

The efficacy of the OSA-18 as a waiting list triage tool for OSA in children

Lisa M. Walter^{1,2} · Sarah N. Biggs^{1,2} · Natascha Cikor¹ · Kathy Rowe^{3,4} · Margot J. Davey^{1,2,5} · Rosemary S. C. Horne^{1,2} · Gillian M. Nixon^{1,2,5}

Received: 15 September 2015 / Revised: 14 November 2015 / Accepted: 23 November 2015 / Published online: 15 December 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract

Purpose Sleep disordered breathing (SDB) is common in children, resulting in extensive waiting lists for specialist clinics. There is an urgent need for a valid method of triaging patients and the OSA-18, a disease-specific tool, is an attractive candidate for this role. We aimed to examine the OSA-18 as a measurement tool in detail and to determine whether the score or aspects of it could be used as a screening tool for SDB in children.

Methods Retrospective analysis of 582 children (6 months to 16.4 years)—216 underwent overnight PSG and 366 overnight oximetry. Confirmatory factor analysis (CFA) and exploratory factor analysis (EFA) were conducted. Receiver operating characteristic curve analysis assessed the diagnostic accuracy of the factors for the presence of OSA. Rasch analysis was used to assess the structure of the items (1–18) and categories of response (Likert scale).

Results The CFA with a forced five-factor structure, revealed three factors with Eigenvalues >1, and explained 73.7 % of the variance. EFA resulted in a two-factor structure, explaining 60.3 % of the variance. Assessment of sensitivity and specificity showed a high false-positive rate, irrespective of the factor structure tested. Rasch analysis showed poor discrimination between adjacent categories on the Likert scale.

Conclusion This study confirmed that the predictive value of the OSA-18 for SDB severity is weak. Some questions perform better than others statistically, and the seven categories of response introduce significant statistical noise, raising the possibility that modification of the OSA-18 may improve its performance in the prediction of OSA severity.

Keywords Pediatric · Obstructive sleep apnea · Quality of life

Introduction

Obstructive sleep apnea (OSA) is very common, affecting 1–3 % of children [1]. First-line treatment for moderate-severe OSA is surgical removal of the tonsils and adenoids (adenotonsillectomy, AT). Given the limited availability of specialist pediatric sleep physicians, pediatric polysomnography (PSG) services, and otorhinolaryngology services, a simple way of estimating the severity of OSA is attractive as a triage or health care system tool.

There are a number of screening instruments available that assess frequency and characteristics of nocturnal respiratory disturbance in children, with varying levels of validity and reliability [2]. The assumption of a correlation between worse health-related quality of life (HRQOL) and OSA severity has led to the use of the OSA-18 as a surrogate diagnostic tool to estimate OSA severity [3, 4]. The OSA-18 is a 7 category Likert-type scale with 18 questions, developed as a practical

Lisa M. Walter and Sarah N. Biggs are co-first authors.

✉ Lisa M. Walter
lisa.walter@hudson.org.au

- ¹ The Ritchie Centre, Hudson Institute of Medical Research, Melbourne, Level 5 Monash Medical Centre, 246 Clayton Road, Clayton, VIC 3168, Australia
- ² Department of Paediatrics, Monash University, Melbourne, Australia
- ³ Murdoch Children's Research Institute, University of Melbourne, Melbourne, Australia
- ⁴ Department of General Medicine, Royal Children's Hospital, Melbourne, Australia
- ⁵ Melbourne Children's Sleep Centre, Monash Children's Hospital, Melbourne, Australia

office-based questionnaire to determine the impact of OSA on children's HRQOL [5]. It is based on five domains: sleep disturbance, physical symptoms, emotional symptoms, daytime function, and caregiver concerns. In the original description of the score, Franco and colleagues reported that HRQOL was negatively correlated with disease severity [5]. They suggested that scores less than 60 reflected a small impact on HRQOL, scores between 60 and 80 a moderate impact, and scores above 80 a high impact.

While there are studies that have proposed the OSA-18 as a valid method for discriminating OSA severity and determining improvement following treatment [6–8], there are also studies which have demonstrated no correlation between the OSA-18 score and disease severity [9–12]. However, most of these studies evaluated the OSA-18 total symptom score, without considering either individual domains or questions. In addition, the original development of the domains was based on clinical experience and has not, to our knowledge, been factor analyzed to determine construct validity. It may be that some items or questions in the OSA-18 relevant to HRQOL are not a valid assessment of OSA severity, thus reducing the efficacy of the total symptom score as a screening tool. The use of a 7-category Likert scale has also not been validated and may be a source of considerable statistical “noise.” In this study, we aimed to determine the most statistically robust factor structure of the OSA-18 and identify aspects of the questionnaire that may contribute to its predictive ability in assessing disease severity in children. We hypothesized that (1) the total score of the OSA-18 would not be a valid method of distinguishing between OSA severities and (2) some items of the OSA-18 would be more robust than others and could improve the efficacy of the OSA-18 as a screening tool for OSA severity in children.

Methods

Ethical approval for this study was granted by the Monash Health and Monash University Human Research Ethics Committees. Written informed consent was obtained from parents at the time of the clinical testing for use of data pertaining to the diagnostic PSG or home oximetry for research purposes. There was no monetary incentive for participation.

This is a retrospective analysis of OSA-18 data from 582 children (6 months to 16.4 years) attending the Melbourne Children's Sleep Centre for assessment of SDB and 41 non-snoring control children recruited from the community. Two hundred sixteen (37 %) children underwent overnight polysomnography (PSG) and further 366 (63 %) had overnight oximetry. At the time of the PSG or oximetry, children were otherwise healthy and not undergoing treatment with nasal steroids, leukotriene receptor antagonists, or antibiotics.

The OSA-18 questionnaire was completed by a parent on the day or evening of the overnight sleep test (PSG or oximetry).

PSG was conducted using established clinical protocols [13]. Briefly, electroencephalogram, left and right electrooculogram, submental electromyogram, left and right anterior tibialis muscle electromyogram, electrocardiogram, respiratory effort measured using respiratory inductance plethysmography, transcutaneous carbon dioxide, oxygen saturation, nasal pressure, and oronasal airflow were recorded (Series E Sleep System Compumedics, Melbourne, Australia). Sleep and respiratory events were scored according to standard clinical guidelines [14, 15]. We have demonstrated that there is no significant clinical difference between studies scored by the different rules used over the period of this analysis [16]. Oximetry was performed overnight at home using Masimo Radical 7 oximeters set at 2 s averaging (Masimo Corporation, CA, USA).

Statistical analysis

OSA-18 questionnaires were analyzed from all 582 children (age range 6 months to 16.4 years) for confirmatory factor analysis (CFA) to confirm the five-factor structure of the original OSA-18 by forcing the five-factor structure using all items of the questionnaire. Exploratory factor analysis (EFA) was conducted only on the 216 (37 %) children who underwent PSG to explore the best-fit factor and item structure, to determine if a more statistically robust factor structure existed for identification of SDB severity [17]. To account for the error variance within items and allow for factor correlation, principal axis factoring (PAF) and oblique rotation (Promax) were used in both the CFA and EFA [18–20]. Initial factor extraction in the EFA was based on the Kaiser test of Eigenvalues >1 , which is the point at which the factor explains more of the variance than the individual items within each factor [21]. Subsequent factor extractions were based on examination of item loadings and trivial factors [21]. Any item with a factor loading <0.4 were removed. Any factors containing two or fewer items (trivial factors) were also removed. Factor reliability was assessed using Cronbach's coefficient [22, 23].

The original OSA-18 sub-scales and the factors derived from the CFA and EFA were assessed against an OAH >5 events/h (moderate-severe OSA) and an OAH >2 events/h (includes cases of mild OSA) in the sub-sample of children who underwent PSG ($N=216$, age range 2 to 12.5 years). An OAH >2 events/h was chosen based on previous literature showing that treatment at this level improves OSA symptoms significantly more than not treating [24]. A second cutoff of OAH >5 events/h was used to determine whether discrimination differed with increasing severity. Receiver operating characteristic (ROC) curve analyses were then used to assess

the diagnostic accuracy of each of the sub-scales and factors in the clinical sample of children ($N = 175$) and non-snoring controls ($N = 41$) who underwent overnight PSG. Diagnostic accuracy is represented by the area under the curve (AUC) and describes the overall probability that the scale will accurately identify a positive case. An AUC value of 1 indicates 100 % accuracy, and an AUC of 0.5 indicates that the probability of a positive case being identified is no better than chance. Cutoff scores for the original OSA-18 and the factors extracted from the EFA were determined based on the best proportional balance of sensitivity and specificity.

Further analysis of the properties of the questionnaire was performed using the Rasch model [25, 26]. A Rasch measurement scale was constructed using Quest [27]. Rasch analysis consists of locating category thresholds for each OSA-18 item and a thorough analysis of fit of the data to the model, overall and at the item level. This analysis served to identify whether the seven questionnaire response categories (from “none of the time” to “all of the time”) were distinct from each other and how each of the 18 items related to the construct measured with the questionnaire.

Results

Demographics

The mean age of the entire cohort ($n = 582$) was 4.5 ± 2.6 years (STD) and ranged from 6 months to 16.4 years (85 % aged ≤ 6 years, 42 % male). There was no significant difference between the mean age of the children who had oximetry only ($n = 366$, 4.4 ± 3.1 years, range 6 months to 16.4 years, 34 % male), and the children who had PSG ($n = 216$, 4.7 ± 1.5 years, range 2 to 12.5 years, 57 % male). There were proportionately more females in the oximetry only group compared to PSG group (73 vs 27 %, respectively, $P < 0.001$). The proportion of males was equivalent across the groups (50 vs 50 %, respectively). Subjects who had PSG were grouped by severity of OSA according to OAHl as follows: controls $n = 41$ (median OAHl 0/h, range 0–0.8); primary snoring $n = 78$ (median OAHl 0.2/h, range 0–1); mild OSA $n = 50$ (median OAHl 2.8/h, range 1.1–5), and moderate-severe OSA $n = 46$ (median OAHl 11.3/h, range 5.1–61.2).

Confirmatory factor analysis

The confirmatory factor analysis ($N = 582$), forcing a five-factor structure, revealed three factors with Eigenvalues >1 and no item with a factor loading <0.4 (Table 1). The total variance explained by the five factors was 73.7 %. This analysis revealed three complex items loading equally on more than one factor—concerned child not getting enough air

(question 16), mouth breathing because of obstruction (question 5), and poor attention span or concentration (question 13). This indicates that the items are not discriminatory (i.e., are measuring more than one construct) and thus introduce covariance across factors.

Overall, the factor structure revealed in this analysis reflects that originally designed by Franco et al. [5]; however, only three factors had Eigenvalues >1 . Factor 1 includes all items from the sleep disturbance sub-scale, plus one item from physical symptoms (mouth breathing) and one from caregiver concerns (concerned child not getting enough air). Factor 2 reflects emotional distress, factor 3 physical symptoms, factor 4 caregiver concern, and factor 5 daytime function.

Exploratory factor analysis

The initial PAF ($N = 582$) on all 18 items of the OSA-18 resulted in a three-factor structure, based on Eigenvalues >1 , accounting for a total of 57.8 % of the variance. Examination of the pattern matrix revealed two items with loadings <0.4 : difficulty getting up in the morning (question 14) and difficulty swallowing foods (question 8). These were removed from subsequent analyses.

A second PAF on the remaining 16 items resulted in a two-factor structure, based on Eigenvalues >1 , accounting for 55.5 % of the variance. The pattern matrix revealed one item with a loading <0.4 : excessive daytime sleepiness (question 12). This factor solution also revealed one trivial factor (containing only two items) with an Eigenvalue <1 . As a result, these items—nasal discharge or runny nose and frequent colds or upper respiratory infection—were removed from further analysis.

The final PAF (13 items) resulted in a two factor structure explaining 60.3 % of the variance. Table 2 shows the factor loadings, Eigenvalues, variance explained, and internal consistency (Cronbach's α) of each factor. This table also shows the reliability of the factor if individual items were removed. The lack of change, or slight decrease in alpha if individual items are removed, indicates that this two-factor structure is the most robust for these items.

Receiver operating characteristic curves

Table 3 shows the results from the ROC analyses ($N = 216$) for each of the domains on the original OSA-18 sub-scales, the five-factor structure of the CFA and the two-factor structure of the EFA against an OAHl >5 and 2 events/h, respectively. Due to the similarity between the original domain structure and the results of the CFA, cutoff scores were not assessed for the factors of the CFA. The domain with the greatest diagnostic accuracy was caregiver concern (AUC = 0.63); however, this still represents poor discrimination. While many of the factors

Table 1 Factor structure, item loading, Eigenvalues, percentage of variance explained, and reliability analysis of OSA-18 following confirmatory factor analysis of the five-factor structure

Item	Item loading	Eigenvalue	% of variance explained	<i>a</i>	if item removed
Factor 1 (sleep disturbance)		7.8	43.58	0.88	
Breath holding spells or pauses in breathing	0.98				0.85
Loud snoring	0.91				0.86
Choking or gasping	0.84				0.85
Concern child is not getting enough air	0.56				0.85
Mouth breathing because of obstruction	0.52				0.88
Restless sleep or frequent awakenings	0.48				0.88
Factor 2 (emotional distress)		2.3	12.72	0.89	
Discipline problems	0.95				0.84
Mood swings or temper tantrums	0.92				0.86
Aggressive or hyperactive problems	0.91				0.84
Poor attention span or concentration	0.48				0.91
Factor 3 (physical symptoms)		1.3	7.20	0.80	
Nasal discharge or runny nose	0.98				0.70
Frequent colds or URI	0.94				0.69
Mouth breathing because of obstruction	0.52				0.77
Difficulty swallowing	0.42				0.81
Factor 4 (caregiver concern)		0.9	5.54	0.89	
Made you frustrated	0.92				0.87
Interfered with ability to perform daily activities	0.91				0.85
Worry about child's general health	0.62				0.85
Concern child not getting enough air	0.50				0.86
Factor 5 (daytime functioning)		0.8	4.67	0.73	
Difficulty getting up in the morning	0.97				0.72
Excessive daytime sleepiness	0.73				0.60
Poor attention span or concentration	0.44				0.62

showed high sensitivity, specificity was very poor. As can be seen in Table 3, while the specified cutoffs adequately identify children with an OAH1 > 5 or >2 events/h, they also have a high probability of producing a false positive (low specificity).

Rasch analysis

Given the poor performance of the OSA-18 and factors on the ROC analysis, we decided to undertake Rasch analysis to examine any potential redundancy within the questionnaire construct. The Rasch analysis revealed that a maximum of 4 response categories for each item was reliably scored, indicating that 7 response categories are not necessary because respondents are not able to discriminate between some of the adjacent categories on the Likert scale (e.g., response 1 was not systematically different to response 2). A thorough analysis of fit of the data to the measurement model in a series of analyses based on

statistical fit indicators confirmed that the 18 items belong to a single scale construct (i.e., are measuring the same construct) after collapsing of categories for each item. This analysis reduced the total number of score points on the survey from 108 to 38, with the survey reliability only decreasing from 0.92 retaining 7 categories to 0.85 when the categories were collapsed. Rasch analysis indicates that questions 2 (breath holding spells), 15 (worry about child's general health), 16 (concern child is not getting enough air), 17 (interfered with ability to perform daily activities), and 18 (made you frustrated) were the most discriminating items. This finding is consistent with caregiver concern domain (questions 15–18) having the greatest diagnostic accuracy in the ROC analysis.

The Rasch analysis has shown that there is a redundancy in the response categories for each item. Despite eliminating redundant scores by collapsing categories, differences between mean Rasch measures are observed only in controls (−4.8 STD 0.75) and not between the SDB

Table 2 Factor structure, item loading, Eigenvalues, percentage of variance explained, and reliability analysis of OSA-18 following third principal axis factoring extraction

Item	Item loading	Eigenvalue	% of variance explained	<i>a</i>	if item removed
Factor 1 (score range 9–63)		6.1	46.7	0.91	
Concern child is not getting enough air	0.92				0.89
Breath holding spells or pauses in breathing	0.81				0.90
Worry about child's general health	0.81				0.90
Choking or gasping	0.79				0.90
Loud snoring	0.76				0.91
Mouth breathing because of obstruction	0.65				0.91
Restless sleep or frequent awakenings	0.57				0.91
Interfered with ability to perform daily activities	0.56				0.90
Made you frustrated	0.50				0.91
Factor 2 (score range 4–28)		1.8	13.7	0.89	
Discipline problems	0.95				0.84
Aggressive or hyperactive problems	0.90				0.84
Mood swings or temper tantrums	0.84				0.86
Poor attention span or concentration	0.60				0.91

groups (mean \pm STD 2.1 ± 1.4 , 1.6 ± 1.3 , and 1.6 ± 1.2 for PS, mild, and MS OSA, respectively).

Discussion

This is the first study to assess the validity of the OSA-18 questionnaire in detail using statistically robust methods of factor analysis and Rasch analysis. The structure of the five factors identified by the confirmatory factor analysis substantially reflects that of the original OSA-18 domains, although only three factors from the CFA showed Eigenvalues >1 . As we hypothesized, the domains of the OSA-18 are a poor predictor of OSA severity. Although some questions performed better than others statistically, a sub-set of questions and domains could not be identified that would be suitable to be used as an abbreviated questionnaire to determine OSA severity. Analysis did however identify aspects of the questionnaire that can guide refinement of the tool to potentially improve its specificity as a triage tool—particularly, a reduction in response categories and selection of fewer questions.

Rasch analysis confirms that the 18 items of the OSA-18 measure a single construct (the scale conducted with the questionnaire data is uni-dimensional). Based on extensive previous research showing the sensitivity of the OSA-18 to discriminate between children with OSA and controls [7, 28–31], we can be confident that it is assessing HRQOL in children with SDB [5]. However, the OSA-18 is highly sensitive but very poorly specific for the presence of either mild OSA as indicated by an OAH > 2 events/h (total score sensitivity 95 %, specificity 30 %) or moderate-severe OSA as

indicated by an OAH > 5 events/h (total score sensitivity 93 %, specificity 25 %). Sensitivity relates to the proportion of children with OSA confirmed by PSG, who the OSA-18 predicted would have OSA. Specificity refers to the proportion of children without OSA confirmed by PSG, who the OSA-18 predicted would not have OSA. The predictive value of the positive test is the proportion of children who the OSA-18 predicted would have OSA and who actually had OSA confirmed by PSG. That is, while the OSA-18 will correctly identify the majority of children who actually have OSA at the designated level of severity (93 % OAH > 5 ; 95 % OAH > 2), it will also predict that a substantial number of children who actually do not have OSA will have the disease (75 % OAH > 5 ; 70 % OAH > 2)—a high false-positive rate. If the main purpose of a triage tool for OSA severity is to differentiate those children most likely to have severe OSA from those with no or less severe OSA, the OSA-18 total score is unsuited for this purpose. However, it must be acknowledged that the current study included children across a large age range and parental demographic data were not available for analyses. Future research examining the efficacy of the OSA-18 in particular age groups and the influence of parental characteristics, as the instrument is parent-report, may be of benefit.

The original validation of the OSA-18 questionnaire by Franco et al. was made against the respiratory disturbance index derived by PSG performed during 90 min of daytime napping in children who had been sleep deprived the preceding night [5]. Daytime naps themselves have a poor negative predictive value for OSA [32]. The finding of a correlation between the OSA-18 and the respiratory

Table 3 ROC curve analysis of the original OSA-18 sub-scales, the two-factor structure of the exploratory factor analysis, and the five-factor structure of the confirmatory factor analysis against OAH1 > 5 andOAH1 > 2. Scores with the greatest diagnostic confidence are presented with ratings of sensitivity and specificity ($N = 216$)

	OAH1 > 5		OAH1 > 2			OAH1 > 5		OAH1 > 2	
	AUC	AUC	Cutoff	Sens	Spec	% positive cases correctly identified	% negative cases correctly identified	% positive cases correctly identified	% negative cases correctly identified
Original OSA-18 sub-scales									
Sleep disturbance	0.55	0.59*	9	0.86	0.15	77	33	84	39
Physical symptoms	0.49	0.57	5	0.91	0.05	86	13	91	16
Emotional symptoms	0.53	0.56	5	0.89	0.05	84	16	87	17
Daytime function	0.61*	0.65**	5	0.96	0.17	86	33	88	38
Caregiver concern	0.63**	0.63**	6	0.91	0.18	86	32	89	37
Total	0.66***	0.71***	35	0.93	0.25	93	25	95	30
CFA factor structure									
Factor 1 (sleep disturbance)	0.56	0.59*							
Factor 2 (emotional distress)	0.56	0.59*							
Factor 3 (physical symptoms)	0.49	0.57							
Factor 4 (caregiver concern)	0.63*	0.63**							
Factor 5 (daytime functioning)	0.61*	0.65***							
EFA factor structure									
Factor 1 (symptoms)	0.58	0.61*	17	0.93	0.14	93	29	93	33
Factor 2 (behavior)	0.56	0.59*	9	0.84	0.18	80	29	81	31
Total	0.66**	0.70***	30	0.91	0.37	89	32	91	37

disturbance index in the 61 children who participated in the original validation study thus has limited capability to be extrapolated to the wider population. Our findings support previous research using nocturnal PSG that reported the OSA-18 had poor validity in predicting OSA in children [10–12]. Ishman et al. [10] and Borgstrom et al. [11] compared the total symptom score with the respiratory disturbance index, the obstructive apnea hypopnea index and the apnea hypopnea index on overnight PSG and calculated sensitivity, specificity, and positive and negative predictive values. Similar findings to our own were reported in these studies, irrespective of age, race, gender, BMI, or the different OSA severity cutoff levels used in the analyses. Constantin et al. [12] took a slightly different approach and compared OSA-18 questionnaires to the McGill oximetry score, which provides a validated approach to oximetry studies for OSA, such that in children higher oximetry scores were associated with higher apnea hypopnea index, desaturation index, lower SaO₂, and higher respiratory arousal index [33]. They calculated sensitivity and negative predictive values for the OSA-18 to detect an abnormal McGill oximetry score and concluded that the OSA-18 did not accurately detect which children had an abnormal McGill oximetry score and could not exclude children with moderate-to-severe OSA.

Furthermore, Mitchell et al. [34] reported that there was no correlation between the OSA-18 total symptom score and the respiratory disturbance index. Our study expanded on the previous research with the use of factor analysis and Rasch analysis, which confirmed the validity of the structure of the questionnaire. Additionally, factor analysis identified questions that did not relate to OSA severity and removed them from further analysis to test the validity of an abbreviated questionnaire.

We had hypothesized that some items would perform better than others in predicting OSA and an abbreviated questionnaire would be an effective screening tool for OSA severity in children. However, in contrast to our hypothesis, removing some questions did not add substantially to the variance explained and did little to improve the performance of the score as a diagnostic test for severity of OSA, as 68 % of children without OSA would still be incorrectly identified having OSA. In a much less robust analysis, Borgstrom et al. found a weak but significant correlation between the score for the sleep disturbance domain and the AHI [11]. However, in this study, there were a very small number of children who had an AHI > 80/h and the authors suggest that the significant correlation was probably only the result of these few children [11]. Recognizing this, the authors concluded that

the OSA-18 had low diagnostic power. We have identified by Rasch analysis that respondents are unable to discriminate between several of the adjacent points on the Likert scale of the OSA-18, for example, between “Hardly any of the time” and “A little of the time” or between “A good bit of the time” and “Most of the time.” Thus, the 7-category Likert scale is a source of statistical “noise” that will contribute to the inability of the total score to predict OSA severity. Using Rasch analysis to collapse these categories however did not by itself improve the scale’s ability to discriminate between different severities of OSA, although it did identify items that were more likely to discriminate high and low scores on the questionnaire.

Conclusion

The OSA-18 was designed to measure disease-specific quality of life but has been adopted as a potential screening tool for OSA severity. This study showed that a forced five-factor CFA reflected the five original factors (sub-scales) of the OSA-18; however, only three of these factors were meaningful to the construct. Using the most statistically robust factor structure of the OSA-18 questionnaire as determined by the EFA did not result in a score that reliably predicted OSA severity, and thus, neither the total score nor individual domains should be used as a screening tool for OSA severity in children. The Rasch analysis indicated that modifying the response categories and using better performing items warrants further investigation to potentially improve the questionnaire for use as a triage tool.

Acknowledgments The authors would like to thank the children and families who participated in this study and the staff of the Melbourne Children’s Sleep Centre for their support. The authors thank Andrew Stephanou (Australian Council for Educational Research) for performing the Rasch analysis.

Compliance with ethical standards Ethical approval for this study was granted by the Monash Health and Monash University Human Research Ethics Committees. Written informed consent was obtained from parents at the time of the clinical testing for use of data pertaining to the diagnostic PSG or home oximetry for research purposes. There was no monetary incentive for participation.

Funding This study was supported by a National Health and Medical Research Council of Australia project grant (491001 and 1010810) and the Victorian Government’s Operational Infrastructure Support Program.

Conflict of interest All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers’ bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

References

- Lumeng JC, Chervin RD (2008) Epidemiology of pediatric obstructive sleep apnea. *Proc Am Thorac Soc* 5:242–252
- Spruyt K, Gozal D (2011) Pediatric sleep questionnaires as diagnostic or epidemiological tools: a review of currently available instruments. *Sleep Med Rev* 15:19–32
- Cincin A, Sakalli E, Bakirci EM, Dizman R (2014) Relationship between obstructive sleep apnea-specific symptoms and cardiac function before and after adenotonsillectomy in children with adenotonsillar hypertrophy. *Int J Pediatr Otorhinolaryngol* 78:1281–1287
- Ishman SL, Tawfik KO, Smith DF, et al. (2015) Screening for pediatric obstructive sleep apnea before ambulatory surgery. *J Clin Sleep Med* 11(7):751–75;
- Franco RA, Jr., Rosenfeld RM, Rao M (2000) First place–resident clinical science award 1999. Quality of life for children with obstructive sleep apnea. *Otolaryngol Head Neck Surg* 123:9–16
- Sohn H, Rosenfeld RM (2003) Evaluation of sleep-disordered breathing in children. *Otolaryngol Head Neck Surg* 128:344–352
- Hasukic B (2013) OSA-18 survey in evaluation of sleep-disordered breathing in children with adenotonsillar hypertrophy. *Mediev Archaeol* 67:111–114
- Fischer Y, Rettinger G, Dorn M (2006) Long term change in quality of life after adenotonsillectomy for pediatric obstructive sleep disorders. *Laryngorhinootologie* 85:809–818
- Kang KT, Weng WC, Lee CH, Lee PL, Hsu WC (2014) Discrepancy between objective and subjective outcomes after adenotonsillectomy in children with obstructive sleep apnea syndrome. *Otolaryngol Head Neck Surg* 151:150–158
- Ishman SL, Yang CJ, Cohen AP, et al. (2014) Is the OSA-18 predictive of obstructive sleep apnea: Comparison to polysomnography. *Laryngoscope* 125(6):1491–1495
- Borgstrom A, Nerfeldt P, Friberg D (2013) Questionnaire OSA-18 has poor validity compared to polysomnography in pediatric obstructive sleep apnea. *Int J Pediatr Otorhinolaryngol* 77:1864–1868
- Constantin E, Tewfik TL, Brouillette RT (2010) Can the OSA-18 quality-of-life questionnaire detect obstructive sleep apnea in children? *Pediatrics* 125:e162–e168
- Nisbet LC, Yiallourou SR, Biggs SN, et al. (2013) Preschool children with obstructive sleep apnea: the beginnings of elevated blood pressure? *Sleep* 36:1219–1226
- Iber C, Ancoli-Israel S, Chesson A. L, Quan SF (2007) The AASM manual for the scoring of sleep and associated events
- Berry RB, Budhiraja R, Gottlieb DJ, et al. (2012) Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated eventsb. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. *J Clin Sleep Med* 8:597–619
- Nixon GM, Hyde M, Biggs SN, Walter LM, Horne RS, Davey MJ (2014) The impact of recent changes to the respiratory scoring rules in pediatrics. *J Clin Sleep Med* 10:1217–1221
- Henson RK, Roberts JK (2006) Use of exploratory factor analysis in published research: common errors and some comment on improved practice. *Educ Psychol Meas* 66:393–415
- Ford JK, MacCallum RC, Tait M (1986) The application of exploratory factor analysis in psychology: a critical review and analysis. *Personnel Psychol* 39:291–314
- Gorsuch RL (1970) A comparison of biquartimin, maxplane, promax, and varimax. *Educ Psychol Meas* 30:861–872
- Gorsuch RL (1997) Exploratory factor analysis: its role in item analysis. *J Pers Assess* 68:532–560

21. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ (1999) Evaluating the use of exploratory factor analysis in psychological research. *Psychol Methods* 4:272–299
22. Cronbach LJ (1951) Coefficient alpha and the internal structure of tests. *Psychometrika* 16:297–334
23. Henson RK (2001) Understanding internal consistency reliability estimates: a conceptual primer on coefficient alpha. *Meas Eval Couns Dev* 34:177–189
24. Marcus CL, Moore RH, Rosen CL, et al. (2013) A randomized trial of adenotonsillectomy for childhood sleep apnea. *N Engl J Med* 368:2366–2376
25. Andrich D (1988) Rasch models for measurement. SAGE Publications CA, USA
26. Bond T, Fox M (2015) Applying the rasch model. Fundamental measurement in the human sciences, 3rd edn. Routledge NY, USA
27. Adams RJ, Khoo ST (1999) ACER QUEST: the interactive test analysis system Australian Council for Educational Research. Camberwell, Victoria
28. Goldstein NA, Fatima M, Campbell TF, Rosenfeld RM (2002) Child behavior and quality of life before and after tonsillectomy and adenoidectomy. *Arch Otolaryngol Head Neck Surg* 128:770–775
29. Mohsen N, Susan A, Shahin B, Soheila D (2014) Sleep related quality of life before and after adenotonsillar surgery in pediatric population. *Int J Pediatr Otorhinolaryngol* 78:330–333
30. Silva VC, Leite AJ (2006) Quality of life in children with sleep-disordered breathing: evaluation by OSA-18. *Braz J Otorhinolaryngol* 72:747–756
31. Gomes Ade M, Santos OM, Pimentel K, et al. (2012) Quality of life in children with sleep-disordered breathing. *Braz J Otorhinolaryngol* 78:12–21
32. Saeed M, Keens T, Stabile M, Bolokowicz J, Davidson Ward S (2000) Should children with suspected obstructive sleep apnea syndrome and normal nap sleep studies have overnight sleep studies? *Chest* 118:360–365
33. Nixon GM, Kermack AS, Davis GM, Manoukian JJ, Brown KA, Brouillette RT (2004) Planning adenotonsillectomy in children with obstructive sleep apnea: the role of overnight oximetry. *Pediatrics* 113:e19–e25
34. Mitchell RB, Kelly J, Call E, Yao N (2004) Quality of life after adenotonsillectomy for obstructive sleep apnea in children. *Arch Otolaryngol Head Neck Surg* 130:190–194