

# Validity and reliability of the modified Tal score in Chilean children. A multicenter study

Soledad Luarte-Martínez, Kinesiologist<sup>a-d</sup>, Iván Rodríguez-Núñez, Kinesiologist<sup>b-d</sup> and Paula Astudillo, Registered Nurse and Master of Science in Health<sup>c</sup>

## ABSTRACT

**Introduction:** In pediatrics, it is decisive to assess the severity of bronchial obstruction; to this end, different clinical scoring scales have been developed, including the modified Tal score.

The objective of this study was to determine its validity and reliability in children seen at two emergency departments of Concepción, Chile.

**Population and methods:** Prospective, cohort study. Children younger than 36 months old diagnosed with bronchitis and/or bronchiolitis during the winter and spring months of 2015 were included. Concurrent criterion validity was determined based on the correlation between the score and oxygen saturation, as a reference standard. Predictive validity was assessed based on the association between the score and the probability of a new visit to the emergency department in the following 7 days. The area under the ROC curve was estimated. Reliability between kinesiologists and physicians was established based on the intraclass correlation coefficient (ICC). A *p* value < 0.05 was considered statistically significant.

**Results:** A total of 102 children were assessed; a weak correlation between the score and oxygen saturation was observed among kinesiologists ( $\text{Rho} = -0.41$ ). An association was observed between the score and the probability of a new visit to the emergency department. The area under the curve measured by both health care providers was > 0.80. Inter-observer reliability between physicians and kinesiologists showed a weak correlation ( $\text{ICC} = 0.17$ ).

**Conclusions:** The modified Tal score shows an adequate predictive validity, but a poor validity when correlated to oxygen saturation, and a weak inter-observer reliability.

**Key words:** validity, reliability, scores, bronchial obstruction.

<http://dx.doi.org/10.5546/aap.2019.eng.e340>

**To cite:** Luarte-Martínez S, Rodríguez-Núñez I, Astudillo P. Validity and reliability of the modified Tal score in Chilean children. A multicenter study. *Arch Argent Pediatr* 2019;117(4):e340-e346.

- a. Department of Kinesiology, School of Medicine, Universidad de Concepción, Concepción.
- b. School of Kinesiology, School of Health Sciences, Universidad San Sebastián, Concepción.
- c. Department of Surgery, Traumatology, and Anesthesiology, Universidad de La Frontera.
- d. Research Task Force in Cardiovascular and Respiratory Health (IDEAS-CVR), Concepción, Chile.

**E-mail address:**  
Soledad Luarte-Martínez, Kinesiologist: soledadluarte@udec.cl

**Funding:**  
Vice-Rectorship of Research and Development (Vicerrectoría de Investigación y Desarrollo, VRID) of Universidad de Concepción, award of Project VRID no. 215.082.050IN.

**Conflict of interest:**  
None.

Received: 7-25-2018  
Accepted: 1-21-2019

## INTRODUCTION

Health measurement plays a critical role in clinical decision-making.<sup>1</sup> This is why measurement instruments have been designed to facilitate the measurement process.<sup>2</sup> Instrument quality should be assessed based on psychometric properties, including validity and reliability.<sup>3</sup> This will allow to establish an adequate clinical decision-making process, optimize health resources, and make a significant impact on health care effectiveness.<sup>3</sup>

Respiratory diseases impose a great burden on health care systems worldwide.

From an epidemiological perspective, a disease with a high incidence that causes the greatest demand for primary health care during the winter season is acute respiratory infection (ARI).<sup>5</sup> In Chile, obstructive bronchial diseases account for the most frequent ARIs, including acute bronchitis and bronchiolitis.<sup>4</sup> Both conditions are common and demand for a high investment in health.<sup>6,7</sup>

Bronchiolitis is the most common respiratory disease in children younger than 2 years.<sup>8</sup> It is known as the first instance of acute wheezing affecting a previously healthy child.<sup>7,9</sup> On its side, acute obstructive bronchitis is characterized by acute bronchial obstruction and wheezing and, sometimes, crackles and rhonchi that mostly affect children younger than 3 years.<sup>7</sup> Different clinical scoring scales<sup>10-13</sup> have been designed to determine the severity of bronchial obstruction in these diseases. In Chile, the Ministry of Health has recommended the modified Tal score, which was obtained from the

instrument originally developed in the 1970s by Bierman and Pierson.<sup>14</sup> The subsequent version developed by Tal et al.,<sup>15</sup> replaced the inspiratory-expiratory ratio by the presence of cyanosis and this was called the Tal score. The version modified by Girardi et al.,<sup>16</sup> in the late 1990s, grouped the respiratory rate item into two age groups: younger and older than 6 months, and this was known as the modified Tal score.

The instrument's methodological structure encompasses 4 items: respiratory rate, retractions, use of accessory muscles, and wheezing. Although this score has been validated before, only partial assessments of its psychometric properties have been established.<sup>11,17,18</sup> This is why it is necessary to supplement such assessments in the current Chilean population.

The objective of this study was to determine the validity and reliability of the modified Tal score in children with acute bronchial obstruction seen at two emergency departments of the city of Concepción, Chile.

## POPULATION AND METHODS

This was a prospective, cohort study. Patients were selected using a consecutive non-probability sampling among those who sought care at the emergency department of Centro de Salud Víctor Manuel Fernández, a primary care facility in the city of Concepción. In addition, patients who were seen at the emergency department of Hospital Guillermo Grant, also in the same city, were selected. The study was conducted between August and November 2015. It is worth mentioning that both sites were within the urban perimeter of the city, so they take care only for the population living in that area.

Inclusion criteria were children younger than 36 months diagnosed with bronchitis, defined as a respiratory condition that presented as bronchial obstruction and/or acute bronchiolitis, i.e., the first episode of bronchial obstruction. Children with chronic pulmonary disease (long-standing pulmonary disease, such as asthma or cystic fibrosis), cardiovascular disease (such as heart disease) or pre-existing neuromuscular condition were excluded.

The sample size was estimated based on a 0.4 correlation coefficient, a 5 % risk for a type I error, and a 20 % risk for a type II error. Thus, the minimum number of subjects was 46 children. However, to increase the statistical power of outcomes above 99 %, a higher number of patients was included.

**Ethical considerations:** The study was approved by the Scientific Ethics Committee of the Concepción Health Service (CEC code: 15-07-40) in August 2015; in addition, children's parents signed an informed consent.

**Outcome measures:** The general characteristics of each child were recorded (sex, age, weight, height, level of maternal education, place of residence, type of housing, attending an educational establishment), together with the percentage of oxygen saturation ( $\text{SaO}_2$ ) measured using a pulse oximeter (Prince-100L).

The severity of bronchial obstruction was determined based on the score obtained from applying the 4 items of the instrument.

**Procedures:** Once children were admitted, their  $\text{SaO}_2$  was recorded. This outcome measure was considered the reference standard to which the instrument's score was related, thus determining the concurrent criterion validity.

Then, the instrument was applied twice: once by a physician and once by a kinesiologist, at each site, in an independent manner and without one observer knowing the score obtained by the other, less than 5 minutes apart, thus determining inter-observer reliability. Participating health care providers had at least two years of clinical experience and were not related to the study.

Then, children follow-up was done through phone calls to establish the need for a new visit to the emergency department in the following 7 days and assess the score's predictive validity, i.e., the instrument's ability to predict a specific outcome.

**Statistical analysis:** The MedCalc Statistical Software, version 16.2.0, was used to assess the normal distribution of data with the Kolmogorov-Smirnov test, which verified the lack of a normal distribution in most outcome measures. Then, descriptive statistics were established by estimating the median, maximum, and minimum values.

The Wilcoxon test was used to verify differences in the instrument's outcome measures between the first and the second observation. Concurrent criterion validity was assessed by estimating Spearman's correlation coefficient (Rho) between the instrument's score and  $\text{SaO}_2$ , a methodological approach used in other studies that explored the validity of these instruments.<sup>11,12,17,18</sup> Given that physicians and kinesiologists assessed selected children, the scores obtained by both were taken into consideration. A common cut-off point was used to assess the correlation: > 0.90 (excellent), 0.75-

0.90 (good), 0.60-0.75 (moderate), and < 0.60 (low).<sup>19</sup>

Predictive validity was assessed using a logistic regression analysis and odds ratio (OR). In the regression model, the instrument's score was an independent outcome measure and the probability of a new visit to the emergency department in the following 7 days was a dependent outcome measure. In addition, a Receiver Operating Characteristic (ROC) curve was generated to differentiate the instrument's ability to determine the probability of a new visit to the emergency department. The area under the curve (AUC) was described as absolute value and 95 % confidence interval (CI); in turn, the score that showed the best sensitivity and specificity value was reported (Youden index [J]). The AUC obtained based on the instrument application by the kinesiologists and physicians was compared using DeLong's non-parametric test.<sup>20</sup>

The inter-observer reliability between the score obtained by physicians and kinesiologists was estimated using the intraclass correlation coefficient (ICC). Mean discrepancy and discrepancy limits between physicians and kinesiologists were also estimated using the Bland-Altman method.<sup>21</sup>

Validity and reliability indices were described as absolute value and 95 % CI. A *p* value of < 0.05 was considered significant.

## RESULTS

A total of 102 children were included in the study. *Table 1* shows the biodemographic characteristics of patients. No statistically significant difference in items was noted between observers in relation to the severity of bronchial obstruction and the total score (*Table 2*).

TABLE 1. General characteristics of the study sample

Outcome measures	Results
Number of patients	102
Age (months old)	10 (1-35)
Weight (kg)	9 (4-22)
Height (cm)	71 (120-49)
Oxygen saturation (%)	97 (92-99)
Attending an educational establishment (n)	34 (33.3 %)
Primary maternal education (n)	15 (14.7 %)
Secondary maternal education (n)	61 (59.8 %)
Tertiary maternal education (n)	26 (25.5 %)

Quantitative outcome measures are described as median and range, whereas qualitative outcome measures are expressed as absolute value and percentage.

**Concurrent criterion validity:** Considering all measurements, a weak correlation between the instrument and  $\text{SaO}_2$  was observed. However, when the instrument was applied by kinesiologists, a greater correlation to the reference standard was observed compared to physicians (*Table 3*). Besides, the correlation between the score items and  $\text{SaO}_2$  was weak and only statistically significant in terms of retractions and respiratory rate (*Table 4*).

**Predictive validity:** Follow-up was done in all children; of them, 4 (3.9 %) reported that they required care in the 7 days after the measurement; a lower  $\text{SaO}_2$  (94.5 % [93-99] versus 97.0 % [93-99]; *p* = 0.0215) and a higher score as measured by kinesiologists (5.0 [3-7] versus 2.0 [0-9]; *p* = 0.0261) and physicians (5.5 [3-7] versus 2.0 [0-7]; *p* = 0.012) were observed compared to those who did not need to make a new visit. The logistic regression analysis revealed that there was an association between the instrument's score and the risk for the need to make a new visit to the emergency department in both observers (*Table 3*).

The instrument's AUC to predict a new visit was > 0.80 in both kinesiologists and physicians (*Figure 1*). No significant difference was observed between them (*Table 3*). Considering all measurements, a modified Tal score > 3 showed a sensitivity of 0.75 (0.35-0.97) and a specificity of 0.70 (0.63-0.77). Lastly, the score obtained by kinesiologists showed a J index > 2, while that obtained by physicians was > 5.

**Inter-observer reliability:** The ICC of the total score estimated by both observers was 0.17 (95 % CI: 0.03-0.35).

Finally, *Figure 2* is a Bland-Altman plot that shows that the mean score difference between assessments was 0.2 points, with discrepancy limits of 5.1 and -4.7.

TABLE 2. Results for each item and total score of the modified Tal score applied by kinesiologists and physicians

Outcome measures	Kinesiologists (n= 102)	Physicians (n= 102)	<i>p</i> value
RR < 6 months old	1 (0-3)	1 (0-3)	0.1039
RR ≥ 6 months old	1 (0-3)	1 (0-3)	0.8076
Wheezing	1 (0-3)	1 (0-3)	0.5230
Cyanosis	0 (0-1)	0 (0-0)	0.1583
Retractions	1 (0-3)	1 (0-3)	0.5134
Total score	2 (0-9)	2 (0-8)	0.3333

Results are expressed as median and range.

RR: respiratory rate.

## DISCUSSION

Considering the prevalence and severity of ARIs in our population have changed in the past 20 years,<sup>5</sup> there is a current need to assess the validity and reliability of the modified Tal score, an instrument that is widely used in Chile and worldwide.<sup>11,13,15</sup>

In relation to concurrent criterion validity, the results indicated that there was a weak correlation between the score and  $\text{SaO}_2$ ; this revealed that the instrument had determined the severity of bronchial obstruction to a low percentage in the studied sample. This was consistent with previous studies, such as the one by McCallum et al.,<sup>11</sup> who observed that the instrument had a low diagnostic ability (ROC AUC: 0.6), considering saturation as a reference criterion.

It is worth noting that the evidence supporting the clinical applicability of  $\text{SaO}_2$  as a reference standard to determine the severity of bronchial obstruction has demonstrated certain inconsistencies; therefore, most likely, using a reference criterion to directly assess this parameter would allow to establish the validity of these instruments with more certainty.<sup>11,12</sup>

TABLE 3. Concurrent criterion validity and predictive validity of the modified Tal score

	Kinesiologist (n= 102)	Physician (n= 102)
Spearman's Rho	-0.41 (-0.56 to -0.24) <sup>d</sup>	-0.16 (-0.34 to 0.04)
AUC	0.83 (0.74 to 0.89) <sup>c</sup>	0.87 (0.79 to 0.93) <sup>d</sup>
Odds ratio	1.78 (1.10 to 2.98) <sup>a</sup>	2.04 (1.14 to 3.65) <sup>b</sup>

Concurrent criterion validity is described using Spearman's Rho correlation coefficient. Predictive validity is expressed using the AUC and odds ratio. AUC: area under the ROC curve. ap = 0.0218, bp = 0.0064, cp = 0.0004, dp < 0.0001.

Results are shown as absolute value and (95 %) confidence interval.

TABLE 4. Correlation coefficient between peripheral oxygen saturation and the results of the items in the modified Tal score

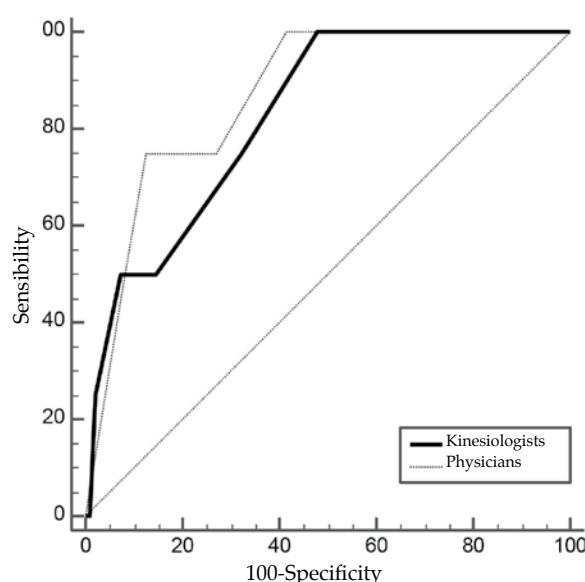
	Wheezing	Cyanosis	Retractions	Respiratory rate
Rho	-0.240	0.025	-0.337	-0.337
p	0.15	0.8	0.001	0.001

In addition, Pavón et al.,<sup>17</sup> observed that the item cyanosis in the modified Tal score showed a weak correlation to saturation. This was consistent with the observations of this study, which showed that cyanosis was not correlated to the reference standard.

In accordance with what has been observed, there were differences in concurrent validity between physicians and kinesiologists. Specifically, kinesiologists showed a weak correlation between the score and saturation, whereas physicians did not observe a correlation between both outcome measures. Such difference may be explained by the clinical training profile of the providers involved in the management of children with obstructive bronchial disease. In Chile, kinesiologists receive, during their instruction, greater training in infant respiratory health, unlike general practitioners.

In relation to predictive validity, results showed an association between the instrument's score and the probability of a new visit to the emergency department in the following 7 days, thus suggesting that the instrument may be capable of predicting the probability of future respiratory complications, regardless of the health care provider applying it. The AUC was above 0.80 for both physicians and kinesiologists. Therefore, there was an important probability of a higher score in children with complications in

FIGURE 1. ROC curve for the modified Tal score as predictor of a new visit to the emergency department



the following 7 days. This was consistent with the results obtained by Duarte-Dorado et al.,<sup>22</sup> who showed that the higher severity score in a bronchial obstruction scale similar to the modified Tal score –which included saturation, wheezing, retractions, and connection to the environment as items– was associated with a greater need for hospitalization.

It is worth mentioning that the J index is the maximum vertical distance between the ROC curve and the diagonal line.<sup>23</sup> In relation to the cut-off points obtained in this study, there is a greater probability of a new visit in the following 7 days, which may suggest the implementation of an early treatment and an adequate monitoring. These observations confirm what has been described in the Chilean guidelines for the management of respiratory infections, which establish a  $> 5$  score as the cut-off point to implement a brief procedure called abbreviated hospitalization, which consists in administering aerosol therapy and respiratory physical therapy in the outpatient setting of the facility.<sup>7</sup>

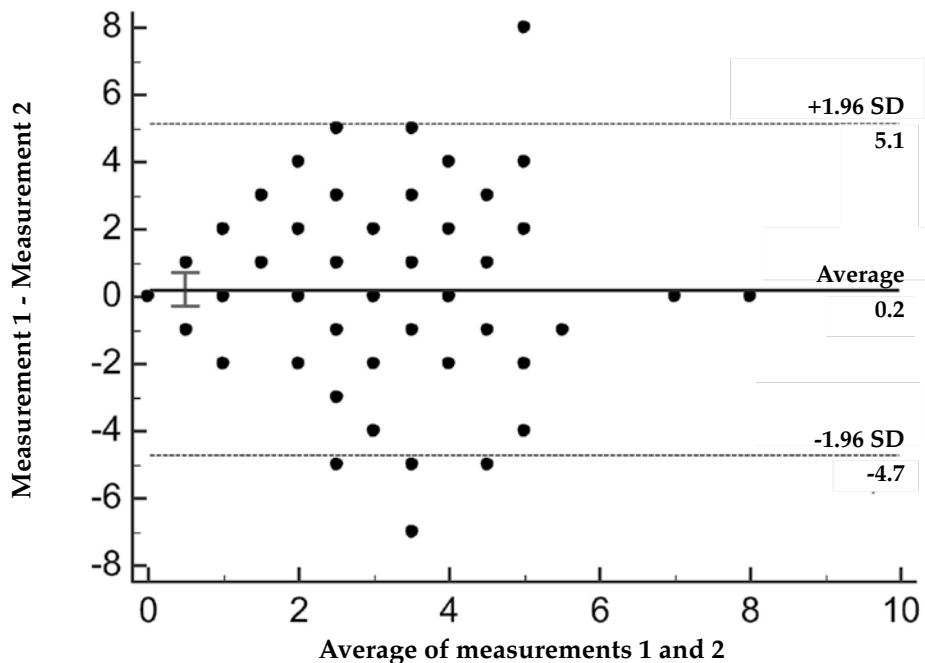
The instrument's reliability has been previously assessed in a Chilean study conducted by Urzúa et al.,<sup>24</sup> who analyzed reliability between

medical students and pediatric residents and concluded that agreement between observers was poor. These results were confirmed in our study, which found an 83 % measurement error between observers (ICC = 0.17), possibly due to the inherent differences in the training on pediatric pulmonology between kinesiologists and physicians.

It is worth pointing out that certain studies consider the special training of observers for the application of instruments in their experimental protocols. These protocols have reported higher reliability levels than those observed in this study. McCallum et al.<sup>10</sup> determined a high inter-observer reliability ( $k: 0.70$ ; 95 % CI: 0.63-0.83) among different health care providers who had previously received training.

One of the limitations of this study is that most of the sample participants had a score between 0 and 5, which is equivalent to a mild bronchial obstruction. This may have affected the possibility of establishing a correlation between the studied outcome measures. Likewise, it was not possible to know the size of the correlation between the outcome measures in the most severe conditions. In addition, few children had complications in

FIGURE 2. Inter-observer reliability for measurement obtained with the modified Tal score



Measurement 1 corresponds to kinesiologists, and measurement 2, to physicians. The average discrepancy is described as absolute value and confidence interval.

SD: standard deviation.

the following 7 days, thus limiting the statistical power of predictive validity. Therefore, future studies are required to confirm such results.

It is also worth noting that most assessed children lived in the urban perimeter of the city, which may have influenced the external validity of our results, considering that the city of Concepción has both an urban and a rural demographic distribution.<sup>25</sup>

Finally, it may be concluded that the modified Tal score applied in children younger than 3 years with acute bronchiolitis or bronchitis has an adequate predictive validity; however, it shows a poor criterion validity when correlated to SaO<sub>2</sub> and a weak inter-observer reliability. It is worth noting that future studies are required to verify the potential impact on measurement properties of a training program about the application of this instrument in the population of interest. ■

## REFERENCES

1. Clancy C, Eisenberg J. Outcomes research: measuring the end results of health care. *Science*. 1998; 282(5387):245-6.
2. Kerlinger F, Lee H. *Investigación del comportamiento. Métodos de investigación en ciencias sociales*. 4ta ed. México: McGraw-Hill; 2002.
3. Alarcón MA, Muñoz NS. *Medición en Salud: Algunas consideraciones metodológicas*. *Rev. Med Chile*. 2008; 136(1):125-30.
4. Foro de las Sociedades Respiratorias Internacionales. *El impacto mundial de la Enfermedad Respiratoria*. 2da ed. México: Asociación Latinoamericana de Tórax; 2017. [Accessed on: December 2<sup>nd</sup>, 2018]. Available at: [http://www.who.int/gard/publications/The\\_Global\\_Impact\\_of\\_Respiratory\\_Disease\\_ES.pdf](http://www.who.int/gard/publications/The_Global_Impact_of_Respiratory_Disease_ES.pdf).
5. Organización Mundial de la Salud. *Medidas de control de infecciones en la atención sanitaria de pacientes con enfermedades respiratorias agudas en entornos comunitarios*. 2009. [Accessed on: December 2<sup>nd</sup>, 2018]. Available at: [http://www.who.int/csr/resources/publications/WHO\\_HSE\\_GAR\\_BDP\\_2009\\_1/es/](http://www.who.int/csr/resources/publications/WHO_HSE_GAR_BDP_2009_1/es/).
6. Chile. Departamento de Estadística e Información de Salud, Ministerio de Salud. 2014. [Accessed on: December 2<sup>nd</sup>, 2018]. Available at: <http://www.deis.cl>.
7. Chile. Ministerio de Salud. *Guía Clínica Auge: Infección Respiratoria Baja de Manejo ambulatorio en menores de 5 años*. 2da ed. 2013. [Accessed on: December 2<sup>nd</sup>, 2018]. Available at: <https://www.minsal.cl/portal/url/item/7220fdc4341244a9e04001011f0113b9.pdf>.
8. Ralston S, Lieberthal A, Meissner H, Alverson B, et al. Clinical practice guideline: the diagnosis, management, and prevention of bronchiolitis. *Pediatrics*. 2014; 134(5): e1475- 502.
9. Jartti T, Smits HH, Bonnelykke K, Bircan O, et al. Bronchiolitis needs a revisit: Distinguishing between virus entities and their treatments. *Allergy*. 2018;1-13. [Epub ahead of print].
10. Ramos Fernández J, Cordón Martínez A, Galindo Zavala R, Urda Cardona A. Validación de una escala clínica de severidad de la bronquiolitis aguda. *AnPediatri (Barc)*. 2014; 81(1):3-8.
11. McCallum G, Morris P, Wilson C, Versteegh L, et al. Severity scoring systems: Are they internally valid, reliable and predictive of oxygen use in children with acute bronchiolitis? *Pediatr Pulmonol*. 2013; 48(8):797-803.
12. Bekhof J, Reimink R, Brand PL. Systematic review: Insufficient validation of clinical scores for the assessment of acute dyspnoea in wheezing children. *Pediatr Respir Rev*. 2014; 15(1):98-112.
13. Destino L, Weisgerber M, Soung P, Bakalarski D, et al. Validity of Respiratory Scores in Bronchiolitis. *Hosp Pediatr*. 2012; 2(4):202-9.
14. Bierman W, Pierson W. The Pharmacologic Management of Status Asthmaticus in Children. *Pediatrics*. 1974;54(2):245-7.
15. Tal A, Bavilski C, Yohai D, Bearman J, et al. Dexamethasone and Salbutamol in the Treatment of Acute Wheezing in Infants. *Pediatrics*. 1983; 71(1):13-8.
16. Girardi G, Astudillo P, Zúñiga F. El programa IRA en Chile: hitos e historia. *Rev Chil Pediatr*. 2001; 72(4):292-300.
17. Pavón D, Castro D, Rodríguez J, Rubilar L, Girardi G. Relation between pulse oximetry and clinical score in children with acute wheezing less than 24 months of age. *Pediatr Pulmonol*. 1999; 27(6):423-7.
18. Puebla Molina S, Bustos L, Valenzuela M, Hidalgo M, et al. La escala de Tal como test diagnóstico y el diagnóstico clínico como gold standard en el síndrome bronquial obstructivo del lactante. *Rev Pediatr Aten Primaria*. 2008; 10(37):45-53.
19. Portney L, Watkins M. *Foundations of clinical research: applications to practice*. 3rd ed. Upper Saddle River, NJ: Pearson/Prentice Hall; 2009.
20. DeLong E, DeLong D, Clarke-Pearson D. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988; 44(3):837-45.
21. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 8(8476):307-10.
22. Duarte-Dorado D, Madero-Orostegui D, Rodríguez-Martínez C, Nino G. Validation of a scale to assess the severity of bronchiolitis in a population of hospitalized infants. *J Asthma*. 2013; 50(10):1056-61.
23. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its Associated Cutoff Point. *Biom J*. 2005; 47(4):458-72.
24. Urzúa BS, Duffau TG, Zepeda FG, Sagredo JS. Estudio de concordancia clínica en educandos de pre y pos título en Pediatría: Puntaje de Tal. *Rev Chil Pediatr*. 2002; 73(5):471-7.
25. Chile. Instituto Nacional de Estadísticas. Censos de Población y Vivienda 2002. [Accessed on: January 24<sup>th</sup>, 2019]. Available at: <http://www.ine.cl/estadisticas/censos/censos-de-poblacion-y-vivienda>.