Improved Behavior and Sleep After Adenotonsillectomy in Children With Sleep-Disordered Breathing

Julie L. Wei, MD; Matthew S. Mayo, PhD; Holly J. Smith, MA; Matt Reese, PhD; Robert A. Weatherly, MD

Objective: To determine changes in behavior and sleep in children before and after adenotonsillectomy for sleep-disordered breathing (SDB) using the validated Pediatric Sleep Questionnaire (PSQ) and Conners’ Parent Rating Scale–Revised Short Form (CPRS-RS).

Design: Prospective, nonrandomized study.

Setting: Ambulatory surgery center affiliated with an academic medical center.

Patients: A total of 117 consecutive children (61 boys and 56 girls) (mean [SD] age, 6.5 [3.1] years) who were clinically diagnosed as having SDB and who had undergone adenotonsillectomy. Complete follow-up data were available in 71 of 117 patients (61%).

Interventions: Parents completed the PSQ and CPRS-RS before surgery and 6 months after surgery.

Main Outcome Measures: Changes in age- and sex-adjusted T scores for all 4 CPRS-RS behavior categories (oppositional behavior, cognitive problems or inattention, hyperactivity, and ADHD index) were determined for each subject before and after surgery. Changes in PSQ scores from a select 22-item sleep-related breathing disorder subscale were also determined.

Results: Preoperatively, the mean (SD) T scores on the CPRS-RS for oppositional behavior, cognitive problems or inattention, hyperactivity, and ADHD index were 59.4 (13.7), 59.5 (13.6), 62.0 (14.4), and 59.9 (13.4), respectively. A T score of 60.0 in any category placed a child in the at-risk group. Postoperatively, T scores for each category were 51.0 (9.6), 51.2 (8.8), 52.4 (10.52), and 50.6 (7.8), respectively. All changes were statistically significant (P<.001) and clinically significant by approximating a change of 1 SD from the baseline score. For the PSQ, the preoperative and postoperative mean (SD) scores were 0.6 (0.1) and 0.1 (0.1), respectively, on a scale of 0 to 1, with scores higher than 0.33 suggesting obstructive sleep apnea. Correlations between sleep and behavior scores were statistically significant before surgery (P=.004 for ADHD index and cognitive problems, P=.008 for oppositional behavior) and after surgery (P=.049 for cognitive problems, P=.03 for oppositional behavior). Higher baseline T scores for the CPRS-RS were associated with larger changes in T scores for the CPRS-RS in all 4 domains (oppositional behavior, cognitive problems or inattention, hyperactivity, and ADHD index).

Conclusions: Children diagnosed as having SDB experience improvement in both sleep and behavior after adenotonsillectomy. The PSQ and CPRS-RS may be useful adjuncts for screening and following children who undergo adenotonsillectomy for SDB.
associated with ADHD, and it has been suggested that treatment of SDB may eliminate ADHD in a subset of children if their habitual snoring and SDB were alleviated. Unlike OSAS, which is defined in part by a specific apnea-hypopnea index (AHI) based on PSG, SDB may be diagnosed clinically and may not consistently meet PSG criteria for an obstructive sleep breathing disorder.

In addition to hyperactivity, children with snoring and SDB have been shown to have neurocognitive impairment and poor school performance. In one study, children without clinically significant hypoxia, as measured by pulse oximetry, but with habitual snoring have been shown to have poorer academic performance than those who do not snore. Sleep-disordered breathing is also associated with enuresis, learning disabilities, daytime sleepiness, and somatic complaints such as headaches. Behavioral and emotional difficulties have been found in children with SDB before interventional, and improvements have been reported after adenotonsillectomy.

According to the recommendations of the American Academy of Pediatrics and American Thoracic Society, a positive finding from PSG or some other objective measure (eg, a cardiopulmonary nocturnal study) is required to establish the diagnosis of OSAS. Although PSG performed overnight in a laboratory setting is considered the gold standard for evaluation of OSAS, there are several reasons why PSG alone may be insufficient to identify all children with SDB. First, patients with primary snoring and upper airway resistance syndrome may show a normal respiratory disturbance index on a standard sleep study. Second, Weatherly et al surveyed practice patterns of otolaryngologists and found that less than 10% of patients received any objective testing and less than 5% of school-aged children received PSG prior to adenotonsillectomy. This practice pattern deviates from the recommendations of the American Academy of Pediatrics. The actual practice pattern of otolaryngologists may reflect the fact that there are not enough sleep centers or resources to have preoperative PSG testing for each child undergoing adenotonsillectomy. Third, children with clinically diagnosed SDB may not consistently meet PSG criteria for OSAS. In the opinion of some otolaryngologists, PSG testing should be reserved for patients of very young age or those with medical comorbidities.

Finally, there are incomplete normative sleep data on the pediatric population, and variability in sleep study performance and interpretation results in an inconsistent diagnosis of SDB or OSAS depending on the definition chosen, which is based on the AHI. Although an AHI as low as 1 may be unusual in healthy children, it is not known whether this value is clinically significant.

Alternative instruments, such as quality-of-life (QOL) measures, could be used for children who may not need PSG testing. Some studies using various instruments have demonstrated that adenotonsillectomy results in marked improvement in QOL, not only in children who have OSAS documented by PSG, but also in those clinically diagnosed as having SDB without confirmation by PSG. Although such instruments are helpful in measuring outcome, a questionnaire specifically for predicting SDB and improvement after surgery would bridge the gap between objective PSG data on every child with nocturnal airway obstruction on one end of the spectrum and clinical suspicion of SDB in children at the other end. Chervin et al developed the Pediatric Sleep Questionnaire (PSQ), and its scales for SDB, snoring, sleepiness, and behavioral problems were shown to have good validity and reliability. The PSQ may be used as a tool in clinical research when PSG may not be feasible, as was the case in our study. We chose the Conners' Parent Rating Scale–Revised Short Form (CPRS-RS) survey to assess behavior in our subjects before and after adenotonsillectomy. This study was modeled after previous studies examining QOL improvement after adenotonsillectomy. We designed this study to examine improvements in both sleep and behavior at 6 months after adenotonsillectomy without using PSG. Patients were excluded if the primary reason for adenotonsillectomy was recurrent infections. Patients were not selected for participation in this study because parents had reported behavioral problems. Children were diagnosed as having SDB by the pediatric otolaryngologists based on medical history and physical examination with or without flexible fiber-optic nasopharyngoscopy. All patients were otherwise healthy and without any medical comorbidity such as medical history of congenital heart disease, neurologic disorders such as cerebral palsy, Down syndrome, or craniofacial abnormalities.

### METHODS

This study was conducted at a pediatric otolaryngology practice based at a university hospital satellite center. The study protocol was approved by the institutional review board at the University of Kansas Hospital in Kansas City. A total of 117 consecutive children who were clinically diagnosed as having SDB and who were scheduled to undergo adenotonsillectomy were enrolled (Table 1), and informed consent for enroll-
PSQ because these were the items found to have the strongest average; 60.0 to 65.0, above average; 65.1 to 70.0, much above average; 56.0 to 60.9, slightly above average; 55.0 to 59.9, just above average; 45.0 to 55.0, on a par with the average; 39.0 to 44.9, below average; and group 4, T scores higher than 70.0 (defined as those having preoperative mean T scores of 50.0 or higher). For each behavior category, the mean T scores were 62.0 (14.5) preoperatively and 52.4 (10.5) postoperatively (P < .001). Reductions in mean T scores were also noted for the behavior categories (oppositional behavior, cognitive problems or inattention, hyperactivity, and ADHD index) (Table 2). The change in mean T scores for all respondents (from before surgery to after surgery) was determined for each behavior category. To see if there were any clinically significant changes 6 months after surgery, we then calculated the Pearson correlation coefficient, for before and after surgery, between the PSQ with and without the behavioral questions and the T scores (for oppositional behavior, cognitive problems or inattention, hyperactivity, and the ADHD index scales). Finally, we assessed the change in PSQ related to age, sex, prior treatment with methylphenidate hydrochloride, and prior diagnosis of attention deficit disorder or ADHD using linear regression. We examined the association of preoperative T scores for each measure with the postoperative change in that specific T score using 1-way analysis of variance and conducted pairwise comparisons across the groups. All analyses were conducted at the P = .05 level of significance.

### RESULTS

A total of 117 patients and their caregivers were enrolled in the study, and 71 completed the 6-month follow-up. Forty-six patients did not complete the study because follow-up questionnaires were not returned by their caregivers. The mean (SD) age of the patients was 6.5 (3.1) years. We found no differences in age, ethnicity, sex, or baseline scores between those who completed the study vs those who did not. Demographic data are shown in Table 1. The change in mean T scores for all 4 CPRS-RS categories is shown in Table 2. From the CPRS-RS profiles, the mean (SD) T scores for all 71 subjects for oppositional behavior were 59.4 (13.7) preoperatively and 51.0 (9.6) postoperatively at the 6-month follow-up (P < .001). For the cognitive problems or inattention category, the mean T scores were 59.5 (13.6) preoperatively and 51.2 (8.8) postoperatively (P < .001). For the hyperactivity category, the mean T scores were 62.0 (14.4) preoperatively and 52.4 (10.5) postoperatively (P < .001). The mean (SD) ADHD index T scores were 59.9 (13.4) preoperatively and 50.6 (7.8) postoperatively (P < .001). Reductions in mean T scores were statistically significant in all 4 categories as shown, and the T score decreased by close to 10 points between preoperative and postoperative values, which is considered clinically significant.

Once we found the reduction in T scores to be statistically significant for the entire group, we explored whether...
the changes were a function of baseline T scores for each category. We wanted to know if the worse the patient's preoperative sleep and behavior scores, the greater the reduction in T scores. The 71 patients were stratified into 4 groups for each CPRS-RS category based on their preoperative mean T score as described in the previous paragraph. The mean change in CPRS-RS T scores in each behavior scale was calculated for each group, and results are shown in Table 3. For example, for the ADHD index, 27 patients were in group 1 (those with a preoperative mean T score ≤ 50.0), which would be considered below average and not indicative of being at risk for ADHD, and their mean (SD) postoperative T score change was 2.9 (9.0). However, for groups 2, 3, and 4, whose preoperative mean T scores were 50.1 to 60.9, 61.0 to 70.0, and higher than 70.0, respectively, we found that the higher the mean preoperative T score, the greater the change in T score after adenotonsillectomy. Pairwise comparisons among the groups were statistically significant in each category except for the comparisons between groups 2 and 3 for oppositional behavior (P = .052; correlation coefficient, 0.23), and hyperactivity (P = .20; correlation coefficient, 0.15).

Although PSG is considered the gold standard for evaluation of OSAS, for select patients PSG testing may not be necessary prior to adenotonsillectomy. There are no clinical practice guidelines regarding the use of PSG and candidacy for adenotonsillectomy, and most otolaryngologists make decisions or recommendations based on medical history that is suggestive of SDB, as well as on physical examination demonstrating adenoid and/or tonsillar hypertrophy. However, predictive accuracy of clinically suspected OSAS may be as low as 30% when PSG testing was performed as well.28,29

In patients with adenotonsillar hypertrophy, adenotonsillectomy is potentially curative for the whole spectrum of obstructive sleep disorders, including primary snoring and upper airway resistance syndrome, which may not show abnormalities on standard PSG. Children who undergo adenotonsillectomy for PSG-defined OSAS demonstrate marked improvement in both QOL scores and PSG parameters compared to those who did not have surgery, but no strong association was identified between QOL scores and PSG parameters.24 Although clinical history is improved after adenotonsillectomy, postoperative PSG may still indicate an abnormality depending on the parameters used for defining an abnormal PSG study. In addition to clinical suspicion of SDB, using an instrument such as the PSQ may help to identify patients who are appropriate candidates for adenotonsillectomy or those who may have residual SDB or associated symptoms that fail to improve after surgery.

### Table 3. Changes in CPRS-RS T Scores Grouped by Preoperative T Scores

<table>
<thead>
<tr>
<th>Behavior Category</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
<th>Pairwise Comparison P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADHD index</td>
<td>2.9 (9.0)</td>
<td>27</td>
<td>-8.0 (6.8)</td>
<td>33</td>
<td>-2.1 (6.0)</td>
</tr>
<tr>
<td>Cognitive problems or inattention</td>
<td>2.9 (6.8)</td>
<td>37</td>
<td>-5.7 (6.3)</td>
<td>28</td>
<td>-11.1 (10.0)</td>
</tr>
<tr>
<td>Hyperactivity</td>
<td>3.2 (6.1)</td>
<td>28</td>
<td>-7.1 (7.4)</td>
<td>37</td>
<td>-6.8 (7.3)</td>
</tr>
<tr>
<td>Oppositional</td>
<td>1.0 (5.5)</td>
<td>36</td>
<td>-6.6 (6.8)</td>
<td>28</td>
<td>-11.7 (7.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001c</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001d</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001e</td>
</tr>
</tbody>
</table>

Abbreviations: ADHD, attention-deficit/hyperactivity disorder; CRPS-RS, Conners’ Parent Rating Scale–Revised Short.

* a T scores are standardized scores for each scale. Ranges for scores for the groups are as follows: group 1, 50.0; group 2, 50.1-60.9; group 3, 61.0-70.0, and group 4, > 70.0. Data for changes are given as mean (SD).

©2007 American Medical Association. All rights reserved.
It has been demonstrated that predictive accuracy for SDB is much higher when parental report of multiple measures of behavior is combined with medical history of snoring. Therefore, we chose the PSQ as a comprehensive parental questionnaire for assessing quality of sleep and the CPRS-RS for assessing child behavior. The Conners’ scales are the most widely used instruments for assessing child behavior, even in the school setting. The CPRS-RS can be used for research, screening, or monitoring treatment effects over time and can also be used as a diagnostic tool in conjunction with other information, such as the short versions of the Conners’ Teacher Rating Scale-Revised. Using both teachers’ and parents’ reports would certainly provide more information regarding the global behavior of the child. However, behavior at school and behavior at home may not correlate and may vary drastically for every child.

Although 46 of 117 subjects did not complete the study (39%), because of the consistent reduction in T scores for each CPRS-RS category in subjects who completed the study, both as a group and for each subgroup, we believe that our findings would be similar had more subjects completed the study. Questionnaires were mailed to the known mailing addresses supplied by subjects, and attempts were made to ensure receipt of questionnaires by the subjects as well as to encourage follow-up. Our data showed that at 6 months after adenotonsillectomy, there was an average reduction in T score by approximately 1 SD in every CPRS-RS category: oppositional behavior, cognitive problems or inattention, hyperactivity, and the ADHD index. This difference of close to 1 SD is considered clinically significant to pediatric behavioral specialists who use this instrument to measure changes in behavior. A T score of 60.0 in any category indicates that a child is at risk for problems in that category, and T score of 65.0 or higher indicates that the child should be formally evaluated for possible ADHD. A hyperactivity index or T score higher than 60.0 results in the child as being rated hyperactive. A T score higher than 70.0 is clinically significant and may be diagnostic for ADHD. Reduction of T scores by 2 SDs is clinically significant and is expected to be reflected by clinical changes in behavior. The mean preoperative T score was 59.0 for oppositional behavior, cognitive problems or inattention, and the ADHD index and 62.0 for hyperactivity. Such scores are indicative of the enrolled patients being at risk for each behavior category and support the association between SDB and behavior because these are patients who underwent adenotonsillectomy for clinically diagnosed SDB.

By stratifying the patients into 4 groups based on their preoperatively mean T scores, we found that the higher the preoperative T score, the greater the reduction in mean T scores after intervention in all 4 behavior categories measured by the CPRS-RS. Group 1 had the lowest mean preoperative T score, 50.0 or lower, and had the smallest room for improvement. We believe the minor postoperative increase in the mean T scores (less than 1 SD) is just a function of random variation in that group. We first performed a global test to determine if the mean change differed across the 4 groups, and because we saw that it did, pairwise comparisons were performed to show where differences lie among the groups. Although the number of subjects in each group was different for each CPRS-RS category, the trend was consistent in every category such that as the mean preoperative T score increased, there was a greater reduction in mean T scores after adenotonsillectomy. We consistently found the greatest reduction in mean T scores in group 4, whose baseline T scores in each category were higher than 70.0. This group had reductions in mean T score for each category, approximating 2 SDs.

Chervin et al found that a PSQ score higher than 0.33 suggests a high risk for SDB with reasonable sensitivity and specificity. We calculated preoperative and postoperative scores for both the full PSQ (the 22-item questionnaire, including the 6 behavioral items duplicated in the CPRS-RS) and the PSQm6. Both sets of scores are presented in Table 2. To demonstrate that SDB is related to child behavior, we wanted to evaluate the PSQ without the behavioral component. However, to demonstrate the clinical usefulness of the PSQ, the additional data on the 22-item instrument is valuable because if there is a strong correlation between the PSQ and CPRS, then perhaps the CPRS would not be necessary. Of the 117 patients who were enrolled in the study, 112 (95.7%) had PSQ and PSQm6 scores higher than 0.33, and 5 (4.3%) had PSQ and PSQm6 scores of 0.33 or lower preoperatively. Our data showed that for this set of patients, preoperative mean PSQ and PSQm6 scores were similar: 0.5 (0.2) and 0.6 (0.1), respectively. Both scores supported the clinical suspicion of SDB based on the medical history and physical examination. At 6 months after adenotonsillectomy, the mean (SD) for the PSQ and PSQm6 scores for the 71 patients who completed the study was 0.1 (0.1). Of these 71 patients, 7 (9.9%) still had PSQ and PSQm6 scores higher than 0.33, whereas 64 (90.1%) had PSQ and PSQm6 scores of 0.33 or less. This finding suggests that most patients experience improved sleep after adenotonsillectomy and alleviation of upper airway obstruction. It is possible that owing to the subjective nature of the caregiver-based, questionnaire-type study, the postoperative scores improved partly owing to expectations of improvement. Based on our study design, at 6 months after adenotonsillectomy, the reductions in mean T scores in CPRS-RS categories may be a reflection of a statistical regression to the mean. However, our results showed a consistent trend in the reduction of postoperative T scores as a function of preoperative T scores, as well as consistent reduction of mean T scores in all CPRS-RS categories, which supports our conclusion that adenotonsillectomy can improve a child’s behavior as measured by an instrument such as the CPRS-RS.

We found correlations between sleep and behavior before and after adenotonsillectomy; that is, as PSQ and PSQm6 scores approach 1, the higher the T score for each of the CPRS-RS behavior categories. This correlation was statistically significant for oppositional behavior, cognitive problems or inattention, and the ADHD index before the surgery, and statistically significant for cognitive problems of inattention and oppositional behavior after the surgery (see Table 3 for P values). Not only did both behavior and sleep improve independently before and after adenotonsillectomy for SDB in our group of patients, but they also improved in correlation with each other. The correlation between the PSQm6 scores and the T scores for each of the CPRS-RS behavior categories support our conclusion that SDB is related to child
behavior. By also calculating changes in the full PSQ score (22 items, including the 6 behavioral items), which were very similar to changes in the PSQm6 score, we have demonstrated the usefulness of the PSQ as a tool for evaluating both sleep and behavior. In fact, because the PSQ correlated very well to the CPRS-RS in this study, it may be unnecessary to use the CPRS-RS instrument in future studies evaluating pediatric sleep and behavior for SDB secondary to adenotonsillar hypertrophy.

This is an observational study that has demonstrated associations between adenotonsillitis and changes in sleep and behavior as shown by the PSQ and CPRS-RS instrument scales. Although this study supports a cause-and-effect relationship between SDB and development of behavioral problems, because it is not randomized or controlled as an intervention trial, this relationship cannot be proven. Without a control group that did not undergo adenotonsillectomy, we cannot prove definitively that the surgical intervention was the cause of the change in behavior as measured by the CPRS-RS. However, the PSQ and CPRS have been used together in a 4-year prospective cohort study demonstrating that symptoms of SDB actually precede development of hyperactivity. Such findings lend validity to our data, which shows that treatment of SDB can lead to amelioration of behavior and sleep problems.

Because PSG testing for each subject was not feasible and was incongruent with clinical practice, we chose a consecutive sample for patients who presented with a medical history of nocturnal airway obstruction and possible obstructive sleep apnea. This exploratory study demonstrated the potential usefulness of the PSQ as a screening tool and as an adjunct to clinical history and physical examination in determining candidacy for adenotonsillectomy. A 3-year follow-up of the same patients may demonstrate whether such improvements in sleep and behavior are maintained over time.

Submitted for Publication: November 11, 2006; final revision received April 10, 2007; accepted May 3, 2007.

Correspondence: Julie L. Wei, MD, Department of Otolaryngology—Head and Neck Surgery, University of Kansas Medical Center, 3901 Rainbow Blvd, Mail Stop 3010, Kansas City, KS 66160 (jwei@kumc.edu).

Author Contributions: Drs Wei, Mayo, Reese and Weatherly and Ms Smith had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Wei, Reese, and Weatherly. Acquisition of data: Wei and Smith. Analysis and interpretation of data: Mayo, Smith, Reese, and Weatherly. Drafting of the manuscript: Wei and Mayo. Critical revision of the manuscript for important intellectual content: Wei, Mayo, Smith, Reese, and Weatherly. Statistical analysis: Mayo. Administrative, technical, and material support: Reese and Weatherly. Study supervision: Wei and Weatherly. Data management and database development: Smith.

Financial Disclosure: None reported.

Additional Contributions: The late Debra Park, PhD, provided assistance in the initial design of this study. Ronald Chervin, MD, provided the use of the PSQ and input on the manuscript.

REFERENCES