Randomised trial showed no difference in behavioural symptoms between surgical methods treating paediatric obstructive sleep apnoea

Isabella Sjölander¹, Anna Borgström², Jan-Olov Larsson³, Hans Smedje³, Danielle Friberg¹

Affiliations: 1. Department of Surgical Sciences, Otorhinolaryngology-Head and Neck Surgery, Uppsala University, Uppsala, Sweden 2. Department of Clinical Science, Intervention and Technology, CLINTEC, Karolinska Institute, Stockholm, Sweden. 3. Department of Women’s and Children’s Health, Karolinska Institutet, Stockholm, Sweden

Corresponding author: Isabella Sjölander, Department of Oto-rhino-laryngology Lasarettsvägen 1, -37- 801 87 Gävle Sjukhus, E-mail: isabella.sjolander@surgsci.uu.se
ABSTRACT

Aim
Our previous randomised controlled trial of children with obstructive sleep apnea (OSA) showed no significant differences between adenotonsillectomy (ATE) and adenotonsillotomy (ATT) in improving nocturnal respiration and quality of life after one year. The aim of this report was to evaluate the effects on behavioural symptoms using the Strengths and Difficulties Questionnaire (SDQ).

Methods
Children between 2 and 6 years with OSA were randomised to ATT or ATE. Parents, blinded to method, answered the SDQ while their child underwent polysomnography before and one year after surgery. Differences between the total SDQ-scores were analysed between the treatment groups.

Results
The SDQ was filled out in 87% of the cases preoperatively, and in 86% postoperatively. At follow up, the mean total SDQ score was 9.6 SD ±5.1 in the ATE group (n=31), and 8.2 ±6.7 in the ATT group (n=37), p = 0.09. The mean total SDQ score for all was preoperatively 10.6 ±5.0, and postoperatively 8.8 ±6.0, p = 0.0002.

Conclusion
There were no significant differences in SDQ-scores between the groups at follow up, indicating that the more conservative ATT is a treatment option in paediatric OSA. The whole group of patients showed a significant improvement after surgery.
Key notes:

- Behavioural symptoms are common among children with obstructive sleep apnoea. This randomised controlled trial is a continuous report of our previous study comparing adenotonsillectomy and adenotonsillotomy as surgical methods for treating paediatric obstructive sleep apnoea.
- This study showed no significant differences in behavioural symptoms evaluated with Strengths and difficulties questionnaire one year after surgery.
- The findings support the choice of adenotonsillotomy, which is the more conservative surgical treatment option.

Keywords: adenotonsillectomy, adenotonsillotomy, paediatric obstructive sleep apnoea, strengths and difficulties questionnaire.
INTRODUCTION

Paediatric obstructive sleep apnoea (OSA) is a serious and common form of sleep disordered breathing in children. OSA is characterised by repeated events of upper airway obstruction during sleep causing decreased airflow and hypoxemia despite continual respiratory effort (1). Hypertrophy of adenoid and tonsils is the most common cause of OSA, which affects around 1% to 5% of all children between two to six years of age (2). Left untreated, OSA is associated with serious complications, such as failure to thrive, hyperactivity, cognitive disturbances, lower quality of life and cardiovascular disorders (3-5). Already in 1976, the first published series of children with OSA were described with learning difficulties, hyperactivity and attention deficits. Studies in the early 1980’s, showed that children with attention deficit hyperactivity disorder who were treated for their sleep-disordered breathing experienced less symptoms related to this (6). Newer studies support these findings (3, 7), as well as suggest that OSA may affect the cortical thickness and grey matter in the brain (8, 9).

Adenotonsillectomy (ATE) is considered the first-line treatment for OSA and is one of the most common surgical procedures among children around the world (10). Studies have shown significant improvement in behavioural problems among children with OSA treated with ATE (11, 12). Tonsillotomy is a more conservative treatment with only a reduction of the protruding mass of the tonsils, instead of total extracapsular excision as in standard tonsillectomy. Adenotonsillectomy (ATT) is accompanied with less bleeding and postoperative pain, as well as decreased costs for society (13). ATT was shown to be equally effective as ATE in treating 79 children with OSA in our prospective, blinded randomized controlled trial (RCT) from 2017, measured with polysomnography and OSA-18 questionnaire before and one year after surgery. The results also showed small differences in morbidity after surgery (14).

In this study, at the same day the children underwent polysomnography, the Strengths and Difficulties Questionnaire (SDQ) was filled in by the parents to assess the behavioural and emotional symptoms of their children. The aim of this report was to evaluate the behavioural difficulties among children with OSA treated with ATT vs. ATE using SDQ.
METHODS

Study population
Children between 2 and 6 years of age with OSA were recruited between November 2011 and April 2015 at the Oto-Rhino-Laryngology Department at Karolinska University Hospital in Stockholm, Sweden. Parents and care providers were blinded to intervention method. For more information regarding study design and the randomisation procedure see Borgström et al. (14). The original inclusion criteria were apnoea hypopnea index (AHI) of ≥5 and ≤30 events per hour from polysomnography (PSG) and tonsil hypertrophy 3 or 4, according to the Brodsky scale (15). Exclusion criteria were craniofacial abnormality, neuromuscular disease, chromosomal abnormality, obesity (body mass index z score >1.67), previous adenotonsil surgery, bleeding disorder, cardiopulmonary disease, history of recurrent tonsillitis or parents with insufficient knowledge of the Swedish language. The parent who accompanied the patient at the sleep lab was asked to fill out the Strengths and Difficulties Questionnaire (SDQ) before surgery and at 1-year follow up. For this study an inclusion criterium was added; only those children with a complete first page of the SDQ before and/or after surgery (Figure 1). The randomised trial was approved by the Swedish Regional Ethics Board in Stockholm, Sweden (Dnr 2011/925-32 and 2013/2274-32) and registered at ClinicalTrials.gov (Trial registration number NCT01676181).

Strengths and Difficulties Questionnaire
The SDQ is a widely used and validated questionnaire designed to screen for behavioural difficulties and strengths among children and adolescents (16, 17). It is a user-friendly questionnaire designed for use in different age groups with versions for different informants including parents, teachers, and self-reports (16, 17). The SDQ was translated into Swedish in 1998 (18), the language used in this study. The questionnaire consists of two pages, the first page includes 25 questions divided into 5 subscales; emotional symptoms, conduct problems, hyperactivity/inattention, peer-relationship problems, and prosocial behaviour with a 3-point rating scale (zero for not true, one for somewhat true, two for certainly true)(19). The second page asks about the impact of symptoms on the child’s everyday function, or about effects of a treatment. The questionnaire used in this study was the version for ages 4 to 17 years. A version of the SDQ adapted for ages 2 to 4 years has been made available after the start of our study, this new version has minor modifications in two questions (17).
In addition to the SDQ, the quality of life questionnaire OSA-18 was filled out before and after surgery (14). This questionnaire consists of 18 questions grouped into five subscales (sleep disturbance, physical symptoms, emotional distress, daytime function, and caregiver concerns). The 18 survey items are scored with a 7-point ordinal scale, where the caregiver is asked to report how often during the previous four weeks their child has had specific symptoms. The total symptom score (TSS) may vary from 18 to 126 points.

Statistics
All statistical analyses were made with the algorithm presented on the SDQ website (17), using Stata 15 (20). The total SDQ-score is obtained by summing up the scores from all the subscales except prosocial. The score ranges from 0 to 40. Replacement of missing data was undertaken using the guidelines from the SDQ developers, where, if at least three of the five SDQ items in a subscale were completed, the remaining two scores were replaced by their mean. When more than three items were missing in a scale level, scores were excluded from the analysis. SDQ-scores are presented with median (range) in text, and mean with standard deviation in tables, and in the text when needed.

SDQ subscales, SDQ total scores, Apnoea-Hypopnoea index (AHI) and OSA-18 total symptom scores were compared within groups before and after surgery with Wilcoxon sign rank test. Similar comparisons between treatment groups after surgery were made with Mann-Whitney U test. Intention-to-treat analyses were performed using “baseline carried forward”. Spearman correlation test was performed to correlate Apnoea Hypopnoea Index (AHI) with total SDQ score after surgery, as well as total score of OSA-18 with total SDQ score after surgery. A multivariable regression model was used to evaluate the impact of gender. P value less than 0.05 were regarded as statistically significant.

RESULTS
A total of 79 patients between 2 to 6 years (mean age 3.6 years) were originally included in the trial. Out of the 79 children, 40 were randomised to ATE and 39 to ATT (Figure 1). The two groups had similar baseline characteristics (Table 1). In the present study, 10 patients were excluded preoperatively due to incomplete baseline SDQ. Out of the 69 children, 37 SDQ were obtained from the ATE group and 32 from the ATT group. At follow up, 68 children (86%) were included, whereof 31 children had undergone ATE and 37 ATT. The second page of the SDQ was filled out by 61 (77%) parents postoperatively. Mean time from baseline to follow up was
15.1 months (range 9-19 months) in the ATE group, and 14.3 months (range 9-21) in the ATT group.

The median total SDQ score was 11.0 (range 4-22) at baseline in the ATE group and 9.5 (1-22), in the ATT group (p=0.52) as shown in table 2. At follow up, the median total score in the ATE group was 9.0 (2-29), and in the ATT group 7.0 (0-35), (p=0.09). Further, the SDQ subscales values were compared between the groups and showed no significant differences at baseline or at follow up, (Table 2). Intention to treat analyses did not change the results. There were no significant correlations between AHI and total SDQ score after surgery, $r=-0.15$ p=0.22. A significant correlation was found between OSA-18 and SDQ score after surgery, $r=0.55$ p=0.00. The results for AHI and OSA-18 total symptom scores are presented in table 2.

The median preoperative and postoperative total SDQ-score for both groups were 10.0 (range 1-22), and 8.0 (range 0-35), respectively (p=0.0002). The distribution of boys in the ATE group and ATT group, was 70% and 66%, respectively. Among girls there was a modest decrease in total SDQ-score pre- and post-surgery (mean 9.9 to 9.6, p=0.26). In contrast, there was a significant difference in total SDQ-score among boys (mean 11.0 to 8.4 (p=0.0001). There was no significant difference between pre- and post-surgical SDQ scores between boys and girls (p=0.17).

**DISCUSSION**

To the best of our knowledge this was the first study to compare SDQ-score among children with OSA treated with either ATE or ATT. Our results showed no significant differences between the ATE and ATT groups. This is of importance since ATT is a more conservative treatment option, with lower risks and less pain for the child as well as lower costs, compared to ATE (13). Our previous randomised trial of the same cohort found that ATT was non-inferior to ATE in treating OSA in terms of polysomnographic parameters (14).

In a study from 2009 by Ericsson et al., 67 children between 4.5 to 5.5 years old with OSA were treated with either ATE or ATT. Their behaviour was assessed using the Child Behaviour Checklist before and after surgery (21). Their study, similar to this study, showed no significant differences between the groups. In a recent prospective cohort study by Kim et al., the results of the SDQ scores in 170 children with sleep-disordered breathing treated with ATE, were compared to the SDQ score results from 150 non-snoring controls(7). At follow up after 15
months there was a reduction in the mean SDQ score from 10.8 to 8.1 in the ATE group, and from 8.0 to 7.8 in the control group. Younger children improved more than older. Thus, their findings of improvements in SDQ after surgery are similar to ours.

In another study by Hill et al, an association was found between the degree of OSA and scores on the SDQ(22). Normal controls had a mean total SDQ score of 6.44, compared to 13.74 in habitual snorers. When snorers were investigated with polysomnography, the authors could significantly correlate children with AHI>5 with higher mean total SDQ score (16.54 ± 4.89), a difference not found significant when OSA was defined as AHI>2. In the present study all children had OSA with an AHI>5, but there was no significant correlation with the total SDQ score and AHI. It is a known issue for clinicians to not be able to find correlations between objective and subjective findings among children with OSA (23, 24). However, in the present study we found a significant correlation between OSA-18 and SDQ total score. This could be explained by some similar characteristics in the two questionnaires such as emotional symptoms, and that the same parent filled out both at the same time and probably rated the problems at the same level.

An important study of paediatric OSA is the “Childhood Adenotonsillectomy Trial” (CHAT) of 400 children (mean age 6.5, 5 to 9 years) in which attention and executive-function score by Developmental Neuropsychological Assessment was primary outcome (25). The children were randomised to ATE or watchful waiting and no significant group difference in changes of neuropsychological assessment after 7 months was found. On the other hand, secondary outcomes showed significant group differences, as ATE resulted in greater improvements in behavior and quality of life. The children in the CHAT-study were slightly older than in our study (with mean age 3.6, 2 to <6 years) and they did not use SDQ, which makes the results somehow difficult to compare. But in our RCT, there were also similar significant improvements in quality of life (OSA-18) in both ATE and ATT groups (14).

In the present study, the median total score of SDQ for the whole cohort before surgery was 10.6. At present there is no absolute cut off for normal or abnormal scores in the SDQ. In a Swedish study on small children (age 1-3), the normal sample had a median of 6.0 and the psychiatrically “identified children” had a median score of 11.0 (26). On the SDQ-website, data regarding SDQ mean score is available from different countries. The SDQ score for Danish children between five to seven years, was 5.45 ±4.17 for girls and 6.42 ± 4.79 for boys, and 6.8 ±4.7 and 7.8 ±5.2, for English girls and boys, respectively (27). By comparison, our preoperative results on girls were mean 9.9 ±4.6 and boys 11.0 ±5.2. Corresponding postoperative total scores were 9.6 ±5.3 for
girls and 8.4 ±6.4 for boys. Our results suggest that several children in our sample had significant behavioural difficulties preoperatively, and some might have had so also postoperatively. Whether OSA was the cause of the high scores cannot be determined from our results, but a significant improvement after surgery indicates that. Another possibility of course, could be that the improvement partly was due to natural maturation of these young children. Thus, changes between baseline and the follow up values are difficult to interpret, and we chose to compare the follow up values between the two treatment groups.

The main strength of this study was the prospective randomised controlled design and that parents were blinded to intervention method which minimises problems with selection bias and confounders. It was also a low dropout rate (14%). Taken together, these strengths add to the reliability of the results. Important to notice is also the fact that there are few other studies of these young children (<6 years) with paediatric OSA, since this age group has a prevalence peak of OSA which can impact severely on the young developing children.

There were several limitations of this study. Firstly, it was a secondary outcome of our previous RCT, which means no power analysis of SDQ-score was made. Thus, our study sample (n=68) may have been too small to show true group differences, especially when comparing gender. However, the two treatment groups of ATE and ATT were of similar size and characteristics, but only included healthy, non-obese children which limits the generalisability. There are several known risk factors for OSA, e.g., obesity and neuromuscular diseases, including these children might have affected the results.

Another limitation concerns the fact that psychological issues among children are difficult to evaluate and ideally you should use several tools. We decided to use the well-known paediatric questionnaires OSA-18 and SDQ, both validated and translated into Swedish. SDQ is an established screening tool to detect behavioural problems in this young age and is considered a good shorter version of the longer Child Behaviour Checklist. However, SDQ is a subjective tool and the parents in the present study were only blinded for surgical method. Further, there was no untreated placebo-group since it would have been unethical in this study of children with severe OSA. Therefore, it is possible that improvements in SDQ after surgery could be explained by expectations from the parents, a surgical placebo-effect. To use teachers’ ratings were not applicable among these pre-school children. Also, as was stated in the methods section we used a questionnaire designed for children 4 to 17 years of age since the one for 2 to 4 years wasn’t accessible at the study start. We do however believe that this had only minor impact since the wording of the two versions are similar.
CONCLUSION
This study investigated behavioural problems using strengths and difficulties questionnaire (SDQ) among young children with obstructive sleep apnoea (OSA). They were treated with either adenotonsillotomy (ATT) or adenotonsillectomy (ATE) and screened before and one year after surgery. We found no significant differences in SDQ total score between the groups after surgery. The whole group of patients showed a significant improvement in SDQ after surgery. Our results should be interpreted with caution as there was no untreated control group, however, they indicate that the more conservative treatment with ATT is a treatment option in non-obese healthy children with OSA.

ABBREVIATIONS
AHI: Apnoea Hypopnea Index
ATE: Adenotonsillectomy
ATT: Adenotonsillotomy
OSA: Obstructive sleep apnoea
PSG: Polysomnography
SD: Standard Deviation
SDQ: Strengths and difficulties questionnaire

ACKNOWLEDGEMENTS
We would like to thank the staff at the Sleep Laboratory at Karolinska University Hospital for their valuable assistance in the study progress and statistician Johan Bring for all your valuable help in interpreting data.

FINANCE
This study did not receive any specific funding.

CONFLICTS OF INTEREST
The authors have no conflicts of interest to declare.
References


20. StataCorp. *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LP; 2017


27. youthinmind. SDQ Normative data. http://sdqinfo.org/g0.html; 2013
Table 1. Baseline characteristics

<table>
<thead>
<tr>
<th></th>
<th>ATE</th>
<th>ATT</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>37</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td><strong>Age months (range)</strong></td>
<td>44 (29-74)</td>
<td>42 (25-80)</td>
<td>0.38</td>
</tr>
<tr>
<td><strong>Female sex, n (%)</strong></td>
<td>11.0 (30)</td>
<td>11.0 (34)</td>
<td>0.30</td>
</tr>
<tr>
<td><strong>Length, cm</strong></td>
<td>98 ± 14</td>
<td>98 ± 10</td>
<td>0.99</td>
</tr>
<tr>
<td><strong>Weight, kg</strong></td>
<td>15.6 ± 3.0</td>
<td>15.2 ± 3.4</td>
<td>0.38</td>
</tr>
<tr>
<td><strong>Tonsil size 1-4</strong></td>
<td>3.3 ± 0.6</td>
<td>3.5 ± 0.6</td>
<td>0.13</td>
</tr>
<tr>
<td><strong>Adenoid size 1-4</strong></td>
<td>2.7 ± 0.8</td>
<td>3.0 ± 0.7</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Values are mean (± SD), except for sex (n(%)) p-value calculated with Mann Whitney U
Table 2. The SDQ subscales and total score, AHI and OSA-18 total symptom score, in mean for each group and in total at baseline and at follow up, the group differences between baseline and follow up values, respectively.

<table>
<thead>
<tr>
<th>SDQ subscales/ AHI &amp; OSA-18</th>
<th>Baseline before surgery</th>
<th>Follow up after one year</th>
<th>Total post</th>
<th>p** Total pre vs. post</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n ATE n=37 ATT n=32 p* ATE vs ATT</td>
<td>n ATE n=31 ATT n=37 p* ATE vs ATT</td>
<td>n Total pre</td>
<td>n Total post</td>
</tr>
<tr>
<td>Emotion</td>
<td>70 2.2 ±2.0 1.9 ±1.6 0.49</td>
<td>69 2.0 ±2.0 1.4 ±1.9 0.08</td>
<td>69 2.0 ±2.0 1.0 ±1.8 0.36</td>
<td></td>
</tr>
<tr>
<td>Conduct</td>
<td>70 3.0 ±1.9 2.8 ±1.8 0.57</td>
<td>68 2.5 ±1.8 2.4 ±2.4 0.42</td>
<td>68 2.5 ±1.8 1.0 ±1.8 0.36</td>
<td></td>
</tr>
<tr>
<td>Hyper</td>
<td>69 4.3 ±2.4 4.0 ±2.3 0.58</td>
<td>69 3.9 ±2.2 3.3 ±2.3 0.25</td>
<td>69 3.9 ±2.2 1.0 ±1.8 0.36</td>
<td></td>
</tr>
<tr>
<td>Peer</td>
<td>70 1.5 ±1.5 1.5 ±1.5 0.90</td>
<td>69 1.1 ±1.3 1.1 ±1.8 0.53</td>
<td>69 1.1 ±1.3 0.5 ±1.8 0.36</td>
<td></td>
</tr>
<tr>
<td>Prosocial</td>
<td>70 7.7 ±1.6 7.5 ±2.1 0.86</td>
<td>68 7.9 ±1.9 7.9 ±2.3 0.65</td>
<td>68 7.9 ±1.9 0.5 ±1.8 0.36</td>
<td></td>
</tr>
<tr>
<td>Impact</td>
<td>32 0.4 ±0.9 0.8 ±1.5 0.79</td>
<td>61 0.2 ±0.7 0.5 ±1.8 0.36</td>
<td>61 0.2 ±0.7 0.0 ±1.8 0.36</td>
<td></td>
</tr>
<tr>
<td>Total SDQ</td>
<td>69 11.0 ±4.8 10.2 ±5.3 0.52</td>
<td>68 9.6 ±5.1 8.2 ±6.7 0.09</td>
<td>68 9.6 ±5.1 0.5 ±1.8 0.36</td>
<td></td>
</tr>
<tr>
<td>AHI</td>
<td>78 14.5 ±7.3 15.2 ±7.3 0.65</td>
<td>72 2.5 ±2.0 4.5 ±6.3 0.11</td>
<td>72 2.5 ±2.0 0.5 ±1.8 0.36</td>
<td></td>
</tr>
<tr>
<td>OSA-18</td>
<td>78 60.9 ±18 65.0 ±19 0.30</td>
<td>69 31.8 ±11.7 36.5 ±11.6 0.11</td>
<td>69 31.8 ±11.7 0.5 ±1.8 0.36</td>
<td></td>
</tr>
<tr>
<td></td>
<td>62.9 ±18</td>
<td>34.1 ±11.8 0.0000</td>
<td>62.9 ±18</td>
<td>34.1 ±11.8 0.0000</td>
</tr>
</tbody>
</table>

_AHI = Apnoea Hypopnea Index. Data are mean (mean ± SD)). p* = calculated with Mann Whitney U for tests between groups. p** Wilcoxon sign rank baseline vs follow up._
Assessed for eligibility (n=176)

97 Excluded
- 68 Not meeting inclusion criteria
- 18 Declined to participate

Randomised (n=79)

Allocated to ATE (n=40)

3 caregivers did not complete the preop SDQ-questionnaire

Analysed (n=31)
9 Caregivers did not complete the follow-up

Allocated to ATT (n=39)

7 caregivers did not complete the preop SDQ-questionnaire

Analysed (n=37)
2 Caregivers did not complete the follow-up

Figure 1. Flow diagram of participants
Figure 2. Box plot SDQ total score before surgery and at follow-up for each group

SDQ total score for each group
before surgery and at follow-up

ATE

ATT

n = 37

n = 31

n = 32

n = 37

SDQ score

before

follow-up