

Psychometric properties of scales used for grading the severity of bronchial obstruction in pediatrics: A systematic review and meta-analysis

Soledad Luarte-Martínez, Kinesiologist, Magister^{a,b}, Iván Rodríguez-Núñez, M.D., Kinesiologist^{b,c}, Paula Astudillo, Nurse, Magister^{b,d,e} and Carlos Manterola, M.D.^{b,d,f}

ABSTRACT

Introduction. In pediatrics, identifying the severity of bronchial obstruction in an early manner is a decisive factor.

Objective. To assess the psychometric properties of the scales for grading the severity of bronchial obstruction in pediatric patients.

Population and Method. This was a systematic review of studies on the validity and reliability of scales for grading the severity of bronchial obstruction conducted in infants and children younger than 3 years old. The search was conducted in Medline, WoS, EMBASE, SciELO, and Google Scholar. The correlation coefficient corresponding to each article was included in a random effects model to establish the criterion validity and reliability using the weighted averages of coefficients as per the sample size.

Results. A total of 9 articles were included, which accounted for 2699 children; 3 articles had an adequate or excellent methodological quality. Four articles established the concurrent criterion validity considering oxygen saturation, with a weighted correlation coefficient of -0.627 (95% confidence interval [CI]: -0.767 to -0.431, $p < 0.001$); 2 articles established the convergent criterion validity, with a weighted correlation coefficient of 0.809 (95% CI: 0.721 to 0.871, $p < 0.001$); 6 articles established the inter-observer reliability, with a weighted correlation coefficient of 0.500 for kappa and 0.891 for the intraclass correlation coefficient.

Conclusion. The assessment of psychometric properties to support the use of scales for grading the construct "severity of bronchial obstruction" showed a moderate to adequate criterion validity. The percentage of agreement among observers in terms of the studied measure (severity of bronchial obstruction) was adequate; however, weaknesses such as the article design should be taken into account since it may affect the internal validity of results.

Key words: result reproducibility, obstructive pulmonary diseases, result reliability, result validity, scales.

<http://dx.doi.org/10.5546/aap.2017.eng.241>

To cite: Luarte-Martínez S, Rodríguez-Núñez I, Astudillo P, et al. Psychometric properties of scales used for grading the severity of bronchial obstruction in pediatrics: A systematic review and meta-analysis. *Arch Argent Pediatr* 2017;115(3):241-248.

- a. Department of Kinesiology, Universidad de Concepción, Concepción, Chile.
- b. Doctoral Program in Medical Sciences, Universidad de la Frontera, Chile.
- c. School of Kinesiology, Universidad San Sebastián, Concepción, Chile.
- d. Department of Surgery, Universidad de La Frontera, Chile.
- e. Doctoral Program in Psychology, Universidad de Girona, Spain.
- f. Biomedical Research Center, Universidad Autónoma de Chile.

E-mail address:

Paula Astudillo, Nurse, Magister:
paula.astudillo@ufrontera.cl

Funding:

This study was funded by the Vice-Rectorship of Research and Development (Vicerrectoría de Investigación y Desarrollo, VRID) of Universidad de Concepción, Project No. 215.082.050IN.

Conflict of interest:

None.

Received: 6-28-2016

Accepted: 11-7-2016

INTRODUCTION

Acute bronchiolitis is a common disease during childhood and is the main cause for admission due to an acute lower respiratory tract infection (ALRTI) among children younger than 2 years old.¹ In Latin America, acute respiratory tract infections are the main reason for pediatric hospitalizations (98% of these infections are secondary to a lower respiratory tract infection).^{2,3} The main ALRTIs include obstructive bronchial diseases, such as acute obstructive bronchial syndrome and bronchiolitis.^{2,4,5} Obstructive bronchial syndrome is characterized by acute respiratory obstruction and wheezing, usually of viral etiology.^{4,6} Bronchiolitis is the first obstructive event among infants and its diagnosis is preferably made based on the patient's history, signs, and symptoms.⁶

One of the factors determining the clinical course of acute respiratory tract infections in infants is the early identification of the severity of bronchial obstruction. For this reason, many clinical scoring scales have been developed based on different domains representative of the signs and symptoms typical of these conditions, which give rise to the construct "severity of bronchial obstruction."⁷ Some of these scales include the acute bronchiolitis severity scale,⁸ Wang's score (WS),⁹ the respiratory distress scale by the Ministry of Health of Argentina (RDSMoHA),¹⁰ the Respiratory Distress Assessment Instrument (RDAI),¹¹ the Children's Hospital of Wisconsin Respiratory Score (CHWRS),¹² the Wood's Clinical Asthma Score (WCAS),¹³ Tal's score,¹⁴

and the Tal's score, modified.¹⁵ The properties of these scales should consider the severity of bronchial obstruction in its entirety and indicate an association between the measurement outcome and the severity of bronchial obstruction, thus categorizing individuals and targeting therapeutic strategies.¹⁶

Several studies have looked into the psychometric properties of scales for grading the severity of bronchial obstruction and showed inconsistent results in terms of validity and reliability.^{10-12,14,15,17} Therefore, it is necessary to establish measurement properties of these type of instruments in a comprehensive manner to assess the usefulness of estimating the severity of bronchial obstruction based on indirect methods in the clinical setting.^{7,17,18}

The objective of this review was to establish the psychometric properties of the scales used for grading the severity of bronchial obstruction in infants and children younger than 3 years old.

POPULATION AND METHOD

Design: This was a systematic review and meta-analysis conducted in accordance with the Preferred Reporting Items in Systematic Reviews and Meta-Analyses (PRISMA) Statement.¹⁹⁻²¹

Study inclusion criteria: Every article about establishing the validity and reliability of scales for grading the severity of bronchial obstruction in infants and children younger than 3 years old seen at a hospital due to obstructive bronchial diseases, with no restrictions in terms of gender or race, was included. Articles written in English, Spanish and Portuguese were taken into consideration.

Study exclusion criteria: Articles were excluded if the subject matter was not relevant, or if they were reviews, discussions, and articles that grouped children with a concomitant cardiovascular or chronic pulmonary disease, or about studies based on diagnostic tests with a wide range of reference criteria, cut-off points, and inconsistent reporting in relation to the area under the curve.

Response outcome measures considered in the studies: a) Concurrent criterion validity: if the scale was correlated to an external criterion ("gold standard"), whether the total score provided by the scale for grading the severity of bronchial obstruction was close to the criterion or not. b) Convergent criterion validity: whether the measurements done with the same feature and different methods correlated. Correlation

values ranged between -1 and +1; the closer the value to 1 (either + or -), the greater the validity; the closer the value to 0, the smaller the validity; the +/- sign depended on the direction of the relationship. c) Inter-observer reliability: whether there was a correlation between the scores obtained from different observers; it showed the percentage of agreement in relation to the measure observed (severity of bronchial obstruction) and corrected the random factor, i.e., the scale's ability to produce the same results regardless of who uses it (values ranged between 0 and 1; the closer the value to +1, the greater the agreement).

Sources of information and study identification: The first search included the following databases: Medline, WoS, EMBASE, SciELO, and Google Scholar, from their initiation to November 2015. The second search included Medline and SciELO, from November 2015 to June 2016.

The following Medical Subjects Headings (MeSh) were used: bronchiolitis, result reproducibility, statistics, viral bronchitis, obstructive pulmonary diseases, and study validation. Also, the following free terms were used: bronchial obstruction, acute bronchiolitis, acute bronchitis, validation, reproducibility, reliability, correlation, agreement, scale, score, clinical score. The Boolean operators AND and OR were also used, and "humans," "infants," and "children" were used as search limits.

Data collection: Data were collected from studies that met the inclusion criteria in a special worksheet developed by two of the investigators, independently of each other. The following data were collected: year and language of publication, sample size, participants' age, severity of bronchial obstruction scale assessed, validity criterion, reliability, and physiological outcome measure used as reference criterion. Discrepancies in data collection were solved by consensus with a third member of the research team.

Methodological quality (MQ) and risk of bias assessment: MQ was assessed by two investigators, independently of each other, using the Consensus-based Standards for the Selection of Health Measurement (COSMIN) checklist to establish the MQ of studies targeted at analyzing the psychometric properties of health measurement parameters. Only the checklist section regarding the assessment of reliability and validity studies was used.^{22,23}

MQ was classified into excellent, adequate, reasonable, and poor. "Excellent" was assigned

if the methodological quality of the study was appropriate. "Adequate" was assigned if relevant information was not provided in the article but the quality was assumed to be adequate. "Reasonable" was assigned if there were concerns regarding the MQ. "Poor" was assigned if there was evidence that the MQ was not adequate. Discrepancies in MQ assessment were solved by consensus with a third member of the research team.

The risk of publication bias was established based on the correlation between the size of the absolute value of the statistical rate that established the measurement property and the sample size using Kendall's tau rank correlation coefficient (CC) (Begg and Mazumdar's rank correlation test). To this end, every coefficient was multiplied by -1. In addition, a funnel plot was developed for criterion validity indexes to establish the risk of selection bias; the vertical line accounted for the coefficient weighted mean, and the diagonal line, for the limits (95% confidence interval [CI]) of the distribution expected in the absence of a selection bias.²⁴

Statistical analysis: The statistical analysis of data was done using the MedCalc software, version 15.8 (MedCalc Software bvba, Ostend, Belgium; <https://www.medcalc.org>; 2015). Descriptive statistics were established using average and standard deviation for quantitative outcome measures and percentages, for categorical outcome measures.

The bivariate correlation (Pearson's r) was used as concurrent validity criterion. The intraclass correlation coefficient (ICC) and the kappa coefficient were used as reliability index.

The meta-analysis of studies that established criterion validity was done based on the Hedges and Olkin's method,²⁵ using Fisher's z transformation from CCs. Inconsistency was estimated using the I² statistics. Considering the discrepancy in terms of article MQ, the meta-analysis was based on the random effects model. For the meta-analysis of studies that established reliability, the weighted average (WA) of the ICC and of the kappa statistics was estimated based on the sample size, according to the sum of each article's weighted coefficient (β). The ICC and kappa WA is the sum of the weighted coefficients as per the following formula:

$$WA = \Sigma\beta = \frac{\Sigma Qi * ni}{\Sigma ni}$$

Where:

i: article.

Q: reliability coefficient used (ICC or kappa).

n: number of subjects.

Σni: sum of all "n" in articles using Q.

Ethical considerations: Authors, study sites, and primary article titles were blinded to prevent any selection and analysis bias.

RESULTS

Study selection: The search obtained 679 articles: 275 were from Medline; 11, from WoS; 17, from EMBASE; 11, from SciELO; and 365, from Google Scholar. Articles were excluded due to duplication and irrelevant topic; 30 potentially relevant abstracts were left. Of these, 19 articles were excluded because they included adults or a concomitant chronic disease with no subset analysis, did not define the scale used nor met some of the selection criteria. Of the 11 articles (full texts), 2 were excluded because they included adults and statistical analysis using the receiver operating characteristic (ROC) curve (Figure 1).

Characteristics of articles: Out of the 9 selected articles, 7 were in English and 2, in Spanish. The year of publication ranged between 1999 and 2015. The sample size ranged between 36 and 1765 participants, and it was not reported in one of the articles. Participants' average age was reported in 8 out of the 9 articles, and the age range, in 5 out of the 9. The WA of age was 4.2 months old (maximum: 6.3, minimum: 1.7).

In relation to the assessed measurement properties, 3 articles only established criterion validity (p= 374, 12.9%); 4, only reliability (p= 2417, 83.3%); and 2, both properties (p= 108, 3.7%) (Table 1).

MQ and risk of bias: Only 3 articles had an adequate or excellent MQ; the other 6 had a reasonable or poor MQ (Table 1). In relation to the risk of publication bias, there was a small correlation between the absolute value of CCs and the studies' sample sizes. In this regard, Kendall's tau CC was -0.447 (p= 0.1415) for the criterion validity studies, and -0.414 (p= 0.1734) for reliability studies, which ruled out any publication bias in these studies (Figure 2. A). For its part, the funnel plot showed that most validity studies were within the confidence limit for a null selection bias. Only one article related to concurrent criterion validity regarding the scales for grading the severity of bronchial obstruction was found to

be outside the confidence limit (Figure 2. B).

Identified scales for grading the severity of bronchial obstruction: The following nine scales were identified: the Kristjansson scale (KS)

($p=54, 1.9\%$),²⁶ Wang's score (WS)²⁶ ($p=54, 1.9\%$), Tal's score²⁷ ($p=112, 3.9\%$), Tal's score, modified by McCallum²⁸ (TSMc) ($p=112, 3.9\%$), Tal's score, modified by Pavón¹⁵ (TSP) ($p=138, 5.1$),

FIGURE 1. Flow chart of primary studies

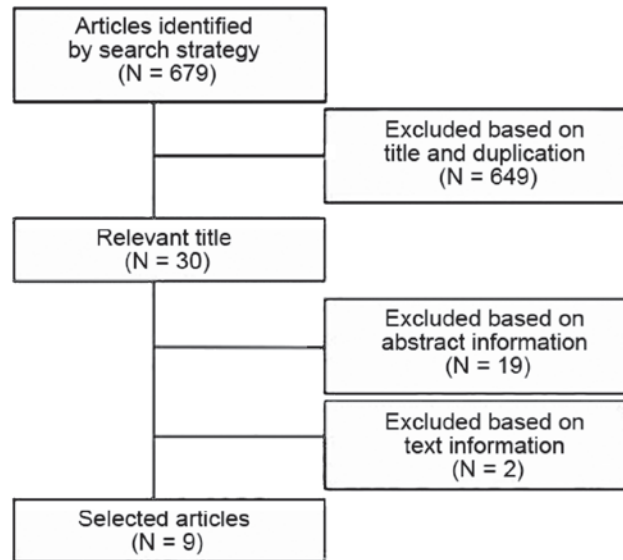


TABLE 1. Characteristics of primary articles included in the systematic review. N= 9

Author, year	No. of cases (% of males)	Age (months old)	Scale	Measurement property	Reference criterion	Statistics	MQ
Ho Jen Chin, 2004	54 (63)	6-18	Kristjansson Wang	Concurrent criterion validity	O ₂ saturation	Pearson's r	Reasonable
				Inter-observer reliability	NA	ICC	Reasonable
Pavón, 1999	138 (66)	1-24	Tal's score, modified	Concurrent convergent criterion	O ₂ saturation	Pearson's r	Reasonable
Fernández, 2015	1765 (62)	< 12 m	Respiratory distress	Inter-observer reliability	NR	ICC	Reasonable
Camargo, 2014	36 (55)	< 24 m	Tal's score, modified	Convergent criterion validity	Wood-Downes modified CAS	Pearson's r	Reasonable
McCallum, 2012	112 (65)	< 24 m	Tal's score, modified	Confiabilidad interobservador	NR	Kappa	Excellent
Duarte, 2012	54 (56)	< 24	Wood-Downes modified CAS	Convergent criterion validity	Tal's score	Spearman's r	Reasonable
				Inter-observer reliability	NA	Kappa	Reasonable
Destino, 2012	195 (60)	< 12	Hospital of Wisconsin Respiratory Score	Inter-observer reliability	NA	ICC	Adequate
Urzúa, 2002	345 (55)	1-24	Tal's score, modified	Inter-observer reliability	NA	Kappa	Poor
Coarasa, 2010	NR	1-24	Respiratory distress scale by the Ministry of Health of Argentina	Concurrent criterion validity	O ₂ saturation	Spearman's r	Excellent

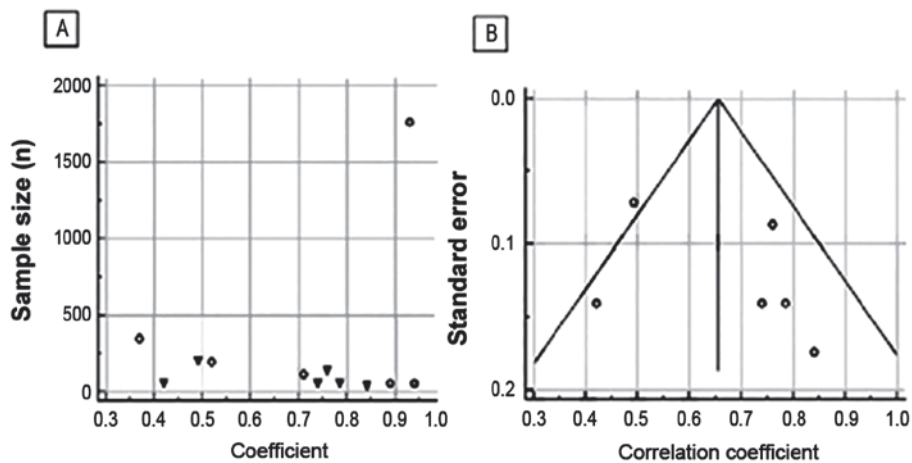
MQ: methodological quality. NR: not reported. NA: not applicable. ICC: intraclass correlation coefficient. O₂: oxygen. CAS: clinical asthma score.

the Respiratory Distress Assessment Instrument (RDAI)^{11,12} (p= 1960, 67.6%), Wood-Downes^{28,29} (modified CAS) (p= 54, 1.9%), the Children’s Hospital of Wisconsin Respiratory Score (CHWRS)¹² (p= 195, 6.7%), and the respiratory distress scale by the Ministry of Health of Argentina (RDSMoHA)¹⁰ (p= 200, 6.9%). Table 2

shows the identified scales, their methodological characteristics, and psychometric properties.

Concurrent and convergent criterion validity: 4 studies (p= 392, 13.5%) established the concurrent criterion validity;^{10,15,27,28} all considered oxygen (O₂) saturation as the reference criterion. One study²⁶ established the concurrent criterion validity of

FIGURE 2. Risk of bias among studies



2.A: Risk of publication bias based on the correlation among validity and reliability indexes and the number of study subjects. Triangles represent studies done to establish validity, and circles, studies done to establish the reliability of the scales for grading the severity of bronchial obstruction.

2.B: Risk of selection bias based on the validity studies established using a funnel plot.

TABLE 2. Characteristics and psychometric properties of identified scales

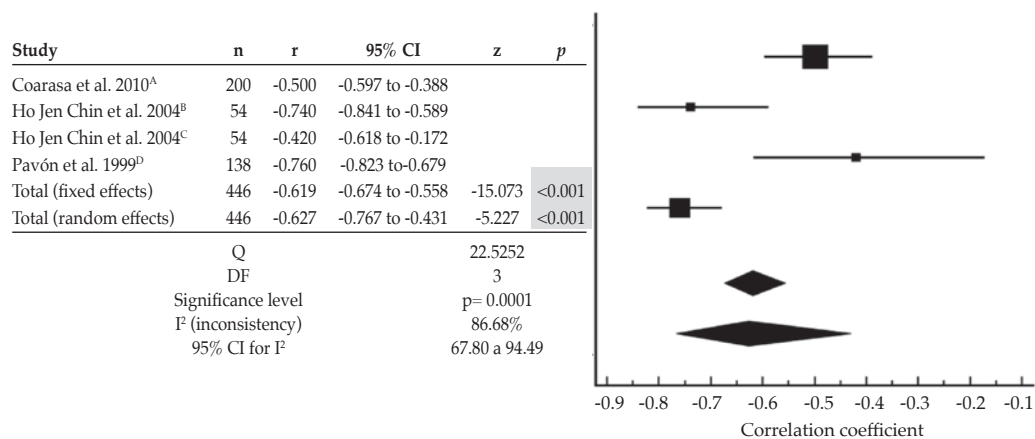
Scale	Assessed items	Scoring range	Validity (r)	Reliability
Kristjansson	RR, retractions, lung sounds, skin color, overall status	0-8	-0.75	ICC: 0.89
Wang’s score	RR, wheezing, retractions, overall status	0-12	-0.42	ICC: 0.99
Tal’s score (original)	RR, wheezing, cyanosis, use of accessory muscles	0-12	NR	Kappa: 0.72
Respiratory distress	Wheezing, retractions	0-17	NR	ICC: 0.93 ICC: 0.30
Wood-Downes (modified CAS)	O ₂ saturation, pulmonary murmur, wheezing, use of accessory muscles, brain function	0-10	0.76	Kappa: 0.89
Children’s Hospital of Wisconsin	Respiratory sounds, dyspnea, retractions, RR, HR, O ₂ requirement, ability to cough up, chest X-ray, lung sounds, surgical status	0-24	NR	ICC: 0.73
Tal’s score, modified by McCallum	RR, wheezing, O ₂ saturation, use of accessory muscles	0-12	NR	Kappa: 0.70
Tal’s score, modified by Pavón	RR by age, wheezing, use of accessory muscles, cyanosis	0-12	-0.76 0.84	Kappa: 0.37
Respiratory distress scale by the Ministry of Health of Argentina	RR, HR, wheezing, chest indrawing	0-12	-0,492	NR

RR: respiratory rate. HR: heart rate. NR: not reported. ICC: intraclass correlation coefficient. O₂: oxygen. r: correlation. Negative coefficients refer to concurrent criterion validity, and positive correlation coefficients, to the reported convergent criterion validity. CAS: clinical asthma score.

two scales, which were considered separately for analysis purposes. Considering the discrepancies among articles in terms of MQ, the random effects model indicated a weighted CC of -0.627 (95% CI: -0.767 to -0.431, $p < 0.001$) (Figure 3). Also, 2 studies ($p = 90$, 3.1%) established the convergent criterion validity using Tal's score³⁰ and the Wood-Downes (modified CAS)²⁹ as reference criterion. In these studies, the random effects model showed a weighted CC of 0.809 (95% CI: 0.721 to 0.871, $p < 0.001$) (Figure 4).

Inter-observer reliability and ICC: Six articles established the inter-observer reliability of the scales for grading the severity of bronchial obstruction. Of these, 3 considered the kappa statistic^{27,28,30} ($p = 511$, 17.6%) and the other 3, the ICC as an index of reliability^{11,12,26} ($p = 2015$, 69.5%). Three articles established the inter-observer reliability of two scales,^{12,27,28} and considered the weighted mean performance of both scales as an index of reliability; the estimated weighted coefficients were 0.500 for kappa and 0.891 for the ICC (Table 3).

FIGURE 3. Concurrent criterion validity for bronchial obstruction scales



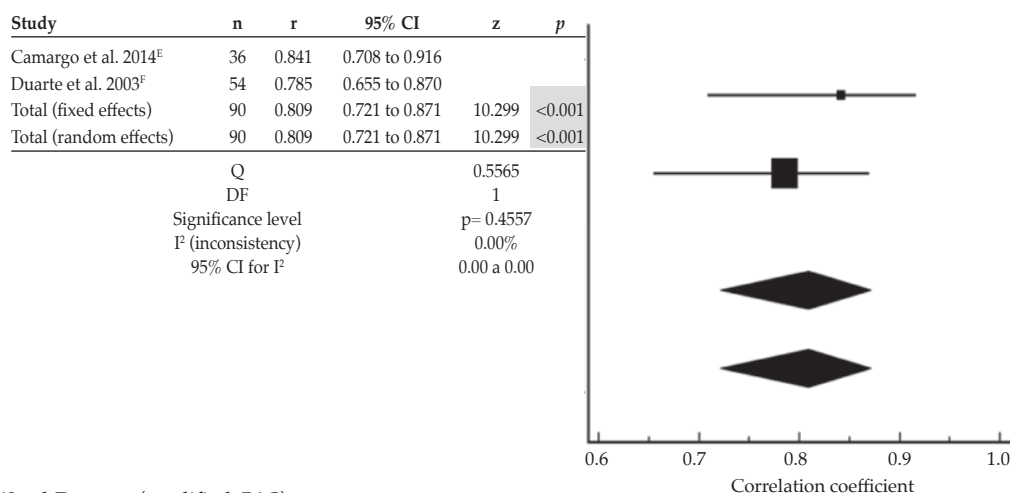
^A Respiratory distress scale by the Ministry of Health of Argentina (RDSMoHA).

^B Kristjansson's scale (KS).

^C Wang's score (WS).

^D Modified Tal's score.

FIGURE 4. Convergent criterion validity for bronchial obstruction scales



^E Wood-Downes (modified CAS).

^F Tal's score.

DISCUSSION

In relation to the evidence of psychometric properties supporting the use of these indirect methods to estimate the severity of bronchial obstruction, it could be said that criterion validity was shown to be moderate to adequate, and the percentage of agreement among observers in relation to the construct (severity of bronchial obstruction) was adequate. The findings observed in the concurrent criterion validity analysis were similar to those of studies done previously using a diagnostic test approach. McCallum et al.²⁸ found that Tal's score had a moderate performance (AUC=0.69), considering peripheral O₂ saturation as the reference standard. On their side, Destino et al.¹² reported a sensitivity and specificity of 65% for the CHWRS, and a ROC curve estimated at 0.68, which is similar to the findings of our study in terms of performance. However, the reference standard used was hospitalization requirement according to the severity of the patient's condition. Also on their side, Puebla et al.³¹ established a sensitivity and specificity of 77% and 88%, respectively, for the modified Tal's score, considering the medical resident's clinical impression as reference standard.

A reference standard widely used to assess the concurrent criterion validity of these scales was O₂ saturation. However, Pavón et al. found that, among the domains included in the modified Tal's score, cyanosis showed the lowest correlation level with peripheral saturation ($r = -0.38$). On the contrary, studies targeted at establishing the scale's internal consistency

recorded acceptable Cronbach's alpha values²⁷ (cyanosis: 0.75, peripheral saturation: 0.72). In relation to convergent criterion validity, it was assessed based on the correlation between two scale scores, one of which is selected as reference standard given its high quality psychometric properties as determined in previous studies.^{28,29}

One of the study limitations is that most studies included in this review had a reasonable MQ, mainly due to weaknesses in their methodological design and conduction, which may affect the internal validity of this study's conclusions. Also, a high percentage of heterogeneity was verified in the concurrent criterion validity meta-analysis, possibly because of the variation in the reference criteria used. It is not possible to rule out the existence of a selection bias in those studies that established this measurement property, which is consistent with what was observed in the funnel plot, where only one of the articles (analyzing the concurrent criterion validity) was outside the confidence limit. In addition, for most studies, observers were trained on how to apply the scale, so the criterion validity coefficients were probably overestimated.

The fact that the severity of bronchial obstruction is adequately detected using several different methods indicates that such feature is real; however, the MQ of studies should support the validity of such conclusions. Therefore, further studies with an improved MQ should be conducted to assess the properties of this measurement instruments.

TABLE 3. Weighted average of reliability coefficients corresponding to the scales identified in the primary articles. N= 6

Author, year	N	Kappa	β
McCallum, 2012	112	0.71*	0.156
Urzúa, 2002	345	0.37	0.250
Duarte, 2012	54	0.89	0.094
Weighted total	511		0.500
Author, year	N	ICC	β
Ho Jen Chin, 2004	54	0.94*	0.025
Destino, 2012	196	0.52*	0.051
Fernández, 2015	1765	0.93	0.815
Weighted total	2015		0.891

β : weighted coefficients. N: sample size.

ICC: intraclass correlation coefficient.

* The reliability coefficient corresponds to the weighted average of both scales.

REFERENCES

- González de Dios J, Ochoa Sangrador C. Conferencia de Consenso sobre bronquiolitis aguda (I): metodología y recomendaciones. *An Pediatr (Barc)* 2010;72(3):221.e1-33.
- OMS. Medidas de Control de Infecciones en la Atención Sanitaria de Pacientes con Enfermedades Respiratorias Agudas en Entornos Comunitarios. Organización Mundial de la Salud. 2009. [Accessed on: November 8th, 2016]. Available at: http://www.who.int/csr/resources/publications/WHO_HSE_GAR_BDP_2009_1/es/.
- Chile. Ministerio de Salud. Guía Clínica AUGÉ Infección Respiratoria Aguda Baja de Manejo Ambulatorio en menores de 5 años. Santiago: Minsal, 2013. [Accessed on: November 8th, 2016]. Available at: <http://www.minsal.cl/portal/url/item/7220fdc4341244a9e04001011f0113b9.pdf>.
- American Academy of Pediatrics Subcommittee on Diagnosis and Management of Bronchiolitis. Diagnosis and management of bronchiolitis. *Pediatrics* 2006;118(4):1774-93.
- Ralston SL, Lieberthal AS, Meissner HC, Alverson BK, et al. Clinical practice guideline: the diagnosis, management, and prevention of bronchiolitis. *Pediatrics* 2014;134(5):e1474-502.
- McConnochie KM. Bronchiolitis. What's in the name? *Am J Dis Child* 1983;137(1):11-3.

7. Bekhof J, Reimink R, Brand PL. Systematic review: insufficient validation of clinical scores for the assessment of acute dyspnoea in wheezing children. *Paediatr Respir Rev* 2014;15(1):98-112.
8. Ramos Fernández JM, Córdón Martínez A, Galindo Zavala R, Urda Cardona A. Validación de una escala clínica de severidad de la bronquiolitis aguda. *An Pediatr (Barc)* 2014;81(1):3-8.
9. Postiaux G, Louis J, Labasse HC, Gerroldt J, et al. Evaluation of an alternative chest physiotherapy method in infants with respiratory syncytial virus bronchiolitis. *Respir Care* 2011;56(7):989-94.
10. Coarasa A, Giugno H, Cutri A, Loto Y, et al. Validación de una herramienta de predicción clínica simple para la evaluación de la gravedad en niños con síndrome bronquial obstructivo. *Arch Argent Pediatr* 2010;108(2):116-23.
11. Fernández RM, Plint AC, Terwee CB, Sampaio C, et al. Validity of bronchiolitis outcome measures. *Pediatrics* 2015;135(6):e1399-408.
12. Destino L, Weisgerber MC, Soung P, Bakalanski D, et al. Validity of respiratory scores in bronchiolitis. *Hosp Pediatr* 2012;2(4):202-9.
13. Martínón-Torres F, Rodríguez-Núñez A, Martínón-Sánchez JM. Heliox Therapy in Infants With Acute Bronchiolitis. *Pediatrics* 2009;109(1):68-73.
14. Tal A, Bavilski C, Yohai D, Bearman JE, et al. Dexamethasone and salbutamol in the treatment of acute wheezing in infants. *Pediatrics* 1983;71(1):13-8.
15. Pavon D, Castro-Rodríguez JA, 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.
16. McDowell I, Newell C. The Theoretical and technical Foundations of Health Measurement. In: *Measuring Health: A Guide to Rating Scales and Questionnaires*. 2nd ed. New York: Oxford University Press; 1996. Pages 10-46.
17. Van Miert C, Abbott J, Verheoff F, Lane S, et al. Development and validation of the Liverpool infant bronchiolitis severity score: a research protocol. *J Adv Nurs* 2014;70(10):2353-62.
18. Liu LL, Gallaher MM, Davis RL, Rutter CM, et al. Use of a respiratory clinical score among different providers. *Pediatr Pulmonol* 2004;37(3):243-8.
19. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 2010;8(5):336-41.
20. Liberati A, Altman DG, Tetzlaff J, Mulrow C, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med* 2009;151(4):W65-94.
21. Liberati A, Altman DG, Tetzlaff J, Mulrow C, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009;339:b2700.
22. Mokkink LB, Terwee CB, Patrick DL, Alonso J, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63(7):737-45.
23. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21(4):651-7.
24. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;343:d4002.
25. Rodríguez I, Zambrano L, Manterola C. Validez de criterio de las escalas de medición de esfuerzo percibido en niños sanos: una revisión sistemática y metaanálisis. *Arch Argent Pediatr* 2016;114(2):120-8.
26. Chin HJ, Seng QB. Reliability and validity of the respiratory score in the assessment of acute bronchiolitis. *Malays J Med Sci* 2004;11(2):34-40.
27. McCallum GB, Morris PS, Wilson CC, Versteegh LA, 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.
28. Duarte-Dorado DM, Madero-Orostegui DS, Rodríguez-Martínez CE, 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.
29. Camargo Crespo C. Validación de una escala de severidad en bronquiolitis viral aguda en una población de lactantes atendidos en el hospital de la Misericordia. Bogotá: Universidad Nacional de Colombia; 2014. [Accessed on: November 8th, 2016]. Available at: <http://www.bdigital.unal.edu.co/12343/1/5599350.2014.pdf>.
30. Urzúa S, Duffau G, Zepeda G, Sagredo S. Estudio de concordancia clínica en educandos de pre y postítulo en Pediatría: Puntaje de Tal. *Rev Chil Pediatr* 2002;73(5):471-7.
31. 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.