interobserver reliability.

CONCLUSION

The results only confirm high reliability between EP and NP for the IC variable, since the MIP presented significant differences in favor of the EP in healthy subjects of the Maule Region. This confirms the complexity of MIP measurement and the insufficiency of the period of time used in training; therefore, in the clinical training period, reflective and discriminative skills related to the recognition of a maximum effort must be strengthened with special concern.

REFERENCES

1. Clancy C, Eisenberg J. Outcomes research: measuring the end results of health care. Ciencia. 1998;282:246-47. doi: 10.1126/science.282.5387.245.
2. Bangdiwala S, Muñoz S. Medición de confiabilidad y validez en instrumentos clínicos. Rev Med Chile. 1997;125:446-73.
3. Streiner DL. Starting at the beginning: An introduction to coefficient alpha and internal consistency. J Pers Assess. 2003;80:99-103. doi.org/10.1207/S15327752JPA8001_18.
4. Alarcón A, Muñoz S. Medición en salud: Algunas consideraciones metodológicas. Rev Med Chile. 2008;136:125-130. doi.org/10.4067/S0034-98872008000100016.
5. Bonagamba GH, Coelho DM, Oliveira AS. Confiabilidade interavaliadores e intra-avaliador do escoliômetro. Rev Bras Fisioter. 2010;14(5):432-7. doi.org/10.1590/S1413-35552010005000025.
6. Côté P, Kreitz BG, Cassidy JD, Dzus AK, Martel J. A study of the diagnostic accuracy and reliability of the Scoliometer and Adam's forward bend test. Spine. 1998;23(7):796-802.
7. Onate J, Cortes N, Welch C, Van Lunen B. Expert versus novice interrrater reliability and criterion validity of the landing error scoring system. J Sport Rehabil. 2010;19(1):41-56.
8. Mc Gaghie W, Issenberg B, Petrusa E, Scalese R. A critical review of simulation-based medical education research: 2003-2009. Med Educ. 2010;44:50-63. doi: 10.1111/j.1365-2923.2009.03547.x.
9. Black L, Hyatt R. Maximal Respiratory Pressure. Normal values and relationship to age and sex. Am Rev Respir Dis. 1969;99:669-702. doi: 10.1164/arrd.1969.99.5.696.
10. Mc Cool F, Tzelepis G. Inspiratory muscle training in the patient with neuromuscular disease. Phys Ther. 1995;75:1006-114.
11. Smith P, Coakley JM, Edward R. Respiratory muscles training in Duchenne muscular dystrophy. Muscle Nerve. 1998;11:784-85. doi: 10.1002/mus.880110716.
12. Escobar J. Entrenamiento muscular inspiratorio en pacientes con distrofia muscular de Duchenne. Una visión controversial. REEM. 2015;2:7-16.
13. Wen AS, Woo MS, Keens TG. How many maneuvers are required to measure maximal inspiratory pressure accurately. Chest. 1997;111:802-07. doi.org/10.1378/chest.111.3.802.
14. Volianitis S, McConnell A, Jones D. Assessment of maximum inspiratory pressure. Prior submaximal respiratory muscle activity ('warm-up') enhances maximum inspiratory activity and attenuates the learning effect of repeated measurement. Respiration. 2001;68(1):22-7. doi: 10.1159/000050458.
15. Sclauser P, Franco P, Fregonezi G, Sheel A, Chung F, Reid W. Reference values for maximal inspiratory pressure: A systematic review. Can Respir J. 2014;21:43-50. doi.org/10.1155/2014/982374.
16. American Thoracic Society / European Respiratory Society. ATS/ERS Statement on Respiratory Muscle Testing. Am J Respir Crit Care Med. 2002;166:518-624. doi:10.1164/rccm.166.4.518.
17. Aldrich T, Spiro P. Maximal inspiratory pressure: does reproducibility indicate full effort? Thorax. 1995;50:40-3.
18. Lechner D, Bradbury S, & Bradley L. Detecting sincerity of effort: a summary of methods and approaches. Phys Ther. 1998;75(8):867-88.
19. FND. Fundación neumológica Colombiana. Laboratorio de función pulmonar. Neumologica.org. 2008; [Online]; 2008 [Visitada con fecha 22 Febrero 2018]. Disponible en: http://www.neumologica.org/pruebas.htm.
20. Miranda M, Muñoz R. Confiabilidad y Validez del Incentivador de Volumen en la medición de la capacidad inspiratoria. REEM. 2014;1(1):27-31.
21. Lisboa C, Leiva A, Pinochet R, Repetto P, Borzone G, Díaz O. Valores de referencia de la capacidad inspiratoria en sujetos sanos no fumadores mayores de 50 años. Arch Bronconeumol. 2007;43(9):485-89.
22. Casanova C, Celli B. ¿Debemos tener en cuenta la capacidad inspiratoria? Arch Bronconeumol. 2007;43(5):245-47.
23. Oliveira V, Figueiredo A, Dos Santos S, Almeida J, Dos Santos, H. Reliability of the measures inter and intra-evaluators with universal goniometer and fleximeter. Fisioter Pesqui. 2014;21(3):229-35. doi.org/10.590/1809-2950/52921032014.
24. Sachs M, Enright P, Hinckley K, Jiang R, Barr R. Performance of maximum inspiratory pressure tests and maximum inspiratory pressure reference equations for 4 race/ethnic groups. Respir Care. 2009;54(10):1321-8.

25. Ruppel G, Enright P. Pulmonary function testing. Respir Care. 2012;57(1):165-75. doi.org/10.4187/respcare.01640.

26. Knudson R, Slatin R, Lebowitz M, Burrows B. The Maximal Expiratory Flow-Volume Curve: Normal Standards, Variability, and effects of age. Am Rev Respir Dis. 1976;113:587-99. doi: 10.1164/arrd.1976.113.5.587.

27. Gotelli N, Ellison A. A primer of ecological statistics 2ed. Massachusetts, Sinauer Associates Inc. Publishers Sunderland, 2004.

28. Ju Lin, Zwei J, Hsu T, Liu Y, Yu H, Tsai S, et al. Correlation of rater training and reliability in performance assessment: Experience in a school of dentristry. J Dental Scienc. 2013;8:256-60. doi.org/10.1016/j.jds.2013.01.002.

29. Rodríguez, I. Confiabilidad de la fuerza muscular respiratoria y flujos espiratorios forzados en adolescentes sanos. Rev Chil Enferm Respir. 2015;31:86-93. doi.org/10.4067/S0717-73482015000200003.

30. Garrido F, Muñoz R. Estudio transversal de confiabilidad inter-evaluador para la evaluación de peak del flujo expiratorio, capacidad inspiratoria y presión inspiratoria máxima. REEM. 2015;1(2):25-32.

31. Adaos C, González A, Slater D, Medina P, Muñoz R, Escobar M. Análisis de presión inspiratoria máxima según tres protocolos en estudiantes voluntarios asintomáticos de la Universidad Católica del Maule, Chile. Rev Chil Enferm Respir. 2017;33:21-30. doi.org/10.4067/S0717-73482017000100004.

32. Figueiredo K, De Lima K, Cavalcanti M, Guerra RO. Interobserver reproducibility of the Berg Balance Scale by novice and experienced physiotherapists. Physiother Theory Pract. 2009;25(1):30-6. doi: 10.1080/09593980802631330.

33. Wainwright S, Shepard K, Harman L, Stephens J. Factors that influence the clinical decision making of novice and experienced physical therapists. Phys Ther. 2010;90(1):75-88. doi: 10.2522/ptj.20100161.

34. Fuentes J, Armijo-Olivo S, Funabashi M, Miciak M, Dick B, Warren S, et al. Enhanced therapeutic alliance modulates pain intensity and muscle pain sensitivity in patients with chronic low back pain: an experimental controlled study. Phys Ther. 2014;94:477-89. doi: 10.2522/ptj.20130118.

35. Cipriany-Dacko LM, Innerst D, Johannsen J, Rude V. Interrater reliability of the Tinetti Balance Scores in novice and experienced physical therapy clinicians. Arch Phys Med Rehabil. 1997;78:1160-4.