Respiratory Monitoring by Means of an Unattended Device in Children With Suspected Uncomplicated Obstructive Sleep Apnea*

A Validation Study

Marco Zucconi, MD; Giliola Calori, PhD; Vincenza Castronovo, PhD; and Luigi Ferini-Strambi, MD

Study objective: **To compare an unattended device for cardiorespiratory monitoring (POLY-MESAM; MAP; Martinsried, Germany) [P-M] with classic nocturnal polysomnography (PSG) for diagnosis of obstructive sleep apnea (OSA) in children.**
Design: **Clinical setting.**
Patients: **Twelve children (age range, 3 to 6 years) with highly suspected uncomplicated OSA who underwent PSG and P-M on 2 consecutive laboratory nights in a balanced manner.**
Measurements: **Respiratory indexes were compared for P-M (automated analysis), hand-scored revised P-M (P-Mrev), and PSG. Analysis of contingency for cutoff levels of respiratory disturbance index (RDI) of 5 and 10 and level of agreement between P-M, P-Mrev, and PSG by the concordance method were evaluated.**
Results: **Nine of twelve children (75%) had a PSG RDI > 5, while 41.7% had an RDI > 10, indicating moderate-to-severe OSA. P-M sensitivity (78%) increased with the increase of the RDI cutoff, and P-Mrev sensitivity reached 100% at the cutoff of 10. The specificity was low for RDI > 5 and increased only modestly at RDI > 10 (P-Mrev, 57%). Seven of 12 children (increasing to 9 children with P-Mrev) and 9 of 12 children (increasing to 11 children with P-Mrev) were correctly classified by the P-M unit when cutoffs of 5 and 10 were considered, respectively. As far as the agreement level is concerned, P-M underestimated the incidence of obstructive hypopnea and overestimated the number of central apnea cases. P-Mrev improved the latter measure.**
Conclusion: **Based on these data, the P-M device cannot be advocated for common use in a clinical setting, but it may have a role in urgent screening for highly suspected moderate-to-severe OSA.**
(CHEST 2003; 124:602–607)

Key words: childhood obstructive sleep apnea; nocturnal polysomnography; respiratory monitoring; unattended device; validation study

Abbreviations: AHT = total number of apnea and hypopnea events; CA = central apnea; CI = confidence interval; DE = desaturation event; NPV = negative predictive value; P-M = POLY-MESAM; P-Mrev = revised POLY-MESAM; OA = obstructive and mixed apnea; ODI = oxygen desaturation index; OH = obstructive hypopnea; OSA = obstructive sleep apnea; PPV = positive predictive value; RDI = respiratory disturbance index; Sao_2 = arterial oxygen saturation; Sao_2% = percentage of arterial oxygen saturation

Attended overnight polysomnography (PSG) performed in a sleep laboratory is the currently accepted technique used for the diagnosis of sleep-disordered breathing in children.[1] However, PSG is an expensive and time-consuming test, and the results may be affected by environmental and instrumentation effects. Ambulatory monitoring with unattended devices is considered reliable for the diagnosis of obstructive sleep apnea (OSA) in particular groups of adult patients,[2–7] but is not yet recommended for children since clinical and validation studies are lacking.[1] Validation for some devices, as nocturnal oximetry in suspected OSA for adenotonsillar hypertrophy, reveals that such devices have predictive value in positive OSA cases, but poorly

*From the Sleep Disorders Center, Department of Neurology, and Statistic Unit, IRCCS H San Raffaele, Milan, Italy.
Manuscript received July 23, 2002; revision accepted February 13, 2003.

Correspondence to: Marco Zucconi, MD, Sleep Disorders Center, Department of Neurology, San Raffaele Scientific Institute and Hospital, Via Stamira d'Ancona 20, 20129 Milan, Italy; e-mail: zucconi.marco@hsr.it

detect negative cases.[8] An unattended device for cardiorespiratory monitoring (POLY-MESAM; MAP; Martinsried, Germany) [P-M] has been validated in adults with suspected OSA showing good sensitivity and specificity, compared to PSG at different event threshold values. However, due to some underestimation of obstructive apneas, it produces some false-negative results in patients with mild-to-moderate OSA.[9,10]

The aim of this study was twofold: to test, in a laboratory setting, the P-M device in comparison to an attended classic nocturnal PSG in children with suspected OSA; and to verify the accuracy of the automatic scoring of the device with or without a revised analysis by visual editing of the raw data. The verification was relative to PSG results.

Materials and Methods

Twelve children (8 male and 4 female; age range, 3 to 6 years) with a clinical and sleep history of highly suspected OSA were enrolled in the study, after an informed consent form was signed by one of the parents. We chose this age range since clinical and overnight sleep monitoring epidemiologic studies document a high prevalence of habitual snoring, ranging from 7 to 12%, and of OSA, ranging from 3 to 12%, in this age group.[11]

From a clinical history, the following features were found (mean ± SD): age, 4.0 ± 0.8 years; body mass index, 16.6 ± 3.5; snoring onset, 18.7 ± 11.3 months; familiarity for habitual snoring in 10 of 12 cases (80%); excessive daytime sleepiness in 6 of 12 patients (50%); and failure to thrive in 3 of 12 patients (25%). Reported apneas, daytime irritability, forced daytime oral respiration, and recurrent upper airway infections were present in all the children. In all the children except one, who previously had an adenoidectomy, there was a moderate-to-severe adenotonsillar hypertrophy.

The whole group underwent full-night PSG (Grass, Digital System Heritage; Astro-Med; West Warwick, RI) and a seven-channel unattended recording with P-M for 2 consecutive nights in the sleep laboratory in a balanced manner: six children underwent P-M the first night and PSG the second night, and the other six children were recorded in the opposite order. The accompanying parent was allowed to stay in the recording room and instructed how not to interfere. The time schedule (9 PM admittance, 10:30 PM lights out, 7 AM lights on) was the same for both days.

The P-M unit is a small device that can be attached to the patient and, in the standard setting, has seven channels of recording: (1) flow sensors for oronasal breath flow (thermistors); (2) a laryngeal microphone for detection of snoring sound; (3) an ECG lead; (4) a stress-sensitive belt for thoracic effort, and two other belts for abdominal effort (5) and body position (6); and (7) a pulse oximeter with a finger probe for detection of percentage of arterial oxygen saturation (Sao_2). A computer-based analysis automatically calculates apnea, hypopnea, and desaturation events, and the index of each variable per hour of recording time, based on the setup of each variable. PSG included two EEG channels (C3-A2 or C4-A1 and O1-A2 or O2-A1), two electrooculographic channels (right outer canthus and left outer canthus), submental electromyogram (genioglossus muscle), ECG (one lead), laryngeal microphone, oronasal (thermistor) flow, thoracic and abdominal effort (belt), body position, and finger probe oximetry.

The evaluation of both PSG and P-M recordings was done by the same examiner in different times, in different orders and in blinded fashion, for each child examined. For both PSG and P-M events, the following were evaluated: obstructive and mixed apnea (OA), central apnea (CA), not related to previous body movements and with oxygen desaturation of > 4%; obstructive hypopnea (OH), considered as a reduction of > 50% in the thermistor signal associated with oxygen desaturation of > 4%; and desaturation events (DEs) > 4%. The total number of apnea and hypopnea events (AHT), respiratory disturbance index (RDI) per hour, and oxygen desaturation index (ODI) per hour were calculated according to the total time in bed for both types of recording. The lowest percentage of Sao_2 (Sao_2%) was calculated as the minimum Sao_2 level associated with events, and the mean low Sao_2% was measured as the mean of the peak Sao_2% level associated with each single event. In order to discern motion artifacts, trend and event graphs of oximetry were printed and the oximeter pulse waveform was recorded.

For the P-M unit, both automated and hand-scored revised data analyses (revised POLY-MESAM [P-Mrev]) were provided and assessed. For P-Mrev, the automatic analysis served as a base for scored apnea and hypopnea events. The reviewer, having an overview of all channels, recorded snoring louder at the end of events, pulse rate, and Sao_2% synchronous variations with respiratory events. He evaluated and corrected the automatic analysis, taking into account possible artifacts and errors.

The analysis of sleep for PSG recording was performed according to Rechtschaffen and Kales.[12] The minimum duration of apnea-hypopnea events was set at 8 s for each scoring. There is little consensus on the level of RDI or ODI that defines a child as "affected" or "nonaffected" with OSA; different levels of 1, 5, 10, or 15 are used in different studies.[1,13,14] We chose different thresholds for RDI and ODI (5 and 10) to measure sensitivity, specificity, positive predictive values (PPVs), and negative predictive values (NPVs). These threshold levels are higher than what is generally considered abnormal,[15] but better serve to identify children with a clinically significant level of respiratory events who exhibit at least a moderate degree of OSA.[13] Affected or nonaffected subjects as defined by different PSG cutoff levels were compared by means of contingency tables. Ninety-five percent confidence intervals (CIs) were calculated according to the efficient-score method, after correction for the continuity, as the Gaussian approximation is not well suited for small proportions.[16] Pearson correlation coefficients were also calculated. Agreement between P-M, P-Mrev, and PSG indexes was analyzed according to the Bland and Altman method of concordance.[17] Since a relationship between the difference scores and the size of the measurements emerged, a logarithmic transformation of the raw data were employed. The hypothesis that the mean difference was equal to zero was examined by a paired t test. The mean and 95% limits of the CI of the difference between the two methods, after anti-log transformation, for each analyzed parameter are presented. The anti-log difference between two values on a log scale is a dimensionless ratio; therefore, the statistical significance was obtained when 95% CI of the difference was entirely above or below 1.

Results

The results (raw data) from the PSG, P-M, and P-Mrev are presented in Table 1. Nine of 12 children (75%) showed an RDI > 5, while 41.7% had an RDI > 10 indicating moderate-to-severe OSA. The only significant coefficients of correlation were between PSG and P-Mrev for RDI (0.57; $p = 0.05$), AHT

Table 1—*Respiratory Data (Raw Data) Obtained From PSG, P-M, and PM-Rev

Patient No.	Time in Bed, min (PSG)	OA	OH	CA	AHT	RDI	DE	ODI	Minimum SaO_2 %	Mean SaO_2 %
PSG data										
1	594	2	0	13	15	1,5	44	9.36	89	92,9
2	533	25	29	8	62	7	74	7.12	84	93,5
3	497	150	58	5	213	25,7	110	13,3	72	87,3
4	496	1	5	14	20	1,4	7	19.12	94	95,4
5	488	215	153	48	316	38,9	216	26,5	81	92,2
6	492	14	47	8	69	8,4	54	6,6	88	94,4
7	553	73	115	85	273	29,6	74	8	85	89
8	497	22	9	11	42	5,1	48	5,8	83	91,5
9	508	110	39	15	164	19,4	326	38,5	76	90,9
10	542	72	74	25	171	18,9	232	25,7	70	88,9
11	499	3	13	13	29	3,5	6	0,7	96	96
12	511	9	47	12	68	8	118	13,9	88	94,8
P-M data										
1	555	10	52	21	83	9	22	2	77	91
2	475	49	121	27	197	25	62	8	88	92
3	374	41	160	1	202	31	106	17	78	90
4	520	3	40	4	47	6	10	1	90	93
5	505	1	63	1	65	7	241	29	67	88
6	471	17	26	0	43	5	120	15	78	90
7	528	7	90	3	110	11	61	7	89	93
8	480	4	16	0	20	3	46	6	85	92
9	505	104	208	2	314	38	532	63	50	71
10	515	29	217	9	255	29	364	42	64	89
11	503	3	54	7	64	7	12	1	92	93
12	395	8	61	2	71	11	168	26	68	86
P-Mrev data										
1		9	30	24	63	7				
2		33	122	30	185	23				
3		42	151	13	206	33				
4		2	40	15	57	7				
5		3	216	6	225	27				
6		13	47	2	62	8				
7		23	134	7	164	18				
8		1	16	7	24	3				
9		142	256	6	404	48				
10		80	301	10	391	45				
11		10	87	12	109	13				
12		11	125	2	138	21				

*For PSG and P-M, all the examined measurements are reported; for P-Mrev, SaO_2 parameters were not analyzed.

(0.57; p = 0.05), and OH (0.592; p = 0.04). Table 2 shows the validity of P-M and P-Mrev indexes at different threshold values from PSG analysis. Sensitivity increased with the increase of the RDI threshold and revised RDI reached 100% at the cutoff of 10, as did the NPV. The specificity was low for RDI > 5 and increased at RDI > 10, but decreased with revised RDI. In terms of correct classification and identification of OSA with a cutoff of RDI > 5, 7 of 12 children were correctly classified by the P-M unit (increasing to 9 children with P-Mrev) and 9 of 12 children (increasing to 11 with children P-Mrev) when a cutoff of 10 was considered. As far as the agreement level between PSG and P-M is con-

Table 2—*Contingency Analysis (95% CI) for P-M and P-Mrev vs PSG for Different Cutoff of RDI*

Variables	RDI > 5	RDI > 10	Revised RDI > 5	Revised RDI > 10
Sensitivity, %	78 (40–96)	80 (30–99)	89 (51–99)	100 (46–100)
Specificity, %	0 (0–69)	71 (30–95)	0 (0–69)	57 (20–88)
PPV, %	70 (35–92)	67 (24–94)	73 (39–93)	63 (26–90)
NPV, %	0 (0–80)	83 (36–99)	0 (0–95)	100 (40–100)

Table 3—*Bland and Altman Analysis for P-M vs PSG

Variables	RDI	AHT	OA	OH	CA	ODI	DE	Minimum Sao_2%	Mean Sao_2%
Mean difference (PSG − P-M)	0.86	0.78	1.98	0.31	6.75	0.82	0.84	1.1	1.04
95% CI of difference	1.5–0.5	1.5–0.4	6.5–0.6	1.0–0.1	22.1–2.1	1.1–0.63	1.1–0.64	1.2–1.0	1.1–0.9
p value	0.41	0.59	0.23	0.05	0.005	0.12	0.18	0.055	0.11

*Since there is an association between the differences of the two methods and the size of the measurements, a logarithmic transformation of the raw data has been employed.

cerned, only two variables (OH and CA) showed significant differences (Table 3), indicating that for OHs, there was an underestimation of P-M compared to PSG, and for CAs, there was an overestimation for P-M compared to PSG. When we considered the revised data, OH remained significantly different from the PSG data. RDI and AHT differences were significant, while CA detection dramatically improved (Table 4; Fig 1).

Discussion

Laboratory PSG is considered the "gold standard" for diagnosis of OSA both in adults and children. While in adults the validity of ambulatory monitoring (level III devices) for identification of OSA at least in some groups of patients or in some occasions is indicated, to date, in children, there is no official statement addressing the role of these units in evaluating OSA.[1] However, a better understanding of the pathology and an increased demand for sleep studies in children with suspected OSA who await surgery (*ie*, adenotonsillectomy) raise the question of reliability for this type of equipment in correctly detecting respiratory events during sleep. With these concerns and with the intention of evaluating a possible diagnostic instrument, we tested the P-M device in a sleep laboratory setting in a sample of children with a high probability of suffering from OSA.

The study shows relevant correlation between data collected by the two systems only when we used the revised data for RDI, AHT, and OA, but not for OH and CA. The sensitivity of the device made it possible to detect suspected OSA especially when revised data and higher thresholds were used. The device was less specific, regardless of the RDI cutoff or the revision of the data; however, despite difficulties with the probes and artifacts on the signals, 11 of the 12 children were correctly classified by the P-Mrev for discrimination of an overt OSA (RDI > 10). The agreement test showed that the P-M device underestimates respiratory events, regardless of the visual reanalysis of the raw data, but the means, using the Bland and Altman method, differed only a few units for each parameter considered. The greatest difference involved CA scoring, which improved using the revised analysis. The point of CA analysis was to underscore that central events were marked, even if no displacement was noted on both thoracic and abdominal sensors. Without EEG, electrooculography, esophageal pressure measures, or inductance plethysmography, the reliability of this measurement is low for signal scoring. During PSG, one can intervene to correct signals, unlike the case of an unattended device. For the hypopneas, the explanation for underscoring events may lie in the definition of hypopnea. Some events were unrecognized by the automatic analysis but were added with the revision of the data, which also considered borderline events (*ie*, a 40% decrease in oronasal flow but with desaturation) as well as for the PSG. Other factors may have influenced the difference in agreement between P-M and PSG. The quality of the signal from the thermistor, in children, is very sensitive to body position or movements during the night and the night-to-night variability of RDI.[18] However, a recent study[19] has verified the variability of respiratory measures in children with suspected OSA and found little clinically significant night-to-night variability and no first-night effect. No study addressed night-to-night variability in respiratory

Table 4—*Bland and Altman Analysis for P-Mrev vs PSG

Variables	RDI	AHT	OA	OH	CA
Mean difference (PSG − P-Mrev)	0.55	0.62	1.6	0.23	1.82
95% CI of difference	0.9–0.3	1–0.3	4.8–0.5	0.6–0.1	3.9–0.8
p value	0.025	0.045	0.37	0.007	0.12

*Since there is an association between the differences of the two methods and the size of measurements, a logarithmic transformation of the raw data has been employed.

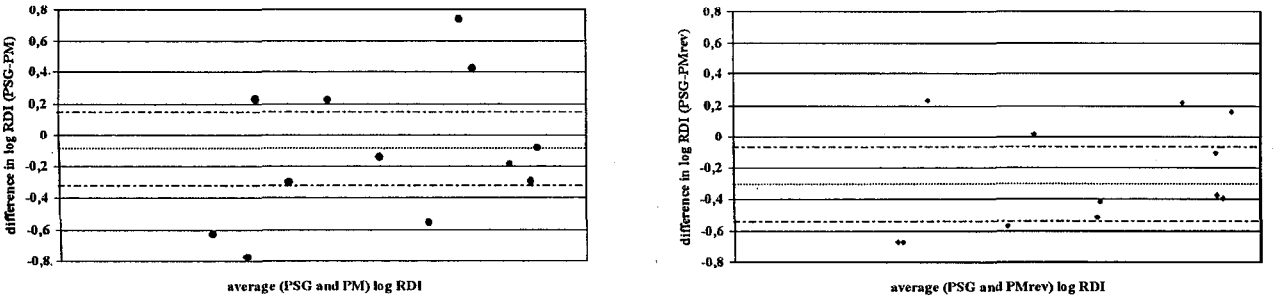

FIGURE 1. Bland and Altman plots of the difference log RDI (PSG − P-M) against the mean log RDI (PSG and P-M) both for P-M and P-Mrev.

parameters measured by unattended devices in children. Of course, our protocol included 2 consecutive nights of recording, with very low potential night-to-night variability.

We are concerned with the major limitations of study, due to the low number of children studied and the very high OSA probability in our sample, confining the results to a group of highly clinically selected OSA candidates. Indeed, it is very difficult for very young children (and parents) to accept two consecutive in-laboratory PSGs or portable monitoring sessions. Thus, an attempt to increase the number of subjects was unproductive. However, the sample is representative of uncomplicated childhood OSA (associated with adenotonsillar hypertrophy and/or obesity).[1] Another limitation for the application and generalization of our results was the laboratory setting of the study. Perhaps the sensitivity and specificity may differ in a home setting as opposed to a sleep laboratory; however, we attempted to simulate a normal and constant environment for the children by obtaining the presence of both parents in the laboratory, television or videorecording viewing before retiring to the sleep room, etc., and there was no problem noted in adapting to the laboratory.

Different portable recordings have been utilized for detecting OSA in children[8] but only a few have been validated. Pulse oximetry has been studied in some samples, obtaining results not sufficient to exclude OSA when the results are negative, since obstructive events in some patients may lead to arousal and consequent sleep disturbances but without significant desaturations.[20,21] However, a large-scale study[22] in OSA candidates for adenotonsillectomy found a 97% PPV but an NPV of 47%, indicating that oximetry was useful only when results are positive. Only ambulatory PSG (as the Compumedics PS2 system [Compumedics; Abbotsford, Victoria, Australia] or the Edentrace I and II [Edentec; Eden Prairie, MN]) showed unequivocal reliability in adults,[19] but they have been tested in a very low number of young children; these devices, however provided satisfactory correlation results.[23,24]

Our results with the P-M device, although not wholly satisfactory, indicate that further validation studies must be conducted before recommending the wide use of this unattended device for OSA diagnosis in children. Larger samples should be evaluated, since the majority of studies in children have involved only a small number of patients. In conclusion, to date, the P-M cannot be advocated for common use in a clinical setting. However, it may have a role in reducing the waiting list for PSG in highly suspected moderate-to-severe OSA in children when treatment is urgently needed.

ACKNOWLEDGMENT: We thank Daniele Bizzozzero, Antonio Massimo, Giampaolo Lecciso, Mauro Secli for assessment of the recordings, and the parents of the children.

REFERENCES

1 American Academy of Pediatrics. Clinical practice guideline: diagnosis and management of childhood obstructive sleep apnea syndrome. Pediatrics 2002; 109:704–712
2 Ferber R, Millman R, Coppola M, et al. ASDA standards of practice: practice parameters for the use of portable recording in the assessment of obstructive sleep apnea. Sleep 1994; 17:378–392
3 Anconi-Israel S, Kripke DF, Mason W, et al. Comparison of home sleep recordings and polysomnograms in older adults with sleep disorders. Sleep 1981; 4:283–291
4 Gyulay S, Gould D, Sawyer B, et al. Evaluation of a microprocessor-based portable home monitoring system to measure breathing during sleep. Sleep 1986; 10:130–142
5 Stoohs R, Guilleminault C. MESAM 4: an ambulatory device for detection of patients at risk for obstructive sleep apnea syndrome (OSAS). Chest 1992; 101:1221–1227
6 Zucconi M, Ferini-Strambi L, Castronovo C, et al. An unattended device for sleep-related breathing disorders: validation study in suspected obstructive sleep apnoea syndrome. Eur Respir J 1996; 9:1251–1256
7 Redline S, Sanders M, Lind B, et al. Methods for obtaining and analyzing unattended polysomnography data for a multicenter study. Sleep 1998; 21:759–767
8 Jacob SV, Brouillette RT. Home monitoring of sleep and breathing in children. In: Loughlin GM, Carroll JL, Marcus

CL, eds. Sleep and breathing in children. New York, NY: Marcel Dekker, 2000; 783–811
9 Verse T, Pirsig W, Junge-Hulsing B, et al. Validation of the POLY-MESAM seven-channel ambulatory recording unit. Chest 2000; 117:1613–1618
10 Marrone O, Salvaggio A, Insalaco G, et al. Evaluation of the POLY-MESAM system in the diagnosis of obstructive sleep apnea syndrome. Monaldi Arch Chest Dis 2001; 56:486–490
11 Ali NJ, Stradling JR. Epidemiology and natural history of snoring and sleep-disordered breathing in children. In: Loughlin GM, Carroll JL, Marcus CL, eds. Sleep and breathing in children. New York, NY: Marcel Dekker, 2000; 555–574
12 Rechtschaffen A, Kales A. A manual of standardized terminology techniques and scoring system for sleep stages in human subjects. Bethesda, MD: National Institutes of Health, 1968; Publication No. 204
13 Redline S, Tishler PV, Schluchter M, et al. Risk factors for sleep-disordered breathing in children: association with obesity, race, and respiratory problems. Am J Respir Crit Care Med 1999; 159:1527–1532
14 Redline S, Sanders M. Hypopnea, a floating metric: implication for prevalence, morbidity estimates, and case finding. Sleep 1997; 20:1209–1217
15 American Thoracic Society. Standards and indications for cardiopulmonary sleep studies in children. Am J Respir Crit Care Med 1996; 153:866–878
16 Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med 1998; 17:857–872
17 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1:307–310
18 Rapoport DM, Kiley JP, Nieto FJ, et al. Night to night variability in sleep and respiratory data collected during NPSG's as part of the Sleep Heart Study [abstract]. Sleep 1998; (Suppl)212
19 Katz ES, Greene MG, Carson KA, et al. Night-to-night variability of polysomnography in children with suspected obstructive sleep apnea. J Pediatr 2002; 140:589–594
20 Williams AJ, Yu G, Santaigo S, et al. Screening for sleep apnea using pulse oximetry and clinical score. Chest 1991; 100:631–635
21 Cooper BG, Veale D, Griffiths CJ, et al. Value of nocturnal oxygen saturation as a screening test for sleep apnoea. Thorax 1991; 46:586–588
22 Brouillette RT, Morielli A, Leimanis A, et al. Nocturnal pulse oximetry as an abbreviated testing modality for pediatric obstructive sleep apnea. Pediatrics 2000; 105:405–411
23 Goodwin JL, Enright PL, Kaemingk KL et al. Feasibility of using polysomnography in children for research: report of the Tucson Children Assessment of Sleep Apnea Study (TuCASA). Sleep 2001; 24:937–944
24 Morton S, Rosen C, Larkin E, et al. Predictors of sleep-disordered breathing in children with a history of tonsillectomy and/or adenoidectomy. Sleep 2001; 24:823–829