Sleep and Breathing (2020) 24:1349–1356
https://doi.org/10.1007/s11325-019-01964-7

SLEEP BREATHING PHYSIOLOGY AND DISORDERS • ORIGINAL ARTICLE

Refining screening questionnaires for prediction of sleep apnea severity in children

Amal Isaiah[1] · Meryam Shikara[1] · Kevin D. Pereira[1] · Gautam Das[2]

Received: 8 July 2019 / Revised: 10 October 2019 / Accepted: 22 October 2019 / Published online: 27 November 2019

Abstract

Purpose Screening instruments are poor predictors of the severity of pediatric obstructive sleep apnea (OSA). We hypothesized that their performance could be improved by identifying and eliminating redundant features.

Methods Baseline scores from three screening questionnaires for pediatric OSA were obtained from the Childhood Adenotonsillectomy Trial (CHAT). The questionnaires included the (i) modified Epworth sleepiness scale (ESS), (ii) the sleep-related breathing disorders subscale of the pediatric sleep questionnaire (PSQ), and the (iii) obstructive sleep apnea-18 (OSA-18) scale. Key features from each questionnaire were identified using variable selection methods. These selected features (SF) were then assessed for their ability to predict the severity of OSA, measured by the apnea-hypopnea index (AHI) and oxygen desaturation index (ODI). In addition, prediction performance of SF was also calculated for AHI > 5 and > 10 and ODI > 5 and > 10, respectively.

Results Four hundred fifty-three children aged 5–10 years were included. The majority of the pairwise correlations among the items within the 3 screening questionnaires were statistically significant. The prediction of AHI and ODI by overall questionnaire scores was poor. Four-item SF, comprising *apneic pauses, growth problems, mouth breathing,* and *obesity* predicted AHI and ODI significantly better than each of the individual questionnaires. Furthermore, SF also predicted AHI > 5 and > 10, as well as ODI > 5 and > 10 significantly better than the original questionnaires.

Conclusions Elimination of redundant items in screening questionnaires improves their prediction performance for OSA severity in children with high pre-test probability for the condition.

Keywords Pediatric obstructive sleep apnea · Screening questionnaires · Prediction

Introduction

Obstructive sleep apnea (OSA) results from intermittent partial or complete upper airway obstruction that disrupts normal sleep patterns and gas exchange [1]. The prevalence of OSA in children is 1 to 4% [2]. Adverse sequelae of untreated OSA on children's health, behavior, and quality of life are well-described, emphasizing the importance of early diagnosis. Polysomnography (PSG) remains the gold standard for diagnosis and assessment of severity of OSA in children. However, PSG is expensive, labor intensive, and not universally available. Instead, clinicians rely on clinical features to screen children in need of a PSG.

Adenotonsillectomy (AT) is the first line of treatment of pediatric OSA, leading to improvement or resolution of upper airway obstruction in most children. Children with severe upper airway obstruction are at risk of (i) increased prevalence of perioperative respiratory adverse events, as well as (ii) persistent OSA after surgery [3]. As the majority of ATs in the USA are performed without quantifying the severity of OSA by PSG [4], preoperative risk stratification may aid in the reduction of unexpected adverse events and improve outcomes. In addition, prediction of OSA severity may reduce the number of screening PSGs that are currently performed, resulting in more efficient use of healthcare resources.

Standardized screening questionnaires provide a semiquantitative estimate of symptom burden in children with high

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11325-019-01964-7) contains supplementary material, which is available to authorized users.

✉ Amal Isaiah
aisaiah@som.umaryland.edu

[1] Department of Otorhinolaryngology—Head and Neck Surgery, University of Maryland School of Medicine, Baltimore, MD, USA

[2] Department of Computer Science and Engineering, University of Texas at Arlington, Arlington, TX, USA

pretest probability for OSA. However, their ability to distinguish children with and without OSA has been relatively modest [5]. Furthermore, these questionnaires fare poorly in their ability to predict the severity of OSA [6].

Questionnaires described for screening children with suspected OSA aim to identify all possible symptoms. For example, in a review of 54 questionnaires described in the literature, the mean number of test items was 29 (95% confidence interval [22, 35]) [5]. The principal effect of the length of these questionnaires is respondent fatigue. In addition, responses to questions near the end of surveys are processed differently, and answers to these questions are faster, shorter, and more uniform than for those placed at the beginning [7]. On the other hand, brief questionnaires used to screen adults with OSA contain no more than eight items and are associated with much higher accuracy [8].

Mitchell et al. [9] assessed the correlation of clinical parameters with PSG indices in children with OSA using the obstructive sleep apnea syndrome-18 disease-specific quality of life instrument (OSA-18), the pediatric sleep questionnaire sleep-related breathing disorder scale (PSQ-SRBD, abbreviated as *PSQ*), and modified Epworth sleepiness scale (ESS) score. While race, obesity, and the PSQ score were associated with higher apnea-hypopnea index (AHI) and oxygen desaturation index (ODI), a multivariable model comprising all significant variables did not show superiority compared to a null model predicted by chance alone.

Screening questionnaires for pediatric OSA are associated with significant length and complexity resulting from the inclusion of potentially redundant items. Variable selection refers to identification of a subset of optimal predictors in correlated datasets [10]. This could identify a parsimonious prediction model, reduce statistical noise, and improve time taken to screen children in primary care practices. Furthermore, this process could result in the formulation of a brief questionnaire, similar in length to popular screening instruments for adults with OSA, which could be used as an epidemiologic tool.

We aimed to analyze the responses to questionnaires administered as part of the Childhood Adenotonsillectomy Trial (CHAT) [11]. We included three questionnaires in our analysis—the PSQ, modified ESS, and OSA-18. We hypothesized that a small subset of features, called *selected features* (SF), identified from the questionnaires predicts OSA severity better than the unabridged original questionnaires.

Methods

Study design and data collection

The study design as well as the outcomes related to the CHAT study, a randomized, controlled, multicenter study comparing early AT to watchful waiting with supportive care in children with OSA, are well described [11, 12]. Details of study procedures are provided in previous publications [6, 11, 12]. Archived data from the study was obtained through a data use agreement from the national sleep research resource (https://sleepdata.org).

Questionnaires

Three screening questionnaires were administered as part of the CHAT study and baseline responses recorded. These included:

(i) Epworth sleepiness scale (ESS) modified for children, a 10-item questionnaire (score from 0 to 24) representing increasing daytime sleepiness (Online Resource Table 1) [13].
(ii) The sleep-related breathing disorders (SRBD) subscale of the pediatric sleep questionnaire (PSQ), a 22-item questionnaire (range, 0–1), with higher scores indicating greater severity (Online Resource Table 2) [14].
(iii) The OSA-18, an 18-item disease-specific quality of life (QOL) questionnaire reflecting symptoms of sleep disturbance (score range, 4–28), physical suffering (range, 4–28), emotional distress (range, 3–21), daytime problems (range, 3–21), and caregiver concerns (range, 4–28). Total scores ranged from 18 to 126, on a scale of worsening QOL (Online Resource Table 3) [15].

Polysomnography

All children underwent full-night PSG according to the American Academy of Sleep Medicine (AASM) guidelines [16]. Scoring was performed according to the AASM pediatric criteria at a central PSG reading center. Two primary response variables—AHI and ODI—were obtained for the current study from the PSG data obtained at baseline. The AHI was defined as the sum of all obstructive and mixed apneas, plus hypopneas associated with a 50% reduction in airflow and either a greater than 3% desaturation or electroencephalographic arousal, divided by the total sleep time in hours. The ODI was defined as the number of times per hour of sleep that oxygen saturation dropped by 3% or more.

Statistical analysis

The intrinsic relationships between individual items within each questionnaire were assessed first by Spearman correlation coefficients and associated *P* values. A correlation matrix was constructed for visualization of the relationships. Statistically significant pairwise correlations were identified (Fig. 1a).

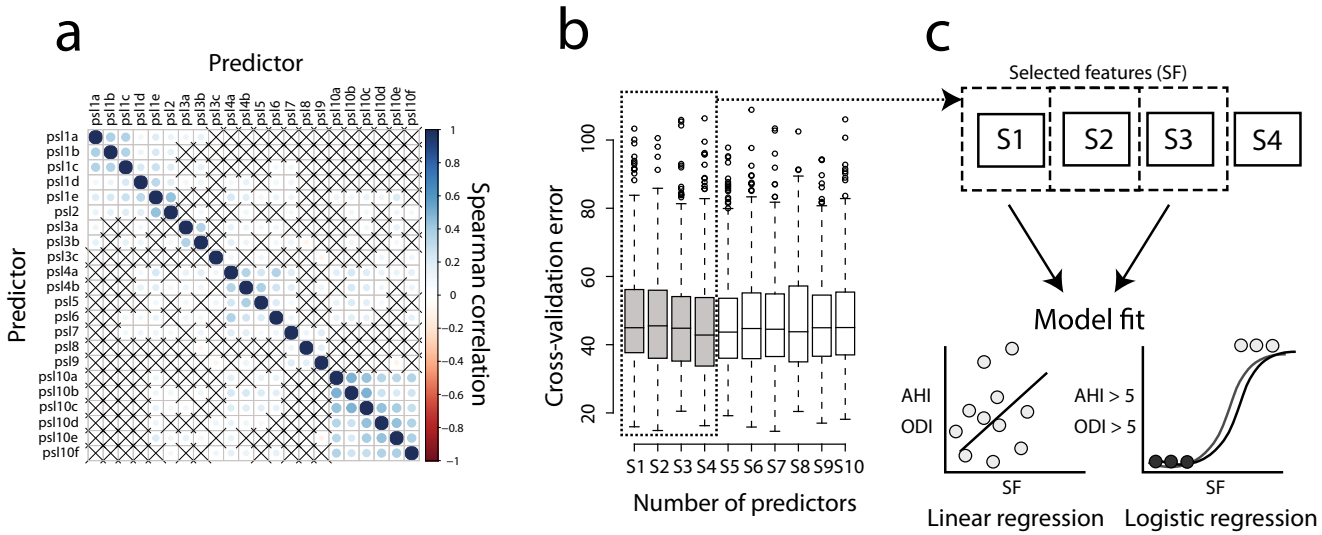

Fig. 1 Workflow for identification of key predictors from screening questionnaires. a identification of redundancy in screening questionnaires. Here, pairwise correlations between items within the responses to the pediatric sleep questionnaire from the Childhood Adenotonsillectomy Trial (CHAT) were assessed using Spearman correlation coefficients. The statistically significant correlations are marked by *X*. **b** shows an example of the cross-validation process within the correlation-adjusted regression analysis, with the *x*-axis demonstrating the number of serially added predictors and *y*-axis showing the average cross-validation error for the outcome apnea-hypopnea index (AHI) or oxygen desaturation index (ODI). **c** S1–S4 are selected features, whose predictive performance is assessed for both AHI and ODI using linear regression, as well as thresholds of 5 and 10 for each of these outcomes using logistic regression

Regression analysis was next performed to quantify the relationship between (i) three sets of predictors—ESS, OSA-18, and PSQ and (ii) two response variables—AHI and ODI. The choice of these variables replicated the key outcomes of polysomnographic testing, including obstruction and hypoxemia, as previously described [9]. Questionnaires that showed a significant relationship with the response variables were explored for ordered importance of the individual predictors within each scale. Correlation-adjusted regression (CAR) analysis, an approach proposed by Zuber and Strimmer [17], was used for canonical ordering of predictors by mitigating the associations between individual predictors. Statistical models with combinations of fewest variables and the least cross-validation error (Fig. 2b) were then selected for stepwise regression.

To relate individual questionnaire items to the severity of OSA, we used the stepwise Akaike Information Criterion (AIC) variable selection [18]. AIC is a well-described metric for relative quality of statistical models. Stepwise AIC variable selection was performed and proceeded as follows. A list of potential effects, including interactions, were added one at a time to an initial model, and AIC statistics were recorded. If AIC was reduced when at least one effect was added, the effect that reduced AIC the most was added to the initial model that, in turn, became the initial model for the next step. If AIC was not reduced by the addition of any effects, the stepwise process stopped. If removal of any effect already in the initial model resulted in a lower AIC, the effect that reduced AIC the most was removed. The selected features (SF) thus identified were combined and then assessed for their predictive performance (Fig. 1c). The outcomes thus tested included (i) AHI and ODI utilizing linear regression and (ii) AHI and ODI values of 5 and 10, each, utilizing logistic regression. Fit characteristics were compared between four models—null, total PSQ, total OSA-18, and SF utilizing a combination of the regression coefficients, proportion of variance explained (R^2) and *P* values. In (ii), receiver operating characteristic (ROC) curves were also plotted, and the classifier performance was determined by area under the curve (AUC). All analysis was carried out using R (version 3.5, The R Project for Statistical Computing, www.r-project.org). All statistical significance testing was two-tailed, and $P < 0.05$ was considered significant.

Results

A total of 464 children, aged 5 to 10 years, were recruited and randomized to AT or watchful waiting with supportive care with follow-up at 7 months. Based on completeness of data, 453 children were included in the current study. The mean [95% confidence interval] ESS, OSA-18, and PSQ scores were 7.7 [7.2, 8.2], 53.6 [51.9, 55.3], and 0.50 [0.48, 0.52], respectively. The total OSA-18 score showed statistically significant positive association with AHI (Spearman $\rho = 0.10$, $P = 0.04$) and ODI ($\rho = 0.11$, $P = 0.03$). Similarly, the PSQ score showed significant correlation with AHI ($\rho = 0.11$, $P = 0.02$) and ODI ($\rho = 0.12$, $P = 0.01$). The ESS score, however, did not show a significant association ($\rho = 0.00$, $P = 0.61$). The relationship between individual items within each questionnaire

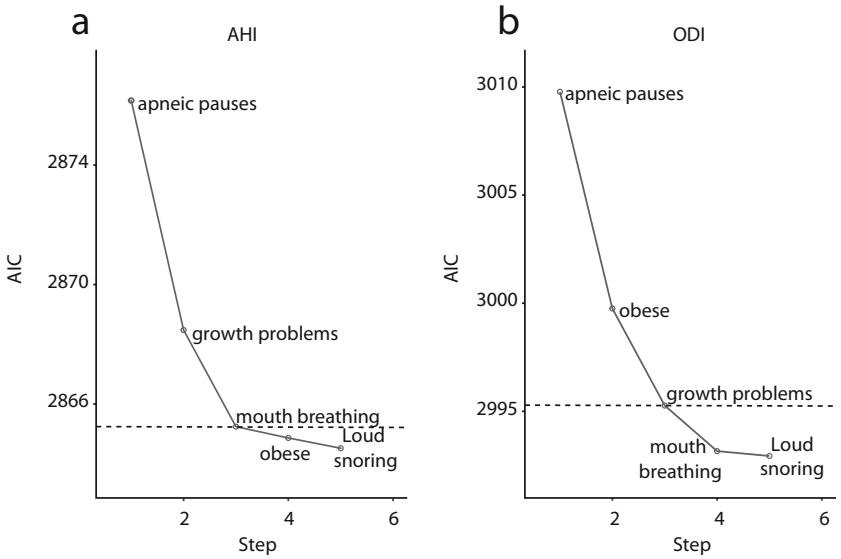

Fig. 2 Identification of best model for prediction of OSA based on feature selection using Akaike information criteria (AIC). A regression model was first created from a set of potentially independent predictor variables by entering and removing predictors identified from the correlation-adjusted regression models, based on AIC, in a stepwise manner until there was no variable left to enter or remove. AIC scores are modeled for apnea-hypopnea index (AHI, a) and oxygen desaturation index (ODI, b) and represented on the *y*-axis as a function of the number of predictors in each step (*x*-axis). An optimal model is characterized by a combination of smallest number of predictors and the lowest AIC scores. In (**a**), *apneic pauses*, *growth problems*, and *mouth breathing*, and in (**b**) *apneic pauses*, *obese*, and *growth problems* similarly provided the best model from a range of available models

was assessed using the pairwise correlation coefficients and their associated *P* values. The correlation matrix demonstrated that more than 50% of the pairwise relationships were statistically significant (Fig. 1a).

Next, CAR analysis was used for ordering of predictors by adjusting for the correlation between individual predictors within the OSA-18 and PSQ, separately (Fig. 1b). For prediction of AHI, comparison of cross-validation prediction errors associated with the CAR regression models showed no further reduction in prediction error beyond the first predictor in OSA-18, OSA1b (witnessed pauses in breathing). Similarly, the magnitude of prediction error did not decrease beyond the serial addition of the first three predictors within PSQ—PSL8 (growth problems), PSL3a (mouth breathing), and PSL9 (overweight). Similarly, for prediction of ODI, CAR analysis showed no further reduction in prediction error after the inclusion of OSA1b. For PSQ, the cross-validation analysis showed that the addition of items beyond PSL8, PSL3a, and PSL9 also did not improve prediction error.

To identify the best model for the prediction of OSA severity, a statistically parsimonious model was created from the variables identified in the CAR analysis using AIC (Fig. 1). Stepwise bidirectional AIC variable selection was used to model the prediction of AHI and ODI, as shown in Fig. 2. An ideal model comprised the smallest number of predictors resulting in the lowest AIC scores possible with progressive addition of each variable. For AHI, combining OSA1b, PSL8, and PSL3a produced the best predictive model, with no further reduction in AIC values with addition of subsequent variables (Fig. 2a). Similarly, for ODI, combining OSA1b, PSL9, and PSL8 provided the best model with no further reduction of AIC score (Fig. 2b).

Next, the predictor subset of four items identified from the two prediction models for AHI and ODI were combined (OSA1b, PSL3a, PSL8, and PSL9), called "SF" (selected features). Subsequently, SF was assessed for its ability to both predict (i) AHI and ODI in a linear regression model, (ii) AHI > 5 and ODI > 5 in logistic regression models, and (iii) AHI > 10 and ODI > 10 in logistic regression models. For the prediction of AHI and ODI, the total scores from PSQ and OSA-18 were not better than the null models (Table 1). However, the model fit for SF was superior compared to the null model, as well as PSQ and OSA-18. Equally, SF was a more robust predictor of both (i) AHI > 5 and ODI > 5 (Table 2) and (ii) AHI > 10 and ODI > 10 (Table 3), as demonstrated by AIC and ROC (Fig. 3) analysis.

Discussion

We examined the performance of commonly used questionnaires for predicting the severity of OSA in children based on clinical findings. The analysis presented in the current study identified the most informative items in three popular questionnaires—PSQ, ESS, and OSA-18. A total of four questions, comprising less than a tenth of the total number of items in three questionnaires provided the best model for prediction of AHI and ODI respectively, based on variable selection.

Table 1 Predicting obstructive sleep apnea (OSAS) severity using total questionnaire scores—results from linear regression. Comparing prediction of apnea-hypopnea index (AHI) and oxygen desaturation (ODI) using total scores obtained from pediatric sleep questionnaire (PSQ), obstructive sleep apnea-18 survey (OSA-18), and selected features identified using feature selection (SF). Columns indicate metrics of regression performance, including the coefficient, R^2 and the P values. Each questionnaire model was compared with the null model using the Akaike information criterion (AIC); lower values indicated better model fit

	AHI				ODI			
Scale	β [95% CI]	P	R^2	AIC	β [95% CI]	P	R^2	AIC
NULL	Not applicable			915	Not applicable			1408
PSQ	0.43 [0.08 to 0.78]	0.02	0.02	911	0.72 [0.10 to 1.36]	0.02	0.01	1405
OSA-18	0.00 [0.00 to 0.00]	0.03	0.01	912	0.01 [0.00 to 0.00]	0.03	0.01	1406
SF	0.09 [0.05 to 0.13]	< 0.001	0.08	880	0.16 [0.10 to 0.22]	< 0.001	0.07	1378

These results demonstrate that the efforts to improve screening performance of questionnaires for pediatric OSA should focus on identification of key features rather than increasing the number of test items to identify symptom burden.

Questionnaires are powerful, practical, and inexpensive tools for screening the general population for conditions with relatively high prevalence. For example, the STOP-Bang questionnaire [19] was created as a reliable and user-friendly screening tool for adult OSA. This questionnaire consists of eight dichotomous (yes/no) questions related to both OSA-related symptoms and patient demographics. The sensitivity of the score with a cutoff of 3 or greater for patients with AHI > 15 and AHI > 30 is 93% and 100%, respectively [8]. In the same study, the respective specificities were 30% and 29%. Since introduction, the questionnaire has been repeatedly validated [8].

In comparison, most pediatric OSA screening questionnaires have modest to low sensitivity and specificity. While PSQ is the most widely used tool in research settings, a sensitivity of 81–85% for AHI > 1 and a specificity of 87% do not provide a screening advantage for the general population [14]. In addition, the PSQ is comprised of more than 20 questions, making it susceptible to respondent fatigue. Similarly, the ESS was shown to have modest sensitivity at 66–76% and a specificity of 41–49% depending on the cutoff chosen [20]. The OSA-18 has the lowest sensitivity (40%)[21] given the main purpose of this tool is to assess the impact of OSA on the quality of life of children before and after AT. Although there are several variables that are informative within these questionnaires, it is possible that the additional variables are correlated with the existing variables and increasing the number of items that does not improve the accuracy.

Our results show for the first time that the questionnaires used for screening for OSA have substantial redundancy. The presence of a large number (> 50%) of statistically significant pairwise relationships between items supports the use of variable selection to identify the key predictors encapsulated within them. In the OSA-18 questionnaire, the only question that was included in the predictive model refers to witnessed apneic events. A study by Wang et al. [22] showed that the prevalence of witnessed apneic events among children with a diagnosis of OSA was 88%. Carroll et al. [23] compared children with primary snoring and patients with confirmed OSA and found significant difference in the prevalence of witnessed apneic events (46% vs. 74%). Among the items in the PSQ, daytime mouth breathing, perceived growth delay, and obesity were identified as significant for predicting AHI and ODI. Brouilette et al. [24] found significant difference in

Table 2 Predicting apnea hypopnea index (AHI) > 5 and oxygen desaturation index (ODI) > 5 using total questionnaire scores—results from logistic regression. Performance of four sets of predictors—(i) null model, (ii) pediatric sleep questionnaire (PSQ), (iii) obstructive sleep apnea-18 survey (OSA-18), and (iv) selected features identified using feature selection (SF). Columns indicate metrics of regression performance, including the coefficient (β with associated 95% confidence interval [95% CI], P value, and the associated odds ratio [95% CI]). In addition, the area under the curve (AUC) for the receiver operating characteristic (ROC) analysis as well as the Akaike information criteria (AIC) is also provided. Plots from the ROC analyses are shown in Fig. 2

	AHI > 5					ODI > 5				
Scale	β [95% CI]	P	OR [95% CI]	AUC	AIC	β [95% CI]	P	OR [95% CI]	AUC	AIC
NULL	Not applicable				607.1	Not applicable				613.7
PSQ	0.68 [0.15, 1.23]	0.20	1.99 [0.70, 5.76]	0.53	607.6	1.41 [0.87, 1.95]	0.008	4.11 [1.44, 12.02]	0.57	609.1
OSA-18	0.00 [0.00, 0.01]	0.83	1.00 [0.99, 1.01]	0.50	609.1	0.01 [0.00, 0.01]	0.04	1.01 [1.00, 1.02]	0.55	612.1
SF	0.21 [0.16, 0.25]	< 0.001	1.21 [1.11, 1.33]	0.62	591.1	0.22 [0.18, 0.26]	< 0.001	1.25 [1.14, 1.37]	0.62	593.4

Table 3 Predicting apnea hypopnea index (AHI) > 10 and oxygen desaturation index (ODI) > 10 using total questionnaire scores—results from logistic regression. Performance of four sets of predictors—(i) null model, (ii) pediatric sleep questionnaire (PSQ), (iii) obstructive sleep apnea-18 survey (OSA-18), and (iv) selected features (SF). Columns indicate metrics of regression performance, including the coefficient (β with associated 95% confidence interval [95% CI], *P* value, and the associated odds ratio [95% CI]). In addition, the areas under the curve (AUC) for the receiver operating characteristic (ROC) analysis as well as the Akaike information criteria (AIC) are also provided. Plots from the ROC analyses are shown in Fig. 2

	AHI > 10					ODI > 10				
Scale	β [95% CI]	*P*	OR [95% CI]	AUC	AIC	β [95% CI]	*P*	OR [95% CI]	AUC	AIC
NULL	Not applicable				484	Not applicable				491
PSQ	1.00 [− 0.25 to 2.25]	0.12	2.71 [0.79 to 5.95]	0.54	484	1.52 [0.27 to 1.64]	0.02	4.57 [1.33 to 16.15]	0.57	487
OSA-18	0.01 [− 0.01 to 0.03]	0.09	1.01 [1.00 to 1.02]	0.56	484	0.01 [− 0.01 to 0.03]	0.03	1.01 [1.00 to 1.03]	0.56	489
SF	0.29 [0.17 to 0.41]	< 0.001	1.33 [1.19 to 1.50]	0.66	459	0.27 [0.15 to 0.39]	< 0.001	1.31 [1.17 to 1.47]	0.65	469

the frequency of the following symptoms between children with OSA vs. controls—96% vs. 2% for observed apneic events and 87% vs. 18% for mouth breathing while awake. Similarly, Carroll et al. [23] found a significant difference in reported daytime mouth breathing between children with primary snoring vs. OSA (61 vs. 85%). Obesity has also been found to be associated with OSA in children. The prevalence of OSA among obese adolescents undergoing bariatric surgery was 55% [25]. Together, these results show that the variables selected for the creation of the best model for the prediction of AHI and ODI are consistent with previous studies demonstrating the relationship between clinical parameters and the severity of pediatric OSA.

Variable selection facilitates ordering of predictors by their relative importance. Lu and Petkova [26] studied factors affecting the performance of variable selection methods in the context of developing screening questionnaires for psychiatric disorders. The use of decorrelation to offset interactions between correlated items in a questionnaire is novel to pediatric OSA. Through a combination of variable and subsequent

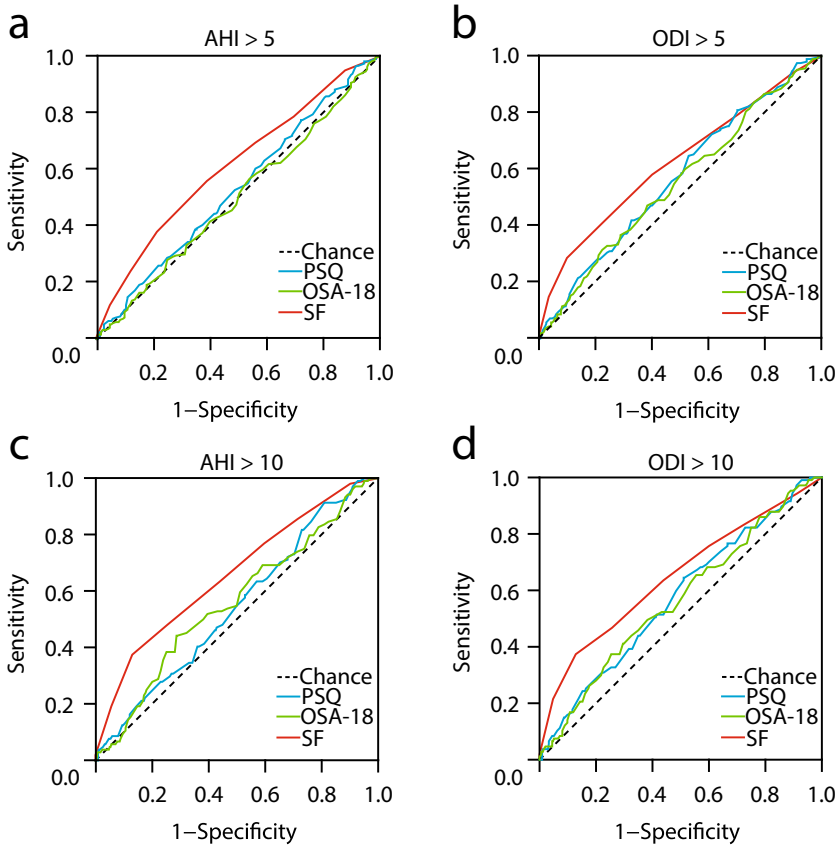

Fig. 3 Prediction of OSA severity compared by receiver operating characteristic curve (ROC) analysis for various predictive models. The prediction performances of pediatric sleep questionnaire-sleep-related breathing disorder scale (PSQ), obstructive sleep apnea-18 (OSA-18) survey, and selected features (SF) identified in the current study (witnessed apneic events, mouth breathing, perceived growth delay, and obesity) were assessed for clinically significant OSA, defined by (**a**) apnea hypopnea index (AHI) > 5, (**b**) oxygen desaturation index (ODI) > 5, as well as (**c**) AHI > 10 and (**d**) ODI > 10. Chance distribution is shown by the diagonal, signifying an area under the curve of 0.50. SF demonstrated the highest predictive accuracy for each threshold of AHI and ODI. The ROC analysis results are shown in Table 3

model selection, only four out of a total of 50 items presented in three questionnaires were identified to be of value in clinical practice to screen for the severity of AHI and ODI.

Together, our results support the use of brief versions of widely used screening instruments, but their universal use needs further validation. Furthermore, the symptoms identified remain clinically relevant and have been previously reported in their ability to distinguish OSA from primary snoring. The rationale for choosing a threshold of 5 for AHI and ODI is justified by the recommended initiation of treatment of OSA at those levels of severity [27]. The use of AHI and ODI threshold of 10 is aligned with the clinical practice guidelines from the American Academy of Otolaryngology—Head and Neck Surgery that recommend overnight observation for children with AHI > 10 due to the increased risk of perioperative respiratory adverse events [28].

The strengths of the study are related to the dataset from the CHAT study. This was a large multicenter, randomized controlled trial at six academic centers [11]. Participants were recruited from pediatric sleep centers, general pediatric clinics, pediatric otolaryngology offices, and the general community. Therefore, the current study maintains the same level of external validity and generalizability as the CHAT study [9].

The principal weaknesses related to this study were the formulation of predictors from children enrolled with a limited range of age and OSA severity. The utility of a questionnaire comprising the features identified in the current study will need prospective validation. As the results presented here are from a secondary analysis of an existing dataset, the logical next step is a population-based study both in the clinic setting, which would entail polysomnographic validation of the score obtained. Although the CAR analysis is one form of variable ordering, there are other algorithms that may provide different results when tested. The use of questionnaires even after feature selection cannot replace clinical judgement in combination with objective assessments in the evaluation and management of pediatric OSA. Hence, input from experts to evaluate the appropriateness of questions chosen for these instruments cannot be overemphasized.

Conclusions

Screening questionnaires for sleep-related breathing disorders in children contain only a small number of variables predicting OSA severity in children. Rational ordering of survey questions and removal of unnecessary items may avoid respondent fatigue. Our results utilizing variable selection could identify only four out of 50 items from the three questionnaires predictive of the severity of OSA in children. These included witnessed apneic events, mouth breathing, perceived growth delay, and obesity. Future studies are required to validate the resultant abbreviated scale for the prediction of OSA severity.

Author contribution Substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data: Isaiah, Pereira, and Das

Drafting the article or revising it critically for important intellectual content: Shikara, Isaiah, Pereira, and Das

Final approval of the version to be published: Isaiah, Shikara, Pereira, and Das

Compliance with ethical standards

Conflict of interest Amal Isaiah has received patent-related royalties from the University of Maryland, Baltimore for inventions related to sleep apnea imaging in adults. Kevin D. Pereira and Gautam Das declare no conflicts of interest.

Ethical approval This article does not contain any studies with human participants performed by any of the authors. De-identified data from the CHAT study was obtained from the National Sleep Research Resource (www.sleepdata.org) following a data use agreement.

References

1. (1996) Standards and indications for cardiopulmonary sleep studies in children. American Thoracic Society. Am J Respir Crit Care Med 153:866–878. https://doi.org/10.1164/ajrccm.153.2.8564147
2. Lumeng JC, Chervin RD (2008) Epidemiology of pediatric obstructive sleep apnea. Proc Am Thorac Soc 5:242–252. https://doi.org/10.1513/pats.200708-135MG
3. Isaiah A, Hamdan H, Johnson RF et al (2017) Very severe obstructive sleep apnea in children: outcomes of adenotonsillectomy and risk factors for persistence. Otolaryngol–Head Neck Surg Off J Am Acad Otolaryngol-Head Neck Surg 157:128–134. https://doi.org/10.1177/0194599817700370
4. Mitchell RB, Pereira KD, Friedman NR (2006) Sleep-disordered breathing in children: survey of current practice. Laryngoscope 116:956–958. https://doi.org/10.1097/01.MLG.0000216413.22408.FD
5. 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. https://doi.org/10.1016/j.smrv.2010.07.005
6. Mitchell RB, Garetz S, Moore RH, Rosen CL, Marcus CL, Katz ES, Arens R, Chervin RD, Paruthi S, Amin R, Elden L, Ellenberg SS, Redline S (2015) The use of clinical parameters to predict obstructive sleep apnea syndrome severity in children: the Childhood Adenotonsillectomy (CHAT) study randomized clinical trial. JAMA Otolaryngol Neck Surg 141:130–136. https://doi.org/10.1001/jamaoto.2014.3049
7. Galesic M, Bosnjak M (2009) Effects of questionnaire length on participation and indicators of response quality in a web survey. Public Opin Q 73:349–360. https://doi.org/10.1093/poq/nfp031
8. Silva GE, Vana KD, Goodwin JL, Sherrill DL, Quan SF (2011) Identification of patients with sleep disordered breathing: comparing the four-variable screening tool, STOP, STOP-Bang, and Epworth Sleepiness Scales. J Clin Sleep Med JCSM Off Publ Am Acad Sleep Med 7:467–472. https://doi.org/10.5664/JCSM.1308
9. Mitchell RB, Garetz S, Moore RH, Rosen CL, Marcus CL, Katz ES, Arens R, Chervin RD, Paruthi S, Amin R, Elden L, Ellenberg SS, Redline S (2015) The use of clinical parameters to predict obstructive sleep apnea syndrome severity in children: the

Childhood Adenotonsillectomy (CHAT) study randomized clinical trial. JAMA Otolaryngol– Head Neck Surg 141:130–136. https://doi.org/10.1001/jamaoto.2014.3049
10. Zou H, Hastie T Regularization and variable selection via the elastic net. J R Stat Soc Ser B Stat Methodol 67:301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x
11. Marcus CL, Moore RH, Rosen CL, Giordani B, Garetz SL, Taylor HG, Mitchell RB, Amin R, Katz ES, Arens R, Paruthi S, Muzumdar H, Gozal D, Thomas NH, Ware J, Beebe D, Snyder K, Elden L, Sprecher RC, Willging P, Jones D, Bent JP, Hoban T, Chervin RD, Ellenberg SS, Redline S, Childhood Adenotonsillectomy Trial (CHAT) (2013) A randomized trial of adenotonsillectomy for childhood sleep apnea. N Engl J Med 368:2366–2376. https://doi.org/10.1056/NEJMoa1215881
12. Redline S, Amin R, Beebe D, et al (2011) The Childhood Adenotonsillectomy Trial (CHAT): rationale, design, and challenges of a randomized controlled trial evaluating a standard surgical procedure in a pediatric population., The Childhood Adenotonsillectomy Trial (CHAT): rationale, design, and challenges of a randomized controlled trial evaluating a standard surgical procedure in a pediatric population. Sleep Sleep 34, 34:1509, 1509–1517. https://doi.org/10.5665/sleep.1388
13. Melendres MCS, Lutz JM, Rubin ED, Marcus CL (2004) Daytime sleepiness and hyperactivity in children with suspected sleep-disordered breathing. Pediatrics 114:768–775. https://doi.org/10.1542/peds.2004-0730
14. Chervin RD, Hedger K, Dillon JE, Pituch KJ (2000) Pediatric sleep questionnaire (PSQ): validity and reliability of scales for sleep-disordered breathing, snoring, sleepiness, and behavioral problems. Sleep Med 1:21–32. https://doi.org/10.1016/S1389-9457(99)00009-X
15. Franco RA, 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 Off J Am Acad Otolaryngol-Head Neck Surg 123:9–16. https://doi.org/10.1067/mhn.2000.105254
16. Berry RB, Budhiraja R, Gottlieb DJ, Gozal D, Iber C, Kapur VK, Marcus CL, Mehra R, Parthasarathy S, Quan SF, Redline S, Strohl KP, Davidson Ward SL, Tangredi MM, American Academy of Sleep Medicine (2012) Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. J Clin Sleep Med JCSM Off Publ Am Acad Sleep Med 8:597–619. https://doi.org/10.5664/jcsm.2172
17. Zuber V, Strimmer K (2011) High-dimensional regression and variable selection using CAR scores. Stat Appl Genet Mol Biol 10. https://doi.org/10.2202/1544-6115.1730
18. Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19:716–723. https://doi.org/10.1109/TAC.1974.1100705
19. Chung F, Yegneswaran B, Liao P, Chung SA, Vairavanathan S, Islam S, Khajehdehi A, Shapiro CM (2008) STOP questionnaire: a tool to screen patients for obstructive sleep apnea. Anesthesiology 108:812–821. https://doi.org/10.1097/ALN.0b013e31816d83e4
20. Borgström 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. https://doi.org/10.1016/j.ijporl.2013.08.030
21. 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. https://doi.org/10.1542/peds.2009-0731
22. Wang DRC, Elkins DTP, Keech DD et al (1998) Accuracy of clinical evaluation in pediatric obstructive sleep apnea. Otolaryngol Neck Surg 118:69–73. https://doi.org/10.1016/S0194-5998(98)70377-8
23. Carroll JL, McColley SA, Marcus CL, Curtis S, Loughlin GM (1995) Inability of clinical history to distinguish primary snoring from obstructive sleep apnea syndrome in children. CHEST 108:610–618. https://doi.org/10.1378/chest.108.3.610
24. Brouilette R, Hanson D, David R, Klemka L, Szatkowski A, Fernbach S, Hunt C (1984) A diagnostic approach to suspected obstructive sleep apnea in children. J Pediatr 105:10–14. https://doi.org/10.1016/S0022-3476(84)80348-0
25. Maninder K, Thomas I, Victor G et al (2012) Obstructive sleep apnea in extremely overweight adolescents undergoing bariatric surgery. Obes Res 13:1175–1179. https://doi.org/10.1038/oby.2005.139
26. Lu F, Petkova E (2014) A comparative study of variable selection methods in the context of developing psychiatric screening instruments. Stat Med 33:401–421. https://doi.org/10.1002/sim.5937
27. Redline S, Tishler PV, Schluchter M, Aylor J, Clark K, Graham G (1999) Risk factors for sleep-disordered breathing in children. Associations with obesity, race, and respiratory problems. Am J Respir Crit Care Med 159:1527–1532. https://doi.org/10.1164/ajrccm.159.5.9809079
28. Mitchell RB, Archer SM, Ishman SL et al (2019) Clinical practice guideline: tonsillectomy in children (Update). Otolaryngol–Head Neck Surg Off J Am Acad Otolaryngol-Head Neck Surg 160:S1–S42. https://doi.org/10.1177/0194599818801757