Are sleep studies worth doing?

Vivienne van Someren, Margarita Burmester, Gus Alusi, Rod Lane

Abstract

Aims—To evaluate a sleep study service for children suspected of having sleep related upper airway obstruction (SRUAO).

Design—Prospective survey.

Setting—Paediatric and ear, nose, and throat clinics of the Royal Free Hampstead NHS Trust.

Subjects—Consecutively referred children with SRUAO symptoms.

Main outcome measures—Sleep study data, referring clinician's impression, and completed symptom questionnaires.

Results—A total of 120 children (aged 6 months to 15.5 years) were studied. Study scores showed that 24 were classified as normal, 42 as mild, 33 as moderate, and 21 as severe SRUAO. In the 106 cases with matching data between clinician's impression and study score, 71 had good agreement, 18 were underestimated by the clinician, and 17 were over estimated. No cases reported as moderate or severe sleep apnoea by the study were referred by the clinician as normal. There were no important associations between parental symptom scores and sleep study scores.

Conclusion—In children with suspected SRUAO, sleep studies do contribute to assessing the need for operation, the likelihood of postoperative respiratory failure, or as a baseline or outcome measure in intervention studies.

(Arch Dis Child 2000;83:76–81)

Keywords: sleep related upper airway obstruction; sleep study; snoring; Visilab video system

children, who do not tolerate many sensors, especially on the face. There have therefore been studies attempting to predict sleep study findings from clinical data.⁷ The results show that one can identify from symptoms and structured questionnaires a group of children with a higher risk of sleep study abnormalities. However, within this group only a minority actually have significant SRUAO. Sleep studies therefore seem to be necessary for further evaluation of the high risk group.

We therefore set up a sleep study service for children suspected by clinicians of having SRUAO. We have evaluated the service from the following points of view: (1) the validity and reliability of the chosen system; (2) an assessment of which variables from the sleep study are most useful; (3) to determine whether clinical symptoms can predict sleep study results; and (4) to contribute to the development of guidelines for indication of operation.

Before setting up the service we carried out a survey of the prevalence of symptoms of upper airway obstruction in the ENT clinic population, as well as validation work on the Visilab video system.

Prestudy clinic survey

Information was collected by questionnaire on new referrals to the Royal Free and the Royal National Throat Nose and Ear Hospitals for an ear, nose, and throat opinion. Parents answered questions about the prevalence of obstructive symptoms. The questions used were those that Stradling *et al* have found to be useful predictors of SRUAO.^{3 4} Clinicians answered questions about treatments chosen and the reasons for that choice. Simple descriptive statistics were used to summarise the data.

Data were collected from 61 children. Age ranged from 0.2 to 14.9 years, with a mean of 5.4 years. Following the consultation 17 children were scheduled for operation and 44 for medical treatment or review. The most common indication for adenotonsillectomy was recurrent infections ($n = 15$), with only three having sleep apnoea as the main reason. Parental pressure to operate was never cited. Clinicians thought that a sleep study might help decision making in eight.

Parents reported a high incidence of SRUAO symptoms. A total of 77% of the children snored and 47% had difficulty getting to or staying asleep. There were also high response rates for daytime sleepiness (31%), hyperactivity (37%), and night time restlessness (64%), all good indicators of SRUAO in previous studies.^{3 4 8}

Thus the clinic survey showed a high prevalence of symptoms of upper airway obstruction in the ENT population.

Department of
Paediatrics and
Neonatology, Royal
Free Hampstead NHS
Trust, Pond Street,
London NW3 2QG, UK
V van Someren

Department of Ear
Nose and Throat
Surgery, Royal Free
Hampstead NHS Trust
G Alusi

Department of
Paediatric Intensive
Care, Great Ormond
Street Hospital for
Children NHS Trust,
Great Ormond Street,
London WC1N 3JH,
UK
M Burmester

Department of
Respiratory
Paediatrics, Great
Ormond Street
Hospital
R Lane

Correspondence to:
Dr van Someren

Accepted 21 March 2000
email: Vvs@btinternet.com

There is increasing awareness that enlarged tonsils and adenoids can lead to sleep related upper airway obstruction (SRUAO) in children. In severe cases, a sequence of snoring, increasing respiratory efforts, and reducing airflow, leading to hypoxaemia, arousal, or both can be seen occurring repeatedly throughout the night. This can lead to physical and developmental morbidity.¹ Initially this phenomenon was described in case series from tertiary sleep centres.² Further studies have shown that some degree of sleep disturbance and hypoxaemia occurs in 26–65% of children attending UK ear, nose, and throat (ENT) clinics.³ Two population studies estimate the prevalence of the problem as approximately 1–2% of children in northern Caucasian populations.^{4 5}

Appreciation that SRUAO is common in children has led to interest in methods of assessment.⁶ Because the airway obstruction is mainly apparent at night, some form of sleep study is required. Traditional polysomnographic sleep studies are difficult to carry out in

Validation of Visilab video system

We considered the various available sleep study systems. We opted for the Visilab system which has a proven track record in adult sleep apnoea and has been used in children.³ A video recording is made along with pulse oximetry. Computer software produces measures of movement (from the video) and snoring (from the microphone recording) and analyses the pulse oximeter output. It has the advantages for paediatric use that only the pulse oximeter sensor is in contact with the child. However, the analysis software was developed for adult snoring and sleep apnoea.

METHODS

Validation work was assessed in three ways.

Firstly, the Visilab system was compared with conventional polysomnography. Ten children with pronounced sleep abnormality aged 0.2 to 6.4 years being assessed at Great Ormond Street Children's Hospital for suspected upper airway obstruction ($n = 6$) or central apnoea ($n = 4$) had simultaneous assessments using Visilab and a more conventional polysomnographic system (Oxcams). Signals recorded by Oxcams were pulse oximetry, ECG, nasal airflow (thermistors), chest and abdominal movement (impedance), and video. Data were stored on computer and standard analyses were carried out by the software. Final diagnosis was provided by the clinician reviewing the record and the video. Data selected for comparison were mean SaO_2 , mean heart rate, and final diagnosis.

Secondly, we analysed possible artefactual interference in the Visilab system. Visilab was originally designed for use in adults in a sleep laboratory. We used it in children on a side room of a children's ward at the Royal Free Hospital. This setting could result in a large amount of artefact, making the calculations of, say, heart rate invalid. Detailed analysis of the data on 15 children was therefore carried out. First, the Visilab software was run in the usual way and mean SaO_2 and movement scores were obtained. Then the video was examined carefully. All artefacts caused by prolonged wakefulness and movements by parents and nurses were noted and these periods were removed from the data. The differences between the first and second results were then examined using Bland Altman plots.

Thirdly, we tested the interobserver reliability for overall diagnosis and grading of study score. The observers reviewed the printouts of mean oxygen saturation, oxygen saturation dips per hour, pulse rate rises per hour, and movements per hour, together with the video recording and continuous display of heart rate, oxygen saturation, and noise levels. A final grading was given according to the criteria in

table 1. Seventeen studies of children assessed at the Royal Free Hospital were independently assessed by MB and VVS and the results compared using the κ statistic.

RESULTS

Comparison with polysomnography

The Visilab mean SaO_2 was 93%, compared to 95% for Oxcams. Mean pulse rates were the same (116). For diagnosis, there were only two discrepancies. One child had mild obstruction detected by the Oxcams system, but the Visilab record was judged normal. One had mixed apnoea on Oxcams, but only obstruction was detected by Visilab. Thus both systems performed similarly in the clinical setting.

Movement artefact

The average SaO_2 level was 96.1. The mean difference in SaO_2 between the unedited and edited studies was 1.13 (95% confidence interval (CI) 0.65 to 1.60) with the edited study having the higher value. This is to be expected as artefact depresses SaO_2 measurements. However, this difference was small and not of practical significance.

The average number of movements per study was 84. The mean difference in movement between the unedited and edited studies was 84 (95% CI 60 to 108). The Bland Altman plot showed that the difference was greater at higher movement scores. These large differences arise because the software cannot distinguish between spontaneous patient movement while asleep, prolonged wakefulness, and movement of parents or nurses. Thus the usefulness of the movement score produced by the software will be limited.

Interobserver reliability

In the 17 studies scored by both MB and VVS, there was perfect agreement on grading in 12. In the other five there was a difference of one grade. The κ value was 0.54, indicating good agreement.

DISCUSSION

Our validation tests show that the Visilab video system is robust in the paediatric clinical setting. The final results are similar to those from the more complex Oxcams system and interobserver reliability is good.

Thus, we concluded that the Visilab system was sufficiently valid and reliable to be used for assessment of children, provided that the problem of movement artefact was considered by the reporter.

Sleep study service

SUBJECTS AND METHODS

The setting was the ENT and paediatric clinics of two London teaching hospitals. The particular Royal Free and Royal National Throat,

Table 1 Mode of classification of sleep study results

	Snoring	Movement	SaO_2 dips
Normal score = 1	Minimal	Little: not related to breathing	None
Mild score = 2	Quiet, intermittent	Some related to breathing	None or occasional
Moderate score = 3	Louder for longer periods	Definitely related to breathing	Some, but not below 90%
Severe score = 4	Loud, almost continuous	Frequent and related to breathing	Many, some below 90%

Nose, and Ear services are mainly secondary, taking referrals direct from general practitioners. The research protocol received ethics approval from the Ethics Review Committee of the Royal Free Hospital.

During the period 12 March 1995 to 9 March 1998, 139 children were referred for possible SRUAO, and overnight inpatient admission was arranged in the children's ward at the Royal Free Hospital. The following data were collected for patients referred: (1) a parental symptom questionnaire; (2) the referring clinician's impression; and (3) the sleep study data. The parental symptom questionnaire was modelled on questions that Stradling *et al* have found to be useful predictors of SRUAO.^{3,4} The answers were graded 1 to 4: 1 = never, 2 = rarely, 3 = sometimes, 4 = often. The referring clinician's impression was graded from the referral letter: 1 = no abnormality, 2 = mild abnormality but no indication for treatment, 3 = moderate abnormality requiring surgical or medical treatment, 4 = severe abnormality requiring urgent surgical treatment. The sleep studies were graded by VVS as described in table 1.

The following analyses were carried out: (1) an examination of the values from the sleep study to see which best predicted the study score; (2) the relation between the study score and symptoms (from the parental symptom questionnaire); and (3) the relation between the study score and the referring clinician's impression.

RESULTS

Sleep study

A total of 139 children were recruited to the study. Nineteen were excluded because of technical problems (mainly failure of the oximetry probe). The remaining 120 children

were aged 6 months to 15.5 years, with a mean of 4.21 years. Study scores showed that 20% were classified as normal, 35% showed mild sleep related upper airway obstruction, 28% moderate, and 18% severe.

The relation between the overall grade and the individual measures of oximetry and movement was examined. For movements per hour (fig 1), although there was a statistically significant difference in movement score between those categorised as 1 and 4, and those categorised as 3 and 4, there was considerable overlap in movement scores between groups. This is not surprising as the movement score is so much affected by artefact.

For pulse rise per hour (fig 2) there were no significant differences between groups 1, 2, and 3. However, groups 1–3 showed significantly less pulse rate rises than group 4. For instance the mean pulse rate rise per hour (95% CI) in group 1 was 43 (39 to 47), while in group 4 it was 57 (48 to 67).

For oxygen dips per hour (fig 3) the same phenomenon is seen. Groups 1, 2, and 3 are not significantly different from each other, but are different from group 4. The mean number of oxygen dips in group 3 was 4.5 (95% CI 3.2 to 5.7), while in group 4 it was 10.4 (95% CI 7.6 to 13.3).

Thus, the individual measure which related best to the overall result was the number of dips per hour. However, the most severe studies also showed more movement and more pulse rate rises per hour.

Comparison between clinician's impression and sleep study score

There were 106 cases in which both sleep study scores and the referring clinician's impression were available for comparison. There was a significant relation between the two measures,

Figure 1 (A) Dot plot of movements per hour and sleep study score. (B) Relation between movements per hour and sleep study score.

Figure 2 (A) Dot plot of pulse rise per hour and sleep study score. (B) Relation between pulse rise per hour and sleep study score.

Figure 3 (A) Dot plot of oxygen dips per hour and sleep study score. (B) Relation between oxygen dips per hour and sleep study score.

$\kappa = 0.1843$, with complete agreement in 45 cases (table 2). However, both sensitivity and specificity of the clinician's impression of moderate or severe SRUAO were low (59% and 73% respectively). In a similar number of cases the clinicians underestimated (17%) and overestimated (16%) study results. In 11 cases the clinicians differed from the sleep study by two or more grades.

Symptom questionnaires

A total of 74 symptom questionnaires were filled in by parents. There was a very high incidence of SRUAO symptoms, making it unlikely

that any individual symptom would prove a good discriminator (table 3). A total of 95% of children suffered from coughs and colds, 85% of the children snored and mouth breathed, and 78% had restless sleep. However, the response rates for hyperactivity (50%) and daytime sleep (47%), together with Stradling's observation that these are good predictors in the general population,⁴ made these symptoms worthy of comparison with the sleep study scores. Only hyperactivity showed more agreement than expected by chance (table 4), but sensitivity and specificity were modest (64% and 55% respectively) with positive predictive

Table 2 Comparison of clinicians' impression and study scores

	Referring clinician's impression*				Row totals
	1	2	3	4	
Study scores					
1	4	16	4	1	25
2	2	23	10	2	37
3	0	14	11	5	30
4	0	4	3	7	14
All groups	6	57	28	15	106
<i>Kappa analysis</i>					
Agreement		42.45%			
Expected agreement		29.45%			
κ		0.1843			
Z		3.30			
Pr>Z		0.0005			

*1 = no abnormality, 2 = mild abnormality but no indication for treatment, 3 = moderate abnormality requiring surgical or medical treatment, 4 = severe abnormality requiring urgent surgical treatment.

Table 3 Results of 74 symptom questionnaires in sleep study

Symptom	Number of positive responses (sometimes or often)	Percentage of positive responses	Missing results
Coughs and colds	71	95	0
Frequent snoring	63	85	0
Mouth breather	63	85	7
Restless sleeper	54	78	5
Problem related to sleeping	56	77	1
Night sweats	53	72	0
Volume of snoring	50	68	0
Sleeps with head tipped back	41	60	6
Trouble getting to/staying asleep	44	59	0
Hyperactivity	37	50	0
Daytime sleep	35	47	0

Table 4 Relation of hyperactivity with sleep study score

	Sleep study score				Total
	1	2	3	4	
Hyperactivity					
1	3	11	6	5	25
2	1	5	0	1	7
3	3	5	14	3	25
4	3	5	3	1	12
Total	10	26	23	10	69
<i>Kappa analysis</i>					
Agreement		33.33%			
Expected agreement		23.67%			
κ		0.1266			
Z		2.03			
Pr>Z		0.0210			

value 45%. Daytime sleep ($\kappa = 20.05$) and all other symptoms both individually and in combination showed no agreement.

DISCUSSION

With regard to the individual variables in the sleep studies, oximetry is the most reliable, because of good movement artefact rejection software. The movement score is the least helpful as it is very dependent on non-sleep periods included in the video. However, there is now a Windows version of software that allows the reporter to determine which parts of the record to reject and which to analyse, and this should increase the usefulness of the movement score. This is important as oximetry alone will miss obstructive episodes leading to arousals without hypoxaemia.⁹ Sleep disturbance in itself is thought to be important in causing behavioural and cognitive problems in children.

The numerical sound score was not examined in detail in this study. Adult snoring has a narrow frequency band, and it is relatively easy to design algorithms to detect it. In children other noises are easily confused with snoring as they snore at higher frequencies and throughout a broader spectrum. Although the software was adapted for higher frequency sound, the numerical analysis was still unhelpful. However, it was straightforward for an observer to see from the printout whether the child was having frequent heavy snoring or intermittent light snoring.

Both under and over reporting were equally distributed, indicating that some children with mild SRUAO may be listed for operation and that some with severe SRUAO would not be recognised and might not receive any treatment. Alternatively they might be listed for routine operation and the potential anaesthetic problems would not be predicted. Children with severe SRUAO are at high risk of postoperative respiratory failure.

The reasons for these discrepancies are to some extent illustrated by the lack of correlation between the symptom questionnaire and the sleep study. It appears that some parents are worried and overemphasise snoring, and others do not appreciate the severity of the obstruction because they do not sleep with their child. Another factor is that the severity of sleep disturbance will vary from night to night, being worse with upper respiratory infections. In addition children might sleep less under study conditions in hospital, leading us to under report the severity of their sleep disturbance.

Overall, it is clear that as in adults,⁹ the sleep study does contribute more information to the clinical evaluation of the child in whom significant sleep related upper airway obstruction is suspected. In general the questionnaires are good at selecting high risk individuals out of the general population, but are unable to discriminate further.

Finally, we do not know how much SRUAO causes harm to a child. It is biologically unlikely that our grade 1–2 (normal and mild) study children are having significant sleep disturbance. It is likely that the threshold for harm is somewhere in the grade 3–4 (moderate and severe) study children who have obvious sleep disturbance and some hypoxaemia. Our statistical analysis suggests that those in group 4 are distinct from groups 1–3. Intervention studies should be targeted at this group.

Overall conclusion

Clinical methods can detect a high risk group for SRUAO but cannot predict exactly which individuals have SRUAO. The Visilab video system is easy to use and robust if its limitations are understood, and can be applied for further evaluation of this high risk group. It is particularly useful when there are definite symptoms of SRUAO but their severity is in doubt.

Hence sleep studies can contribute to assessing the need for operation, the likelihood of postoperative respiratory failure, or as a

baseline or outcome measure in intervention studies.

We would like to thank Dr David St George and Dr Lindsay Forbes for statistical help. This work was undertaken with the Royal Free Hampstead NHS Trust who received a proportion of its funding from the NHS executive; the views expressed in this publication are those of the authors and not necessarily those of the Trust or the NHS Executive.

- 1 Ali NJ, Pitson D, Stradling JR. Sleep disordered breathing: effects of adenotonsillectomy on behaviour and psychological functioning. *Eur J Pediatr* 1996;155:56–62.
- 2 Guilleminault C, Korokbin R, Winkle R. A review of 50 children with obstructive sleep apnoea syndrome. *Lung* 1981;159:275–87.
- 3 Stradling JR, Thomas G, Warley AR, Williams P, Freeland A. Effect of adenotonsillectomy on nocturnal hypoxaemia, sleep disturbance, and symptoms in snoring children. *Lancet* 1990;335:249–53.
- 4 Ali NJ, Pitson DJ, Stradling JR. Snoring, sleep disturbance, and behaviour in 4–5 year olds. *Arch Dis Child* 1993;68:360–6.
- 5 Gislason T, Benediktsdottir B. Snoring, apneic episodes, and nocturnal hypoxaemia among children 6 months to 6 years old. An epidemiological study of lower limit of prevalence. *Chest* 1995;107:963–6.
- 6 Gaultier C. Sleep-related breathing disorders 6. Obstructive sleep apnoea syndrome in infants and children: established facts and unsettled issues. *Thorax* 1995;50:1204–10.
- 7 van Someren VH, Hibbert J, Stothers JK, Kyme MC, Morrison GA. Identification of hypoxaemia in children having tonsillectomy and adenoidectomy. *Clin Otolaryngol* 1990;15:263–71.
- 8 Croft CB, Brockbank MJ, Wright A, Swanston AR. Obstructive sleep apnoea in children undergoing routine tonsillectomy and adenoidectomy. *Clin Otolaryngol* 1990;15:307–14.
- 9 Ryan PJ, Hilton MF, Boldy DA, *et al.* Validation of British Thoracic Society guidelines for the diagnosis of the sleep apnoea/hypopnoea syndrome: can polysomnography be avoided? *Thorax* 1995;50:972–5.

FETAL AND NEONATAL EDITION

July 2000 issue

The following articles—being published in the July 2000 issue of the *Fetal and Neonatal* edition of the *Archives of Disease in Childhood*—may be of general interest to paediatricians.

ETHICS

Drafting guidelines for the withholding or withdrawing of life sustaining treatment in critically ill children and neonates

V F Larcher, L Doyal

ORIGINAL ARTICLES

Increasing rates of cerebral palsy across the severity spectrum in north-east England 1964–1993

A F Colver, M Gibson, E N Hey, S N Jarvis, P C Mackie, S Richmond for the North of England Collaborative Cerebral Palsy Survey

Fetal growth and infantile colic

C Søndergaard, E Skajaa, T B Henriksen