A Randomized Trial of Adenotonsillectomy for Childhood Sleep Apnea

Carole L. Marcus, M.B., B.Ch., Renée H. Moore, Ph.D., Carol L. Rosen, M.D., Bruno Giordani, Ph.D., Susan L. Garetz, M.D., H. Gerry Taylor, Ph.D., Ron B. Mitchell, M.D., Raouf Amin, M.D., Eliot S. Katz, M.D., Raanan Arens, M.D., Shalini Paruthi, M.D., Hiren Muzumdar, M.D., Susan S. Ellenberg, Ph.D., and Susan Redline, M.D., M.P.H., for the Childhood Adenotonsillectomy Trial (CHAT)

BACKGROUND

Adenotonsillectomy is commonly performed in children with the obstructive sleep apnea syndrome, yet its usefulness in reducing symptoms and improving cognition, behavior, quality of life, and polysomnographic findings has not been rigorously evaluated. We hypothesized that, in children with the obstructive sleep apnea syndrome without prolonged oxyhemoglobin desaturation, early adenotonsillectomy, as compared with watchful waiting with supportive care, would result in improved outcomes.

METHODS

We randomly assigned 464 children, 5 to 9 years of age, with the obstructive sleep apnea syndrome to early adenotonsillectomy or a strategy of watchful waiting. Polysomnographic, cognitive, behavioral, and health outcomes were assessed at baseline and at 7 months.

RESULTS

The average baseline value for the primary outcome, the attention and executive-function score on the Developmental Neuropsychological Assessment (with scores ranging from 50 to 150 and higher scores indicating better functioning), was close to the population mean of 100, and the change from baseline to follow-up did not differ significantly according to study group (mean [±SD] improvement, 7.1±13.9 in the early-adenotonsillectomy group and 5.1±13.4 in the watchful-waiting group; P=0.16). In contrast, there were significantly greater improvements in behavioral, quality-of-life, and polysomnographic findings and significantly greater reduction in symptoms in the early-adenotonsillectomy group than in the watchful-waiting group. Normalization of polysomnographic findings was observed in a larger proportion of children in the early-adenotonsillectomy group than in the watchful-waiting group (79% vs. 46%).

CONCLUSIONS

As compared with a strategy of watchful waiting, surgical treatment for the obstructive sleep apnea syndrome in school-age children did not significantly improve attention or executive function as measured by neuropsychological testing but did reduce symptoms and improve secondary outcomes of behavior, quality of life, and polysomnographic findings, thus providing evidence of beneficial effects of early adenotonsillectomy. (Fundied by the National Institutes of Health; CHAT ClinicalTrials.gov number, NCT00560859.)
The childhood obstructive sleep apnea syndrome is associated with numerous adverse health outcomes, including cognitive and behavioral deficits. The most commonly identified risk factor for the childhood obstructive sleep apnea syndrome is adenotonsillar hypertrophy. Thus, the primary treatment is adenotonsillectomy, which accounts for more than 500,000 procedures annually in the United States alone. Nevertheless, there has been no controlled study evaluating the benefits and risks of adenotonsillectomy, as compared with watchful waiting, for the management of the obstructive sleep apnea syndrome.

The Childhood Adenotonsillectomy Trial (CHAT) was designed to evaluate the efficacy of early adenotonsillectomy versus watchful waiting with supportive care, with respect to cognitive, behavioral, quality-of-life, and sleep factors at 7 months of follow-up, in children with the obstructive sleep apnea syndrome. Our primary outcome was a neurobehavioral measure of attention and executive function, a domain that has been shown to be sensitive to intermittent hypoxemia related to the obstructive sleep apnea syndrome. Given the prevalence of this syndrome among black children and obese children, we also evaluated whether the relative efficacy of the treatment differed according to race, weight, or baseline severity of the syndrome.

METHODS

STUDY DESIGN AND PATIENTS

We conducted this multicenter, single-blind, randomized, controlled trial at seven academic sleep centers. Methodologic details have been published previously and are provided in the full protocol and in the Supplementary Appendix (available with the full text of this article at NEJM.org).

Eligible children were 5 to 9 years of age, had the obstructive sleep apnea syndrome without prolonged oxyhemoglobin desaturation, and were considered to be suitable candidates for adenotonsillectomy. The obstructive sleep apnea syndrome was defined as an obstructive apnea–hypopnea index (AHI) score of 2 or more events per hour or an obstructive apnea index (OAI) score of 1 or more events per hour. Children with an AHI score of more than 30 events per hour, an OAI score of more than 20 events per hour, or arterial oxyhemoglobin saturation of less than 90% for 2% or more of the total sleep time were not eligible, owing to the severity of the polysomnographic findings. Exclusion criteria included recurrent tonsillitis, a z score based on the body-mass index (the weight in kilograms divided by the square of the height in meters) of 3 or more, and medication for attention deficit–hyperactivity disorder (ADHD).

Children were randomly assigned to early adenotonsillectomy (surgery within 4 weeks after randomization) or a strategy of watchful waiting. At the baseline visit, children with coexisting conditions that could exacerbate the obstructive sleep apnea syndrome (e.g., allergies and poorly controlled asthma) were referred for treatment as needed.

STUDY OVERSIGHT

The study was approved by the institutional review board at each participating site. Written informed consent was obtained from caregivers, and assent from children who were 7 years of age or older. An independent data and safety monitoring board reviewed interim data on safety and study quality. An external medical monitor adjudicated treatment failures, defined as changes in clinical status requiring a change in the assigned therapy. All the authors vouch for the completeness and accuracy of the data and the fidelity of the study to the protocol (available at NEJM.org). There was no commercial support for this study.

ASSESSMENTS

Children underwent standardized polysomnographic testing with scoring at a centralized sleep reading center, cognitive and behavioral testing, and other clinical and laboratory evaluations at baseline and 7 months after randomization. At both examinations, caregivers were asked to complete survey instruments, and teachers were mailed behavioral assessments (see the Supplementary Appendix).

OUTCOMES

The primary study outcome was the change in the attention and executive-function score on the Developmental Neuropsychological Assessment (NEPSY; scores range from 50 to 150, with 100 representing the population mean and higher scores indicating better functioning). This test has well-established psychometric properties and comprises three tasks (tower building, visual
attention, and auditory attention) performed under the supervision of a psychometrist.

Other outcomes included caregiver and teacher ratings of behavior (Conners’ Rating Scale Revised: Long Version Global Index, comprising Restless—Impulsive and Emotional Lability factor sets [caregiver-rated T scores range from 28 to 101, and teacher-rated T scores range from 37 to 131, with higher scores indicating worse functioning]),12 and the Behavior Rating Inventory of Executive Function [BRIEF] Global Executive Composite T score, comprising summary measures of behavioral regulation and metacognition [caregiver-rated scores range from 28 to 101, and teacher-rated scores range from 37 to 131, with higher scores indicating worse functioning]);13; symptoms of the obstructive sleep apnea syndrome, as assessed by means of the Pediatric Sleep Questionnaire sleep-related breathing disorder scale (PSQ-SRBD), in which scores range from 0 to 1, with higher scores indicating greater severity14; sleepiness, as assessed with the use of the Epworth Sleepiness Scale modified for children, in which scores range from 0 to 24, with higher scores indicating greater daytime sleepiness15; global quality of life (caregiver-rated total score from the Pediatric Quality of Life Inventory [PedsQL], in which scores range from 0 to 100, with higher scores indicating better quality of life)16; disease-specific quality of life (total score on the 18-item Obstructive Sleep Apnea-18 assessment tool, in which scores range from 18 to 126, with higher scores indicating worse quality of life)17; generalized intellectual functioning (General Conceptual Ability score from the Differential Ability Scales-II [DAS], in which scores range from 30 to 170, with higher scores indicating better functioning)18; and polysomnographic indexes.

STATISTICAL ANALYSIS

We calculated that with a sample of 400 children, randomly assigned in a 1:1 ratio to early adenotonsillectomy or a strategy of watchful waiting, the study would have 90% power to detect an effect size of 0.32 or more (on the basis of an estimate from one previous study19) for the primary outcome. We planned to enroll 460 children to compensate for withdrawal from the study. Children who crossed over to the other treatment were included in their assigned study groups for the primary analysis, consistent with the intention-to-treat principle.

Primary and secondary outcomes were evaluated with the use of an analysis of covariance with adjustment for the stratification factors of age, race, weight status, and study site. Additional prespecified analyses included adjustments for other factors and restrictions to certain subgroups (see the Supplementary Appendix). Models evaluating possible effect modification of treatment according to race, obesity status, AHI score, and age were tested by including terms for interactions between the two groups and by the effect of each of these factors on each of the study outcomes. A sensitivity analysis was performed with the use of multiple imputation to assess the effect of missing values on the primary outcome.20

STUDY OVERVIEW

Figure 1 shows the enrollment and randomization of the participants. From January 2008 through September 2011, a total of 464 children underwent randomization. Follow-up visits were conducted for 400 children (86%), with 397 children having measurements of attention and executive function on the NEPSY that could be evaluated. A comparison of children who completed the study and those who did not showed a significant difference only with respect to race; black children were less likely to complete the study (P=0.04), but this trend was evident in both study groups.

Baseline characteristics are shown according to study group (Tables 1 and 2, and Table S1 in the Supplementary Appendix). Baseline demographic and clinical characteristics were generally well balanced between the study groups, and cognitive and behavioral scores were close to population means. Nearly half the participants were overweight or obese. A similar number of children in each group used nasal glucocorticoids (19 children in the early-adenotonsillectomy group and 8 in the watchful-waiting group) or montelukast (7 in the early-adenotonsillectomy group and 8 in the watchful-waiting group) for allergic rhinitis or asthma; the data were the same at baseline and at 7 months of follow-up.
STUDY OUTCOMES

The baseline attention and executive function score on the NEPSY was close to the population mean of 100 in both groups. Average scores increased in both groups; the difference between the groups favored early adenotonsillectomy but was not significant ($P = 0.16$). A sensitivity analysis to assess the possible effect of missing data yielded results that were essentially identical to those presented in Table 2.

There was a significantly greater improvement on the caregiver-reported Conners’ Rating Scale among children randomly assigned to early adenotonsillectomy than among those assigned to watchful waiting. Teacher-reported data for this measure, which were available for 212 children, also showed significantly greater improvement in the early-adenotonsillectomy group. The caregiver-reported BRIEF score was lower (indicating an improvement) in the early-adenotonsillectomy group, with a small increase in score in the watchful-waiting group; the teacher-reported version, which was available for 207 children, showed changes that paralleled the caregiver data but were not significantly different between the groups.

Symptoms of the obstructive sleep apnea syndrome were measured with the use of the PSQ-SRBD and the Epworth Sleepiness Scale, and generic and disease-specific measures of quality of life were assessed by means of the PedsQL and OSA-18, respectively. All these instruments showed a significantly greater reduction in symptoms in the early-adenotonsillectomy group than in the watchful-waiting group. The DAS score did not change significantly in either study group (data not shown).

The AHI score improved in both groups but significantly more so in the early-adenotonsillectomy group. Similar results were observed for the oxygen desaturation index (the number of times per hour of sleep that the blood oxygen...
level drops by 3% or more from baseline) and the level of hypercapnia. The early-adenotonsillectomy group had significantly larger decreases in the arousal index and in the percentage of sleep time in stage N1 (light sleep), consistent with improved sleep continuity. However, no changes were observed in stage N3 or in rapid-eye-movement sleep. Normalization of the obstructive sleep apnea syndrome, as defined by a reduction in both the AHI score to fewer than two events per hour and the OAI score to fewer than one event per hour,\textsuperscript{7-10} was more common in the early-adenotonsillectomy group than in the watchful-waiting group (79% vs. 46%, P<0.001).

### EFFECT OF BASELINE CHARACTERISTICS

For all the outcomes, relative improvements with early adenotonsillectomy, compared with watchful waiting, were similar in subgroups defined according to age or obesity status. However, relative improvements with early adenotonsillectomy were significantly smaller among black children than among children of other races on the Conners’ Rating Scale, the BRIEF, and the PSQ-SRBD scale (see the Supplementary Appendix). Among children with more severe obstructive sleep apnea syndrome at baseline (i.e., AHI score above the median), those randomly assigned to early adenotonsillectomy had a greater improvement in AHI score than did those assigned to watchful waiting (Fig. S1 in the Supplementary Appendix). There was no interaction between the severity of the obstructive sleep apnea syndrome and treatment with respect to the attention and executive-function score on the NEPSY or the behavioral or symptom outcomes.

### SUBGROUP ANALYSES

Substantial subgroup differences with regard to the normalization of polysomnographic findings were observed within each study group. Regardless of the assigned treatment, normalization of polysomnographic findings was seen less frequently in black children than in children of other races, in obese children than in nonobese children, and in children with a baseline AHI score above the median than in those with a baseline AHI score at or below the median (Fig. 2).

### OTHER EXPLORATORY ANALYSES

No consistent relationships were observed between changes in AHI score and changes in other study outcomes, nor were substantive differences seen in models adjusted for additional covariates or restricted to children with low cognitive scores at baseline. The exclusion from the analyses of 24 children who did not receive the assigned intervention (8 children in the early-adenotonsillectomy group who declined surgery and 16 in the watchful-waiting group who underwent early surgery) did not yield any appreciable changes in the study results (see the Supplementary Appendix).

### ADVERSE EVENTS

There were 15 serious adverse events after randomization, 6 of which occurred in children randomly assigned to early adenotonsillectomy. There were no other serious adverse events.

---

**Table 1. Baseline Characteristics of Patients Who Completed the Study.**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Watchful Waiting (N = 203)</th>
<th>Early Adenotonsillectomy (N = 194)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age — yr</td>
<td>6.5±1.4</td>
<td>6.5±1.4</td>
</tr>
<tr>
<td>Male sex — no. (%)</td>
<td>106 (52)</td>
<td>89 (46)</td>
</tr>
<tr>
<td>Race — no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>108 (53)</td>
<td>103 (53)</td>
</tr>
<tr>
<td>White</td>
<td>76 (37)</td>
<td>67 (35)</td>
</tr>
<tr>
<td>Other</td>
<td>19 (9)</td>
<td>24 (12)</td>
</tr>
<tr>
<td>Hispanic ethnicity — no. (%)†</td>
<td>17 (8)</td>
<td>15 (8)</td>
</tr>
<tr>
<td>Height — cm</td>
<td>124.7±10.5</td>
<td>125.1±11.2</td>
</tr>
<tr>
<td>Height z score</td>
<td>0.6±1.0</td>
<td>0.7±1.0</td>
</tr>
<tr>
<td>Weight — kg</td>
<td>30.1±11.7</td>
<td>31.2±13.1</td>
</tr>
<tr>
<td>Weight z score</td>
<td>1.0±1.2</td>
<td>1.0±1.3</td>
</tr>
<tr>
<td>Weight class — no. (%)‡</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overweight or obese</td>
<td>94 (46)</td>
<td>93 (48)</td>
</tr>
<tr>
<td>Obese</td>
<td>67 (33)</td>
<td>68 (35)</td>
</tr>
<tr>
<td>Failure to thrive</td>
<td>3 (1)</td>
<td>4 (2)</td>
</tr>
<tr>
<td>Maternal educational level less than high school — no. (%)</td>
<td>64 (32)</td>
<td>62 (32)</td>
</tr>
<tr>
<td>Annual household income &lt; $30,000 — no. (%)</td>
<td>82 (40)</td>
<td>73 (38)</td>
</tr>
</tbody>
</table>

* Plus–minus values are means ±SD. There were no significant differences between the study groups. The characteristics of the sample included in the primary outcome analysis were similar to those of the baseline population, except that black children were less likely to complete the study (P=0.04), but this trend was evident in both study groups (Table S1 in the Supplementary Appendix).

† Race reported by caregivers.

‡ Overweight or obese was defined as a body-mass index (BMI; the weight in kilograms divided by the square of the height in meters) in the 85th percentile or higher, obese as a BMI in the 95th percentile or higher, and failure to thrive as a BMI in less than the 5th percentile.
assigned to early adenotonsillectomy and 9 in those assigned to watchful waiting (Table 3, and Results and Table S3 in the Supplementary Appendix). A total of 8 events were associated with perioperative complications, 3 of which occurred in children who were randomly assigned to watchful waiting but who had crossed over to surgery. Nine treatment failures were identified, all in the watchful-waiting group.

**DISCUSSION**

This large, randomized, controlled trial of therapy for the pediatric obstructive sleep apnea...
The obstructive sleep apnea syndrome included rigorous assessments of cognitive and behavioral measurements, sleep apnea symptoms, and sleep. After a 7-month intervention period, school-age children with the obstructive sleep apnea syndrome without prolonged oxyhemoglobin desaturation who underwent surgery did not have significantly greater improvement in attention and executive function, as measured by means of neuropsychological testing, than did children in the watchful-waiting group. However, surgery resulted in greater reductions in symptoms and greater improvements in behavior, quality of life, and polysomnographic findings, with effect sizes in the moderate-to-large range. Polysomnographic findings were normalized in the majority of children (79%) in the early-adenotonsillectomy group, although polysomnographic abnormalities also resolved in 46% of the children randomly assigned to watchful waiting. However, surgery resulted in greater reductions in symptoms and greater improvement in behavioral and polysomnographic outcomes than did those in the watchful-waiting group.

Potential cognitive and behavioral effects of the obstructive sleep apnea syndrome are of major concern to parents of children with this syndrome. The plausibility of the obstructive sleep apnea syndrome contributing to cognitive deficits is supported by research showing impaired learning in juvenile rats exposed to intermittent hypoxemia and by imaging studies showing cerebral neuronal injury in children with the syndrome. Previous studies have shown differences in the cognitive function of children with the obstructive sleep apnea syndrome, as compared with controls. However, baseline cognitive scores of children with the obstructive sleep apnea syndrome fell within the normal range, a finding similar to that in the current study. Cognitive and behavioral abnormalities have been shown to be reduced after adenotonsillectomy in some, but not all, nonrandomized studies, with inconsistency in the reported effect after treatment. Previous studies have been limited by small samples, lack of randomization or appropriate controls, heterogeneous study groups, and sole reliance on parent questionnaires rather than including neuropsychological testing.

We observed no significant difference between the two groups in the change from base-
line to follow-up in our primary outcome, the attention and executive-function score of the NEPSY; thus our trial is a negative one. However, other tests showed evidence of changes in behavior. There were greater improvements in the BRIEF, which assesses executive function and behavior on the basis of children’s performance in activities of daily living, among children randomly assigned to early adenotonsillectomy than among those assigned to watchful waiting. Since caregivers were aware of the intervention, it is possible that the improvements in the BRIEF score in the early-adenotonsillectomy group were influenced by parental expectations. Alternatively, the children treated with early adenotonsillectomy might have been better able to attend to tasks in less controlled, real-world settings (such as those in which the BRIEF scores were assessed) than in the closely supervised environment of the assessment of the NEPSY attention and executive-function domain, in which children directly interacted with a psychometrist. Results of previous research showing more robust effects on caregiver ratings of behavior than on psychometric tests are consistent with either possibility.33

There were also significantly larger improvements in the scores on the Conners’ Rating Scale, measuring restlessness and impulsiveness and emotional lability, in the early-adenotonsillectomy group than in the watchful-waiting group. The improvements, which were considered to be small to moderate in size, were observed in both the caregiver-reported and teacher-reported scores, suggesting that the positive behavioral change is due not solely to the influence of unblinded assessment or parental expectation.

The reduction in symptoms of the obstructive sleep apnea syndrome and improvement in quality of life in children treated surgically are noteworthy, owing to the importance of symptoms and health-related quality of life to patients and their families. Polysomnographic findings improved markedly after adenotonsillectomy, with more children in the early-adenotonsillectomy group than in the watchful-waiting group having normalizations of polysomnographic findings (79% vs. 46%). Normalization of the polysomnographic findings in the large majority of children randomly assigned to early adenotonsillectomy is consistent with some,34,35 but not all,36,37 studies. Previous nonrandomized studies may have been limited by differential follow-up of more symptomatic children.

Nearly half the children in the watchful-waiting group showed normalization of the AHI score. This improvement may have been due to growth of the airway or regression of lymphoid tissue, routine medical care, or regression to the mean. This study did not show a correlation between the severity of the obstructive sleep apnea syndrome, as measured by means of polysomnography, and neurobehavioral outcomes — a finding similar to that in other studies.15,25 The lack of correlation may be due to the influence of insufficient sleep or other environmental or genetic influences.

Studies on the efficacy of adenotonsillectomy in obese children have shown contradictory results.38 In CHAT, obese children in both groups had lower rates of normalization of polysomnographic findings than did nonobese children, although the prevalence of residual obstructive sleep apnea syndrome in the obese group postoperatively (33% of children) was lower than has

### Table 3. Adverse Events.*

<table>
<thead>
<tr>
<th>Event</th>
<th>Watchful Waiting (N = 203)</th>
<th>Early Adenotonsillectomy (N = 194)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>225</td>
<td>160</td>
</tr>
<tr>
<td>Tonsillar hemorrhage</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Postoperative pain</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Asthma</td>
<td>18</td>
<td>3</td>
</tr>
<tr>
<td>Lower respiratory tract illness</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Upper respiratory tract or ear illness</td>
<td>90</td>
<td>67</td>
</tr>
<tr>
<td>Cough</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>Gastrointestinal tract illness</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Dehydration</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>ADHD</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Other infection</td>
<td>40</td>
<td>17</td>
</tr>
<tr>
<td>Hypersomnolence</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Exacerbation of sleep-apnea symptoms</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>34</td>
<td>27</td>
</tr>
</tbody>
</table>

* Serious adverse events included tonsillar hemorrhage in three children randomly assigned to early adenotonsillectomy and in one child randomly assigned to watchful waiting who crossed over to early adenotonsillectomy; see Table S3 in the Supplementary Appendix for details. ADHD denotes attention deficit–hyperactivity disorder.
been reported in some studies,\textsuperscript{5,38,39} possibly because CHAT was limited to preadolescents and excluded extremely obese patients. Polysomnographic findings, as well as symptoms and behavior, improved more with early adenotonsillectomy than with watchful waiting, however, in both obese and nonobese children. These results support a strategy of early adenotonsillectomy in both obese and nonobese children for the treatment of the physiological disturbances of the obstructive sleep apnea syndrome and associated symptoms, but they underscore the need to carefully follow obese children after surgery.

Black children have been reported to have more severe cases of the obstructive sleep apnea syndrome than white children; we also observed more severe cases of this syndrome among black children at baseline. Black children in both groups had lower rates of normalization of polysomnographic findings than did children of other races but, similar to children of other races, had relatively greater improvement on polysomnographic findings with early adenotonsillectomy than with watchful waiting. However, early adenotonsillectomy, as compared with watchful waiting, was associated with less relative improvement in caregiver-reported measures of behavior and symptoms in black children than in children of other races. Differences persisted in analyses that were adjusted for obesity, baseline behavioral scores, and household income and in analyses that were restricted to children in whom the obstructive sleep apnea syndrome resolved. The reasons for this racial disparity are unclear. Possible explanations include differences in parents’ expectations, coping mechanisms, or perceptions of their child’s behavior and the presence of risk factors for behavioral problems unrelated to the obstructive sleep apnea syndrome.

Children with more severe cases of the obstructive sleep apnea syndrome showed larger absolute improvements in polysomnographic findings with early adenotonsillectomy than with watchful waiting, although, as a group, they were less likely to show normalization of the findings than were children with less severe cases. More severe cases of the obstructive sleep apnea syndrome were not associated with differences between the two groups in cognitive, behavioral, or symptom outcomes.

Surgery was associated with a low rate of perioperative complications; the treatment failure rate was also low but was limited to the watchful-waiting group. Thus, this trial supports the overall safety of both early adenotonsillectomy and watchful waiting but suggests the need for clinical monitoring of children who are being treated conservatively.

The strengths of this study include the large sample, randomized design, standardization of measurements, blinding of key personnel, wide geographic and racial representation, and high follow-up rates. Data from both caregivers and teachers provided independent behavioral assessments.

The study also had limitations. It did not include children younger than 5 years of age, in whom the obstructive sleep apnea syndrome is common. Since children who had prolonged oxyhemoglobin desaturation or who were taking medications for ADHD were excluded, the study results cannot be extrapolated to these vulnerable groups. It is possible that the follow-up period was not long enough to show the full response to surgery. There are insufficient longitudinal data to determine when recovery is maximal. It is possible that neurobehavioral sequelae related to hypoxemia may result from neuronal damage and take longer to resolve.

A total of 200 interaction tests were conducted, yielding 23 significant interactions; 10 significant interactions would be expected by chance. Therefore, the results of exploratory analyses must be viewed cautiously.

Among school-age children with the obstructive sleep apnea syndrome without prolonged oxyhemoglobin desaturation, early adenotonsillectomy, as compared with a strategy of watchful waiting with supportive care, did not result in significantly greater improvement in our prespecified primary outcome, scores on a formal test of attention and executive function after a period of 7 months. However, early adenotonsillectomy was associated with significant improvements in several other prespecified secondary outcomes, many with effect sizes considered to be moderate to large\textsuperscript{40} and thus likely to be clinically significant, including polysomnographic findings; caregiver-reported measures of executive function, behavior, and sleep apnea symptoms; and teacher reports of behavior. Beneficial effects of early adenotonsillectomy were observed...
in nonobese children as well as in obese children. Normalization of polysomnographic findings in a large number of children in the watchful-waiting group and an absence of significant cognitive decline in this group indicate that medical management and reassessment after a period of observation may be a valid therapeutic option.

Supported by grants (HL083075, HL083129, UL1 RR024134, and UL1 RR024989) from the National Institutes of Health.

Dr. Marcus reports receiving a loan of research equipment from Philips Respironics and Ventus Medical. Dr. Gozal reports consulting fees from Arena Pharmaceuticals, Guideline Global, OrbiMed Advisors, Procter & Gamble, and Zansors; having patents and patent pending, owned by his institution, for analysis algorithms relevant to diagnosis of sleep disorders and having a pending, owned by his institution, for a novel device to treat obstructive sleep apnea (none of these patents are currently licensed); and receiving gifts to his institution for educational purposes from Philips Respironics and Fisher and Paykel. No other potential conflict of interest relevant to this article was reported.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank the Childhood Adenotonsillectomy Trial (CHAT) research staff, including Jean Arnold, Mary Ellen Carroll, Mary Anne Cornaglia, Beth Ann Compton, Casey Critchlow, Judith Emanuelovich, Melissa Fernando, Theresa Friederich, Amanda Goodman, Xiaoling Hou, Elise Hodges, Laurie Karamessinos, Kim Lacy, Megan McDougall, Daniel Mobley, Michelle Nicholson, Angela Orlando, Deborah L. Ruzicka, Gauri Sathe, Nancy Scott, Susan Szovec, Omarya Vega, Xingmei Wang, and Catherine Williams; the families of the children enrolled in the study; and the members of the data and safety monitoring board, including Lynn Taussig, M.D., (chair), Thomas Anders, M.D., Julie Buring, Sc.D., Karina Davidson, Ph.D., Estelle Gauda, M.D., Steven Piantadosi, M.D., Ph.D., Bennett Shaywitz, M.D., Benjamin Wilfond, M.D., Tucker Woodson, M.D., and Robert Zeiger, M.D.

APPENDIX

The authors’ affiliations are as follows: the Department of Pediatrics, Sleep Center (C.L.M.), Department of Child and Adolescent Psychiatry and Behavioral Sciences (N.H.T.), and Department of Otolaryngology (L.E.), Children’s Hospital of Philadelphia, and the Department of Biostatistics and Epidemiology, School of Medicine (R.H.M., S.S.E.), University of Pennsylvania, Philadelphia; the Department of Statistics, North Carolina State University, Raleigh (R.H.M.); the departments of Pediatrics (C.L.R., H.G.T.) and Otolaryngology (R.C.S.), Rainbow Babies and Children’s Hospital, University Hospitals Case Medical Center, Case Western Reserve University, Cleveland; the Departments of Psychiatry and Psychology (B.G.), Pediatrics and Neurology (B.G., T.H., R.D.C.), and Otolaryngology (S.L.G.), University of Michigan, Ann Arbor; the Departments of Otolaryngology (R.B.M.) and Pediatrics (S.P., K.S.), Cardinal Glennon Children’s Medical Center, St. Louis University, St. Louis; the Departments of Otolaryngology and Pediatrics, University of Texas Southwestern Medical Center, Dallas (R.B.M.); the Departments of Pediatrics (R. Amin, D.B.) and Otolaryngology (P.W.), Cincinnati Children’s Hospital Medical Center, Cincinnati; the Departments of Pediatrics (E.S.K.), Medicine and Psychiatry (J.W.), and Otorhinolaryngology (D.J.), Boston Children’s Hospital, and the Department of Medicine, Brigham and Women’s Hospital and Beth Israel Deaconess Medical Center, Harvard Medical School, Boston (S.R.); the Departments of Pediatrics (R. Arens, H.M.) and Otolaryngology—Head and Neck Surgery (J.P.B.), Children’s Hospital at Montefiore and Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, NY; and the Department of Pediatrics, University of Chicago, Chicago (D.G.).

REFERENCES

16. Varni JW, Seid M, Kurtin PS. PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. Med Care 2001;39:800-12.
17. Franco RA Jr, Rosenfeld RM, Rao M.


