JAMA Otolaryngology–Head & Neck Surgery | **Original Investigation**

Predictors of Behavioral Changes After Adenotonsillectomy in Pediatric Obstructive Sleep Apnea
A Secondary Analysis of a Randomized Clinical Trial

Amal Isaiah, MD, PhD; Adam J. Spanier, MD, PhD; Lynn M. Grattan, PhD; Yan Wang, MD, DrPH; Kevin D. Pereira, MD, MS

IMPORTANCE Adenotonsillectomy (AT) is associated with improved behavior in children with obstructive sleep apnea (OSA). However, it is unknown whether polysomnographic parameters are superior to the parent-reported severity of sleep-disordered breathing (SDB) in predicting behavioral changes after AT.

OBJECTIVE To ascertain whether polysomnographic parameters vs parent-reported severity of SDB are better predictors of treatment-related behavioral changes in children with OSA.

DESIGN, SETTING, AND PARTICIPANTS This ad hoc secondary analysis of the Childhood Adenotonsillectomy Trial (CHAT) downloaded and analyzed data from January 1 to January 31, 2020. Children aged 5 to 9 years with a polysomnographic diagnosis of OSA were enrolled in the CHAT and subsequently randomized to undergo either early AT or watchful waiting with supportive care. All outcome measures were obtained at baseline and at follow-up (7 months after randomization).

INTERVENTIONS Early AT vs watchful waiting with supportive care.

MAIN OUTCOMES AND MEASURES Postrandomization changes between the baseline and follow-up periods were derived from (1) T scores in 4 validated behavioral assessments (Conners Global Index parent and teacher versions, Behavior Rating Inventory of Executive Function metacognition index, and Child Behavior Checklist of total, internalizing, and externalizing behavior subscales); (2) 8 aggregated polysomnographic parameters representing the severity of obstruction, hypoxemia, sleep quality, and structure; and (3) the parent-reported severity of SDB measured by the Pediatric Sleep Questionnaire–Sleep-Related Breathing Disorder (PSQ-SRBD) scale. The treatment-related changes in each of the behavioral outcomes attributable to changes in SDB severity (represented by the subjective PSQ-SRBD score and objective polysomnographic parameters) were measured and compared using mediation analysis.

RESULTS A total of 453 children were assessed at baseline, of whom 234 were girls (52%) and the mean (SD) age was 6.6 (1.4) years. The postrandomization changes in 7 of 8 behavioral outcome measures between the baseline and follow-up periods were partially mediated by the changes in PSQ-SRBD scores (range of nonzero causally mediated effects, 2.4-3.5), without contribution from any of the polysomnographic parameters.

CONCLUSIONS AND RELEVANCE This secondary analysis of a national randomized clinical trial found that most treatment-related behavioral changes in children with OSA were mediated by the changes in parent-reported SDB severity alone. These findings suggest that polysomnographic parameters provide clinicians with limited means to predict the improvement in neurobehavioral morbidity in OSA.

TRIAL REGISTRATION ClinicalTrials.gov Identifier: NCT00560859

Author Affiliations: Department of Otorhinolaryngology—Head and Neck Surgery, University of Maryland School of Medicine, Baltimore (Isaiah, Pereira); Department of Pediatrics, University of Maryland School of Medicine, Baltimore (Isaiah, Spanier, Wang, Pereira); Department of Neurology, University of Maryland School of Medicine, Baltimore (Grattan); Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore (Grattan, Wang).

Corresponding Author: Amal Isaiah, MD, PhD, Department of Otorhinolaryngology—Head and Neck Surgery, University of Maryland School of Medicine, 16 S Eutaw St, Ste 500, Baltimore, MD 21201 (aisaiah@som.umaryland.edu).

Sleep-disordered breathing (SDB) comprises a spectrum of respiratory problems during sleep, the central feature of which is nocturnal upper-airway obstruction that results in snoring and respiratory pauses.[1] Although SDB continues to be the principal indication for adenotonsillectomy (AT) performed in the United States,[2] universal objective characterization of SDB using polysomnography before the surgical procedure has garnered some support.[3,4] The motivation for objective testing arises, in part, from the perception that subjective assessments of SDB may be inferior to polysomnography in their discriminatory ability to assess the associated morbidity. In addition, the potential for bias in subjective assessments may result in substantial variation in the preoperative selection criteria for children undergoing AT.

Despite being the criterion standard for SDB stratification, use of preoperative polysomnography remains below 10% principally because of its cost and the resources needed.[5] Nonetheless, a small number of studies, which demonstrated the discernible positive association between the polysomnographic severity of SDB measured by the apnea hypopnea index (AHI) and standardized outcomes such as neurobehavioral tests, provide some rationale for using polysomnography to guide treatment.[6-8] Furthermore, some guidelines recommend treatment of SDB in children that is solely based on the AHI,[9] raising concerns regarding the overuse of polysomnography and potential inappropriate selection of surgical candidates.

The Childhood Adenotonsillectomy Trial (CHAT) was a randomized clinical trial that assessed early AT compared with watchful waiting with supportive care (WWSC) for the treatment of obstructive sleep apnea (OSA), the most severe form of SDB.[10,11] In this trial, early AT was associated with statistically significantly greater improvements in the domains of behavior and quality of life, thus supporting surgical procedure as the first line of treatment of pediatric OSA. However, few studies have assessed the underlying mechanism of the treatment outcome beyond the potential changes in the severity of upper-airway obstruction as measured by polysomnography.

Although previous studies showed that most OSA-related outcomes that spanned the domains of cognition, behavior, quality of life, blood pressure, metabolism, and autonomic function were not causally associated with the treatment-related changes in the AHI,[12,13] it remains unknown whether the polysomnographic parameters representing the severity of hypoxemia or sleep disruption could predict the treatment-related changes in outcomes. Therefore, given that childhood behavior may be affected by sleep disruption with or without intermittent hypoxemia,[14] we examined in the present study whether any of the polysomnographic parameters representing these physiological measures are superior to parent-reported symptoms of SDB in assessing the changes in behavioral outcomes after OSA treatment.

Methods

Study Design

The CHAT design has been reported elsewhere, and a primary analysis and a series of secondary analyses have been published.[10,11] The data use agreement for the current study was obtained by the first author (A.I.) from the National Sleep Research Resource.[15] Local institutional review board approval was not sought for this study given that the terms of use of the deidentified data were governed by the National Sleep Research Resource and the institutional review board of the Brigham and Women's Hospital. Patient consent was not obtained given that no identifiable information was used and that all fields were anonymized at source. The data set was downloaded and analyzed in January 2020.

Briefly, children aged 5 to 9 years were enrolled in CHAT after a polysomnographic diagnosis of OSA, defined as an AHI of 2 or more events per hour or an obstructive apnea index of 1 or more events event per hour with clinical symptoms of SDB. These children were subsequently randomized to undergo either early AT or WWSC. All outcome measures were obtained at baseline and at follow-up (7 months after randomization). Children with an AHI of 30[16] were excluded because they needed urgent treatment. The complete inclusion and exclusion criteria are described in Redline et al.[10]

Outcome Variables

The 4 behavioral measures used in the current ad hoc secondary analysis were the Conners Global Index parent and teacher versions[17-19] (parent version score range: 21-90; teacher version score range: 41-90); the parent- and teacher-reported Behavior Rating Inventory of Executive Function (BRIEF) metacognition index (MI)[20] (parent version score range: 32-91; teacher version score range: 39-104); and the parent-reported Child Behavior Checklist of total, internalizing, and externalizing behavior subscales (total score range: 24-81; internalizing subscale score range: 33-89; externalizing subscale score range: 33-79).[21] For each subscale, higher scores indicate worse behavior.

These validated, standardized, and widely used instruments provided estimates of children's psychological difficulty in the emotional, social, and behavioral domains. The raw scores were converted to T scores based on age- and sex-based normative data.

Key Points

Question Are polysomnographic parameters superior to parent-reported symptoms of upper-airway obstruction in predicting posttreatment behavioral outcomes in children with obstructive sleep apnea (OSA)?

Findings In this secondary analysis of the Childhood Adenotonsillectomy Trial involving 453 children, parent-reported symptoms of upper-airway obstruction were better indicators of most changes in children's behavior than were polysomnographic parameters.

Meaning Results of this secondary analysis suggest that most treatment-related behavioral changes in children with OSA were mediated by the changes in parent-reported sleep-disordered breathing severity alone.

Mediator Variables and Covariates

Standard overnight polysomnographic parameters were mediators collected according to the American Academy of Sleep Medicine guidelines.[15] A mediator is an intermediate variable between the dependent and independent variables to which the changes in the primary outcome variable can be attributed. These mediators included sleep stage, thoracoabdominal movement, oronasal airflow, oxygen saturation as measured by pulse oximetry (SpO_2), leg and chin movement, body position, and cardiac rhythm, and they were collected from each study site and scored at a centralized reading center. The key variables were minimum SpO_2 (lowest oximetry value during sleep), total sleep time below 92% (sleep time with SpO_2 below 92%), sleep duration (determined from sleep onset to awakening), sleep efficiency (percentage of total study time spent asleep), percentage of rapid-eye-movement (REM) sleep (calculated from REM and total sleep duration), and percentage of stage 3 or 4 sleep (calculated from stages 3 and 4 of sleep and total sleep duration). The AHI was the total number of obstructive apnea and hypopnea events per hour. The arousal index was the number of arousal events per hour. An obstructive apnea was at least a 90% reduction in oronasal thermal airflow signal lasting a minimum of 2 breaths during baseline breathing despite respiratory effort. A hypopnea was a decrease in airflow of at least 30% for a minimum of 2 breaths with either an arousal or a 3% decrease in SpO_2.

Another set of mediator variables was the parent-reported SDB severity. The severity of SDB was measured using the Pediatric Sleep Questionnaire–Sleep-Related Breathing Disorder (PSQ-SRBD) scale, a validated, popular, and widely available instrument, comprising 22 items (score range: 0.46-0.95, with higher scores indicating worse SDB).[22] This parent-completed tool assesses prominent symptom clusters, including snoring, daytime sleepiness, and attention or hyperactivity.[23]

Demographic covariates included the child's age (in years), biological sex at birth, self-selected race/ethnicity, and body mass index (BMI) percentile score; the caregiver educational status; and the total household income before taxes (US dollars). The BMI was calculated as weight in kilograms divided by height in meters squared, and the BMI percentile was determined by comparison with the age- and sex-based BMI charts from the Centers for Disease Control and Prevention.[24] A BMI in the 95th percentile or higher was considered a criterion for obesity. The maternal educational status was dichotomized into high school graduate or more or less than high school.

Adjustment for Missing Data

Children were excluded if most of their outcome data or covariates were missing. If less than 30% of data were missing, a multiple-imputation procedure was performed by probabilistically filling the missing fields, accounting for the associated uncertainty. Multiple imputation was performed using the expectation-maximization with bootstrapping approach implemented by Amelia II, an R-based package.[25]

Statistical Analysis

Baseline characteristics were described by proportions (No. [%]) for categorical variables and mean (SD) for continuous vari-

ables. Study covariates included age, sex, race/ethnicity, BMI percentile, household income, and caregiver educational status. The changes in behavioral measure outcomes between baseline and follow-up were compared between the study groups (those who received early AT vs those randomized to WWSC) using analysis of covariance after adjustment for covariates. The effect size associated with each comparison was subsequently calculated using Cohen d, providing an estimate of the mean standardized difference in the outcome measure between the trial groups. Effect size thresholds by established conventions were Cohen d = 0.2 (small), Cohen d = 0.5 (medium), and Cohen d = 0.8 (large).[26]

The association between each of the polysomnographic parameters and the behavioral measures was assessed using a generalized linear model because of the expected skew in OSA severity attributed to the inclusion criteria. This model takes into consideration the underlying asymmetry in the distribution of the AHI data owing to the substantially fewer children exceeding the median value of the AHI in the CHAT.[27] The change in the proportion of variance (adjusted R^2), representing the effect size associated with the predictor, was measured between the baseline covariates-only model and the baseline model with an added predictor variable.

We used causal mediation analysis to estimate an intermediate pathway, which provided a mechanistic explanation for the observed changes in the outcomes of interest. The general outline of mediation analysis is shown in **Figure 1**, including the potential pathways of the mediators (Figure 1A) and the estimated direct and causally mediated effects (Figure 1B). A detailed explanation of causal mediation analysis applied to the CHAT was provided in a previous publication.[12] Briefly, causal mediation[28] is a conceptual extension of mediation analysis described by Baron and Kenny.[29] A mediator is an intermediate variable between the dependent and independent variables to which the changes in the primary outcome variable can be attributed. We performed causal mediation analysis by jointly modeling 2 separate regressions. The first regression estimated the association between the change in the mediator and the covariates of interest after including the pretreatment value of the primary outcome. The second regression estimated the association between the postrandomization value of the primary outcome and the covariates, including the change in the putative mediator. The causally mediated effect was the product of the 2 regression coefficients, and the associated 95% CIs were estimated using bootstrapping with 1000 replicates.

Causal mediation analysis infers sequential ignorability. This premise was achieved in part by the CHAT design, which randomly produced the changes in the mediator. Another premise of sequential ignorability was satisfied by mitigating the implication of potential unmeasured confounders by including all of the covariates in the analysis. The robustness of this assumption is measured by a sensitivity analysis that estimates the threshold of potential confounding from unmeasured competing variables, at which the mediation effect becomes null.

The sample size of CHAT was adequate for an additional secondary analysis, including mediation analysis[12]: 453 pa-

Figure 1. General Causal Mediation Analysis and Magnitude of Changes in Behavior After Treatment of Obstructive Sleep Apnea
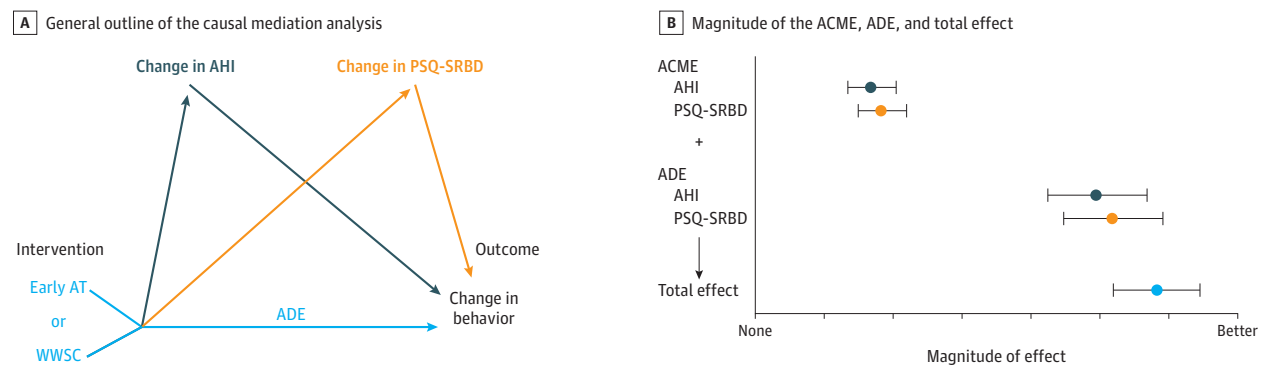
A, The average causal mediation effect (ACME) were the changes in a mediator variable (eg, Apnea Hypopnea Index [AHI] or the Pediatric Sleep Questionnaire–Sleep-Related Breathing Disorder [PSQ-SRBD] score) after the intervention (early adenotonsillectomy ([AT] vs watchful waiting with supportive care) was provided. The magnitude of ACME was obtained from the products of the coefficients of the 2 regressions (mediator modeled by the intervention and outcome modeled by the mediator) after controlling for potential confounders. The average direct effect (ADE) was the outcome of treatment that was not attributable to the putative mediation pathways. The ACME and ADE provided the total effect, which was estimated as the size of the intervention effect. B, The magnitude of the ACME, the ADE, and the total effect is shown with the error bars representing the associated bootstrapped measure of uncertainty (95% CIs).

tients yielded more than 80% power, with a type I error rate of 5% for detecting mediation effects.[30]

All statistical analyses were performed in R, version 3.6 (R Foundation for Statistical Computing), and the R-based mediation package.[31] Data analysis was conducted from January 1, 2020, to January 31, 2020.

Results

A total of 453 children were assessed at baseline and 407 children were assessed at follow-up. The mean (SD) age was 6.6 (1.4) years, 234 participants were girls (52%) and 219 were boys (48%), and 249 self-identified as Black individuals (55%). Obesity was reported in 150 children (33%). A total household income less than $30 000 was reported for 182 children (40%), and a maternal educational status of less than high school was reported for 140 children (31%). Baseline characteristics are summarized in **Table 1**.

The proportion of missing data was less than 5% for all variables except the teacher-reported behavioral outcomes and the household income values that approached 30%. These outcomes were, therefore, imputed before further analysis of the 379 children included based on the threshold for missing data. Regression models revealed a positive association between the baseline PSQ-SRBD score and the parent-reported severity of behavioral impairment (**Figure 2** and **Table 2**). However, no discernible associations could be identified between any of the polysomnographic parameters, including the AHI, and the behavioral outcomes (Figure 2 and Table 2). For example, parent-reported Conners Global Index scores were better predicted by PSQ-SRBD scores (regression coefficient for PSQ-SRBD, 37.0 [95% CI, 29.7-39.8]; regression coefficient for AHI, 0.0 [95% CI, –0.4 to 0.4]). The regression plots in Figure 2A-D illustrate the associations between the severity of upper-airway obstruction represented by the PSQ-SRBD score and the AHI and

Table 1. Baseline Study Characteristics of Children Enrolled in the Childhood Adenotonsillectomy Trial

Variable[a]	No. (%)
Age, mean (SD), y	6.6 (1.4)
Sex	
Male	219 (48)
Female	234 (52)
Race/ethnicity[b]	
White	156 (35)
Black	249 (55)
Other	48 (11)
BMI percentile for age, mean (SD)[c]	70.2 (30.8)
Obesity	150 (33)

Abbreviation: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared).

[a] Continuous variables are described as mean (SD), and categorical variables as numbers (%).

[b] Race/ethnicity was self-selected by the primary caregiver.

[c] A BMI in the 95th percentile or higher was considered a criterion for obesity.

the covariate-adjusted behavioral outcomes. Figure 2E shows the changes in overall variance (adjusted R^2) associated with the baseline model with the added predictor variable vs the baseline covariates-only model. For example, addition of PSQ-SRBD was associated with an adjusted R^2 of 0.3 vs an adjusted R^2 of 0.0 for the addition of AHI to the baseline covariates-only model. Specifically, these results demonstrate that the parent-reported severity of SDB measured by the PSQ-SRBD scale predicted the baseline behavioral outcomes without contribution from any of the polysomnographic parameters.

Early AT was associated with statistically significant improvements in both mediator variables (polysomnographic parameters and parent-reported SDB severity). The mean (SD)

Figure 2. Comparison of Polysomnographic Parameters to Parent-Reported Severity of Sleep-Disordered Breathing
in Predicting Baseline Behavioral Outcomes
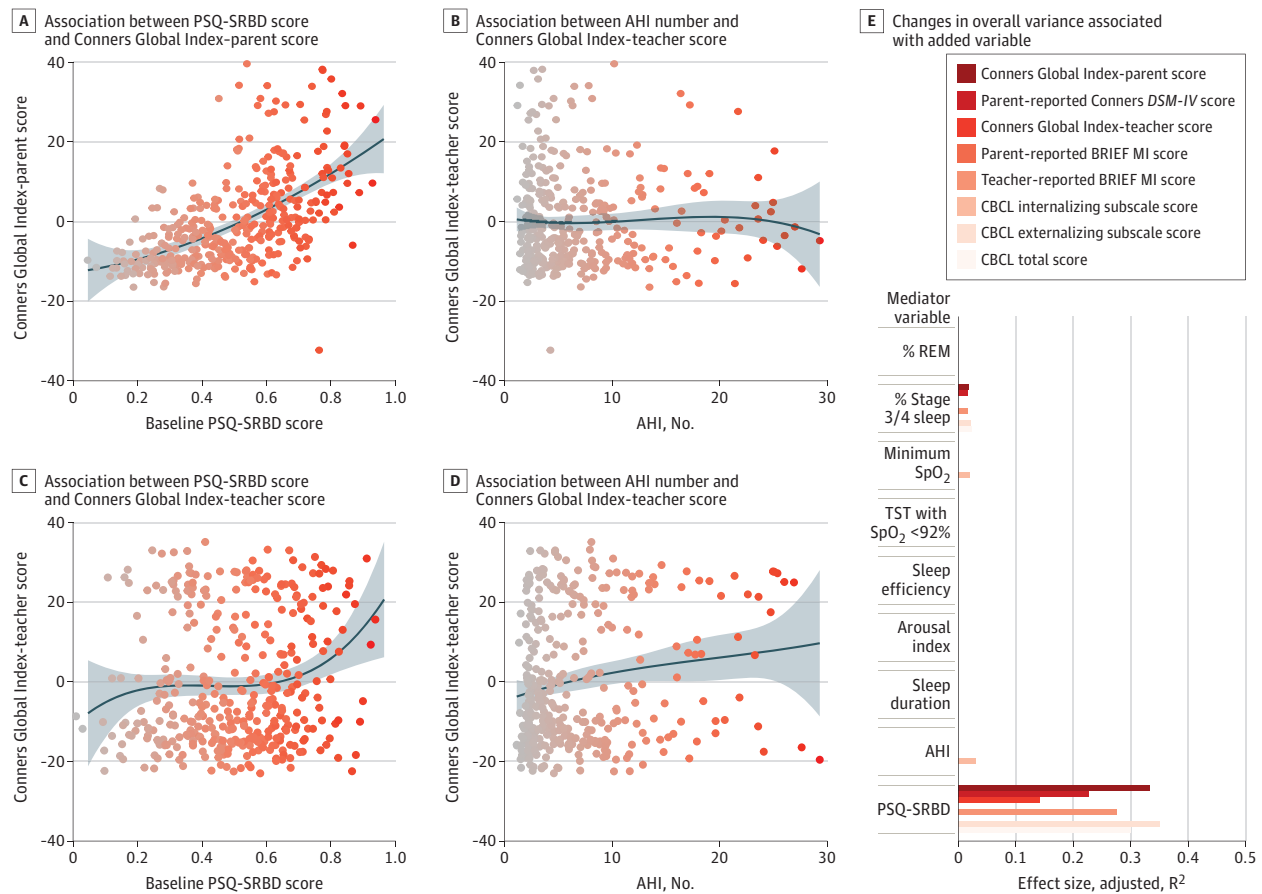
Shown are results of a generalized linear model fitted to 2 representative behavioral outcome measures (Conners Global Index parent and teacher versions) and the Pediatric Sleep Questionnaire–Sleep-Related Breathing Disorder (PSQ-SRBD) score (panels A and C) and the Apnea Hypopnea Index (AHI) (panels B and D), after adjustment for all covariates, including demographic and socioeconomic characteristics. The model was fitted with a smoothing spline, with the error bars spanning 1 SD for the horizontal extent of the distribution. Eight polysomnographic parameters were compared with the PSQ-SRBD score in predicting the various baseline behavioral outcomes (panel

E). The null-effect model included a regression comprising the covariates only and the baseline behavioral outcome of interest. The effect size was measured by estimating the change in proportion of variance (adjusted R^2) by adding the clinical or polysomnographic parameter to the covariates-only model. BRIEF MI indicates Behavior Rating Inventory of Executive Function metacognition index; CBCL, Child Behavior Checklist; *DSM-IV*, *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition); REM, rapid-eye-movement sleep; SpO$_2$, oxygen saturation as measured by pulse oximetry; and TST, total sleep time.

decrease in the AHI score was 5.2 (6.3) for those who underwent early AT and was 0.7 (6.4) for those randomized to WWSC (Cohen *d* = 0.6) between the baseline and follow-up periods. A greater mean (SD) decrease was identified for PSQ-SRBD scores in the early AT group vs WWSC group (0.3 [0.2] vs 0.0 [0.3]), also resulting in a greater effect size (Cohen *d* = 1.5).

The changes in behavioral outcomes are listed in Table 2. Although the parent-reported BRIEF MI score demonstrated the greatest improvement (Cohen *d* = 0.5) in the early AT group, smaller improvements were identified for other parent-reported behavioral outcomes. For example, the effect size for parent-reported Conners Global Index score was 0.3 in the early AT group. No statistically significant changes, as shown by the effect sizes, were identified for the teacher-reported BRIEF MI or the Conners Global Index scores. For example, the effect

sizes for teacher-reported Conners Global Index and BRIEF MI scores were 0.1 for each outcome.

Figure 3A compares the mediation effects of the PSQ-SRBD score and the AHI in capturing the treatment-related changes in each of the behavioral measures (eTable in the Supplement). Seven of 8 treatment-related changes in behavioral measures were causally attributable to the changes in PSQ-SRBD score (range of nonzero average causal mediation effects [ACME], 2.4-3.5; ACME represents the change in the mediator for unit change in the outcome variable and is therefore unitless). In contrast, the change in only 1 outcome (teacher-reported BRIEF MI score) was causally attributable to the changes in the AHI (ACME; 95% CI, –0.7 to –3.0). Figure 3B shows that the treatment-related changes in PSQ-SRBD score predicted the postrandomization changes in behavioral out-

Table 2. Summary of Changes in Behavioral Outcomes in the Childhood Adenotonsillectomy Trial

| Outcome measure | T score at baseline, mean (SD)[a] | Baseline regression coefficient (95% CI)[b] | | Postrandomization change in T score by treatment group, mean (SD)[c] | | |
		AHI	PSQ-SRBD	Early AT (n = 184)	WWSC (n = 195)	Effect size[d]
Parent-reported Conners Global Index	52.5 (11.7)	0.0 (-0.4 to 0.4)	37.0 (29.7 to 39.8)	-2.8 (10.2)	-0.3 (9.4)	0.30
Parent-reported Conners *DSM-IV*	60.0 (16.7)	0.1 (-0.3 to 0.8)	38.7 (33.6 to 43.8)	-3.1 (8.6)	-0.6 (9.2)	0.30
Teacher-reported Conners Global Index	50.3 (11.4)	0.6 (0.0 to 1.2)	10.3 (0.9 to 19.7)	-3.5 (19.6)	-1.6 (20.4)	0.01
Parent-reported BRIEF MI	59.1 (16.8)	-0.1 (-0.5 to 0.3)	34.7 (29.2 to 40.1)	-3.7 (8.1)	0.3 (9.0)	0.50
Teacher-reported BRIEF MI	59.6 (20.8)	0.2 (-0.4 to 0.8)	2.8 (-6.8 to 12.4)	-1.0 (22.8)	0.0 (23.5)	0.01
Parent-reported CBCL internalizing problem subscale	52.0 (11.6)	-0.1 (-0.5 to 0.3)	24.6 (18.5 to 30.7)	-4.0 (11.3)	-0.7 (9.5)	0.39
Parent-reported CBCL externalizing problem subscale	51.6 (11.3)	-0.3 (-0.7 to 0.1)	30.6 (24.9 to 36.3)	-2.3 (8.5)	-1.3 (7.9)	0.14
Parent-reported CBCL total problems	53.1 (11.0)	-0.1 (0.3 to 0.5)	36.2 (31.1 to 41.3)	-4.2 (8.7)	1.1 (8.3)	0.37

Abbreviations: AHI, Apnea Hypopnea Index; AT, adenotonsillectomy; BRIEF, Behavior Rating Inventory of Executive Function (parent version score range: 32-91; teacher version score range: 39-104; higher scores indicate worse behavior); CBCL, Child Behavior Checklist (total score range: 24-81; internalizing subscale score range: 33-89; externalizing subscale score range: 33-79; higher scores indicate worse behavior); *DSM-IV, Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition); MI, metacognition index; PSQ-SRBD, Pediatric Sleep Questionnaire–Sleep-Related Breathing Disorder (score range: 0.46-0.95, with higher scores indicating worse sleep disordered breathing); WWSC, watchful waiting with supportive care.

[a] Mean T scores for each of the behavioral outcomes were calculated by comparing the raw scores with the age- and sex-based normative data at baseline (n = 453).

[b] Regression coefficients with the 95% CIs are provided for the association between the severity of sleep-disordered breathing represented by the AHI or the PSQ-SRBD score.

[c] The relative magnitude of the postrandomization change in the behavioral outcomes and the associated effect size using Cohen *d* were calculated from the analysis of covariance adjusted for baseline values (n = 379).

[d] Effect size conventions were based on thresholds of Cohen *d* = 0.2 (small), Cohen *d* = 0.5 (medium), and Cohen *d* = 0.8 (large).

comes. For example, the change in PSQ-SRBD score predicted the change in parent-reported Conners Global Index score, whereas the change in AHI did not make such a prediction (ACME for PSQ-SRBD: 3.4 [95% CI, 2.2 to 4.9]; ACME for AHI: 0.29 [95% CI, -0.40 to 0.92]). None of the polysomnographic parameters representing sleep structure, obstruction, or hypoxemia mediated the treatment-related changes in outcomes. For example, the ACME for REM sleep was 0.0 (95% CI, -0.4 to 0.1). The sensitivity analysis (eFigure in the Supplement) also demonstrated that the mediation effects associated with the parent-reported behavioral outcomes were more robust than those reported by the teacher. No interactions were found between the 2 groups for any of the outcomes. For example, the ACME for PSQ-SRBD was 3.4 (95% CI, 1.3-5.3) in the WWSC group and was 3.4 (95% CI, 1.9 to 5.1) in the early AT group when the parent-reported outcomes were compared between the trial groups for Conners Global Index score. This finding indicates that the mediation effects associated with the changes in the AHI and the PSQ-SRBD score were not changed by the trial group to which children were randomized.
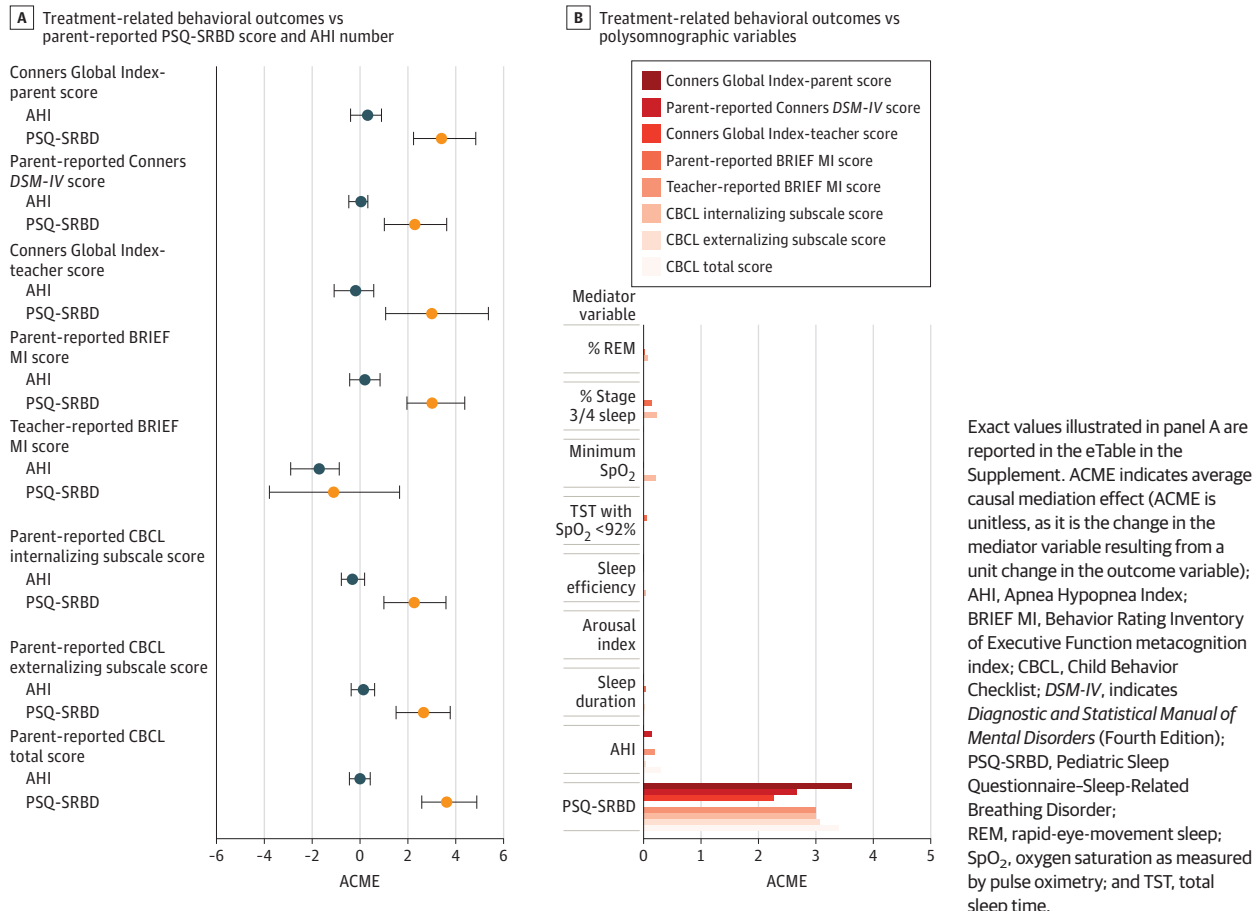
Discussion

In this secondary analysis of behavioral outcomes in the CHAT, the parent-reported severity of SDB measured by the PSQ-SRBD scale predicted most of the baseline behavioral measures in children with OSA. The results of this study demonstrated that the treatment-related changes in behavioral outcomes were causally attributable solely to the changes in parent-reported SDB severity. Furthermore, the

lack of identifiable associations between any of the polysomnographic parameters and the baseline or postrandomization behavioral measures suggested that polysomnographic parameters cannot be used in isolation to gauge the behavioral morbidity associated with pediatric OSA or its treatment-related changes.

The lack of a statistically significant association between polysomnographic severity of SDB and behavioral morbidity is a departure from the findings in some,[32,33] but not all,[6,34] previous studies. For example, Smith et al[35] reported that parent-reported frequency of snoring was a stronger predictor of neurobehavioral outcomes than the AHI. This result may be explained by the differences in demographic and socioeconomic characteristics, the age range, and the subsets of children (eg, those with primary snoring). Other investigators also hypothesized that polysomnographic parameters, such as sleep structure or hypoxemia, may explain the lack of dose dependence between behavioral problems and the AHI.[11] Therefore, the results from the current study are particularly illuminating because no discernible associations were found between any of the polysomnographic parameters and behavioral outcomes.

This analysis demonstrated an association between PSQ-SRBD scores and parent-reported behavioral outcomes, which was not replicated for the teacher-reported outcomes. This discordance was observed at baseline (Figure 2) and some of the treatment-related mediation effects (Figure 3). In the current study, the effect sizes were minuscule for the teacher-reported BRIEF MI and the Conners Global Index scores, potentially because of the extent of familiarity with a child.[36] The large CIs may explain the lack of precision in the data collected from the teacher, although

Figure 3. Comparison of Treatment-Related Changes in Behavioral Outcomes vs Parent-Reported Symptoms of Sleep-Disordered Breathing and Polysomnographic Parameters in Predicting Postrandomization Behavioral Outcomes

A Treatment-related behavioral outcomes vs parent-reported PSQ-SRBD score and AHI number

B Treatment-related behavioral outcomes vs polysomnographic variables

Exact values illustrated in panel A are reported in the eTable in the Supplement. ACME indicates average causal mediation effect (ACME is unitless, as it is the change in the mediator variable resulting from a unit change in the outcome variable); AHI, Apnea Hypopnea Index; BRIEF MI, Behavior Rating Inventory of Executive Function metacognition index; CBCL, Child Behavior Checklist; DSM-IV, indicates Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition); PSQ-SRBD, Pediatric Sleep Questionnaire–Sleep-Related Breathing Disorder; REM, rapid-eye-movement sleep; SpO₂, oxygen saturation as measured by pulse oximetry; and TST, total sleep time.

it is also possible that the expectations of the parents were modified by their participation in a randomized clinical trial.

Although the association of OSA with morbidity spans multiple domains, the most widely reported adverse association of OSA is with childhood behavior,[37] potentially affecting classroom performance and overall psychosocial functioning. A proposed model[37] links the severity of SDB to behavior through the structure and function of the prefrontal cortex, a key substrate for neurobehavioral function. This model supports a mechanistic hypothesis that polysomnography may contain useful information that captures, at least in part, this association by measuring the severity of both episodic hypoxia as well as sleep fragmentation.[14] This model also led the investigators of the CHAT to hypothesize that the changes in neurobehavioral and other outcome measures after treatment of OSA were causally attributable to the changes in severity of OSA or the resolution of OSA as measured by polysomnography.[10] Some treatment guidelines suggest that the principal surrogate end point associated with the treatment of OSA is the improvement in the AHI, with all children with an AHI higher than 5 requiring treatment, and postoperative polysomnography necessary for the implicit characterization of treatment response.[9] The results from the current secondary analysis, as well as 2 other studies[12,13] that also used the CHAT data set, instead support an approach in which polysomnographic data should not be used independent of the parent-reported symptoms of SDB.

Most of the 500 000 ATs performed each year in the US are solely based on clinical symptoms of SDB.[5] Pediatric polysomnography, which is used in less than 10% of the children undergoing AT, is promoted for its objective characterization of SDB and the associated cardiopulmonary variation. In some guidelines, a protocol of stepwise escalation of management has been suggested until the polysomnographic resolution of SDB.[9] This present study demonstrated that the association between parent-reported symptoms of SDB and childhood behavior had no obvious polysomnographic basis. The key practical implication of these results is that surgical candidacy for AT that is solely based on polysomnographic severity of SDB as measured by the AHI or other parameters is an unreliable predictor of behavioral outcomes. Instead, standardized symptom-based questionnaires, such as the PSQ-SRBD scale, should be considered as a useful adjunct in predicting and monitoring behavioral outcomes in children undergoing AT.

Strengths and Limitations

This study has some strengths, which are mainly associated with the analysis of a national, diverse, and rigorous data set. Causal mediation analysis is a robust approach for interrogating the mechanistic basis of treatment-related outcomes of OSA, specifically using bias mitigation efforts such as randomization, standardized follow-up, and central interpretation of polysomnographic studies. Thus, causal mediation analysis provides good generalizability.

This study has some limitations, including the narrow range of age and the AHI as well as the short duration of follow-up. In addition, parental reporting may be biased or inflated because of treatment-related expectations. Furthermore, the shared variance across behavioral questions in the PSQ-SRBD scale may account for some of the baseline and the mediation associations between PSQ-SRBD scores and behavioral outcomes. Among the various outcome measures, the nonresponse rate for the teacher-reported questionnaires, representing a relatively unbiased measurement, was high at approximately 30%.

Conclusion

This study reported the highest-quality evidence of the association between polysomnographic parameters and behavioral outcomes in children with OSA. Given that bad behavior is the most consistently reported adverse effect of untreated upper-airway obstruction, we demonstrated that clinical assessment of SDB remains the sole outcome measure of treatment-related changes in behavioral morbidity. Data from the CHAT, the only national randomized clinical trial in the domains of behavior and quality of life, appear to suggest that polysomnographic thresholds, including their relative change over the course of treatment, provide clinicians with limited means to predict the improvement in neurobehavioral morbidity in OSA.

REFERENCES

1. Katz ES, D'Ambrosio CM. Pathophysiology of pediatric obstructive sleep apnea. *Proc Am Thorac Soc*. 2008;5(2):253-262. doi:10.1513/pats.200707-111MG

2. Mitchell RB, Archer SM, Ishman SL, et al. Clinical practice guideline: tonsillectomy in children (update). *Otolaryngol Head Neck Surg*. 2019;160(1, suppl):S1-S42. doi:10.1177/0194599818801757

3. Aurora RN, Zak RS, Karippot A, et al; American Academy of Sleep Medicine. Practice parameters for the respiratory indications for polysomnography in children. *Sleep*. 2011;34(3):379-388. doi:10.1093/sleep/34.3.379

4. Marcus CL, Brooks LJ, Draper KA, et al; American Academy of Pediatrics. Diagnosis and management of childhood obstructive sleep apnea syndrome. *Pediatrics*. 2012;130(3):576-584. doi:10.1542/peds.2012-1671

5. Mitchell RB, Pereira KD, Friedman NR. Sleep-disordered breathing in children: survey of current practice. *Laryngoscope*. 2006;116(6):956-958. doi:10.1097/01.MLG.0000216413.22408.FD

6. Smith DL, Gozal D, Hunter SJ, Philby MF, Kaylegian J, Kheirandish-Gozal L. Impact of sleep disordered breathing on behaviour among elementary school-aged children: a cross-sectional analysis of a large community-based sample. *Eur Respir J*. 2016;48(6):1631-1639. doi:10.1183/13993003.00808-2016

7. Chervin RD, Weatherly RA, Ruzicka DL, et al. Subjective sleepiness and polysomnographic correlates in children scheduled for adenotonsillectomy vs other surgical care. *Sleep*. 2006;29(4):495-503.

8. Yan XH, Zhao Y, Wang J, et al. Associations among sleep symptoms, physical examination, and polysomnographic findings in children with obstructive sleep apnea. *Eur Arch Otorhinolaryngol*. 2020;277(2):623-630. doi:10.1007/s00405-019-05719-8

9. Kaditis AG, Alonso Alvarez ML, Boudewyns A, et al. Obstructive sleep disordered breathing in 2- to 18-year-old children: diagnosis and management. *Eur Respir J*. 2016;47(1):69-94. doi:10.1183/13993003.00385-2015

10. Redline S, Amin R, Beebe D, et al. The Childhood Adenotonsillectomy Trial (CHAT): rationale, design, and challenges of a randomized controlled trial evaluating a standard surgical procedure in a pediatric population. *Sleep*. 2011;34(11):1509-1517. doi:10.5665/sleep.1388

11. Marcus CL, Moore RH, Rosen CL, et al; Childhood Adenotonsillectomy Trial (CHAT). A randomized trial of adenotonsillectomy for childhood sleep apnea. *N Engl J Med*. 2013;368(25):2366-2376. doi:10.1056/NEJMoa1215881

12. Isaiah A, Pereira KD, Das G. Polysomnography and treatment-related outcomes of childhood sleep apnea. *Pediatrics*. 2019;144(4):e20191097. doi:10.1542/peds.2019-1097

13. Isaiah A, Bertoni D, Pereira KD, Diaz-Abad M, Mitchell RB, Das G. Treatment-related changes in heart rate variability in children with sleep apnea. *Otolaryngol Head Neck Surg*. 2020;162(5):737-745. doi:10.1177/0194599820907882

14. Biggs SN, Nixon GM, Horne RSC. The conundrum of primary snoring in children: what are we missing in regards to cognitive and behavioural morbidity? *Sleep Med Rev*. 2014;18(6):463-475. doi:10.1016/j.smrv.2014.06.009

15. Dean DA II, Goldberger AL, Mueller R, et al. Scaling up scientific discovery in sleep medicine: the National Sleep Research Resource. *Sleep*. 2016;39(5):1151-1164. doi:10.5665/sleep.5774

16. Isaiah A, Hamdan H, Johnson RF, Naqvi K, Mitchell RB. Very severe obstructive sleep apnea in children: outcomes of adenotonsillectomy and risk factors for persistence. *Otolaryngol Head Neck Surg*. 2017;157(1):128-134. doi:10.1177/0194599817700370

17. Conners CK. *Conners' Rating Scale Manual*. Multi-Health Systems, Inc; 1989.

18. Conners CK, Sitarenios G, Parker JD, Epstein JN. The revised Conners' Parent Rating Scale (CPRS-R): factor structure, reliability, and criterion validity. *J Abnorm Child Psychol*. 1998;26(4):257-268. doi:10.1023/A:1022602400621

19. Conners CK, Sitarenios G, Parker JD, Epstein JN. Revision and restandardization of the Conners Teacher Rating Scale (CTRS-R): factor structure, reliability, and criterion validity. *J Abnorm Child Psychol*. 1998;26(4):279-291. doi:10.1023/A:1022606501530

20. Gioia GA, Isquith PK, Guy SC, Kenworthy L. *Behavior Rating Inventory of Executive Function (BRIEF)*. Psychological Assessment Resources; 2000.

21. Achenbach TM, Rescorla L. *Manual for the ASEBA School-Age Forms & Profiles: An Integrated System of Multi-Informant Assessment*. ASEBA; 2001.

22. Chervin RD, Hedger K, Dillon JE, Pituch KJ. Pediatric sleep questionnaire (PSQ): validity and reliability of scales for sleep-disordered breathing, snoring, sleepiness, and behavioral problems. *Sleep Med*. 2000;1(1):21-32. doi:10.1016/S1389-9457(99)00009-X

23. Chervin RD, Weatherly RA, Garetz SL, et al. Pediatric sleep questionnaire: prediction of sleep apnea and outcomes. *Arch Otolaryngol Head Neck Surg*. 2007;133(3):216-222. doi:10.1001/archotol.133.3.216

24. Kuczmarski RJ, Ogden CL, Grummer-Strawn LM, et al. CDC growth charts: United States. *Adv Data*. 2000;(314):1-27.

25. Honaker J, King G, Blackwell M, Amelia II. A program for missing data. *J Stat Softw*. 2011;45(7):1-47. doi:10.18637/jss.v045.i07

26. Cohen J. A power primer. *Psychol Bull*. 1992;112(1):155-159. doi:10.1037/0033-2909.112.1.155

27. Faraway JJ. Generalized linear models. In: Faraway JJ. *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. 2nd ed. Chapman and Hall/CRC; 2016. doi:10.1201/9781315382722-10

28. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychol Methods*. 2010;15(4):309-334. doi:10.1037/a0020761

29. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*. 1986;51(6):1173-1182. doi:10.1037/0022-3514.51.6.1173

30. Schoemann AM, Boulton AJ, Short SD. Determining power and sample size for simple and complex mediation models. *Social Psychological and Personality Science*. 2017;8(4):379-386. doi:10.1177/1948550617715068

31. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. mediation: R package for causal mediation analysis. *J Stat Softw*. 2014;59(5):1-38. doi:10.18637/jss.v059.i05

32. Reeves G, Blaisdell C, Lapidus M, et al. Sleep architecture and behavioral abnormalities in children and adolescents. *Int J Adolesc Med Health*. 2010;22(4):535-545. doi:10.1515/IJAMH.2010.22.4.535

33. Lewin DS, Rosen RC, England SJ, Dahl RE. Preliminary evidence of behavioral and cognitive sequelae of obstructive sleep apnea in children. *Sleep Med*. 2002;3(1):5-13. doi:10.1016/S1389-9457(01)00070-3

34. O'Brien LM, Mervis CB, Holbrook CR, et al. Neurobehavioral correlates of sleep-disordered breathing in children. *J Sleep Res*. 2004;13(2):165-172. doi:10.1111/j.1365-2869.2004.00395.x

35. Smith DL, Gozal D, Hunter SJ, Kheirandish-Gozal L. Frequency of snoring, rather than apnea-hypopnea index, predicts both cognitive and behavioral problems in young children. *Sleep Med*. 2017;34:170-178. doi:10.1016/j.sleep.2017.02.028

36. Kohler MJ, Kennedy JD, Martin AJ, Lushington K. Parent versus teacher report of daytime behavior in snoring children. *Sleep Breath*. 2013;17(2):637-645. doi:10.1007/s11325-012-0736-9

37. Beebe DW, Gozal D. Obstructive sleep apnea and the prefrontal cortex: towards a comprehensive model linking nocturnal upper airway obstruction to daytime cognitive and behavioral deficits. *J Sleep Res*. 2002;11(1):1-16. doi:10.1046/j.1365-2869.2002.00289.x