Is there a cost-effective way to diagnose mild sleep-disordered breathing?

A. Bachour*, J. Herrala†, P. Maasilta*

*Department of Pulmonary Medicine, Helsinki University Central Hospital, Finland and †Department of Pulmonary Medicine, Tampere University Hospital, Finland

Abstract  
Question of the study: To determine the utility and the cost-effectiveness of oesophageal pressure, respiratory flow and movement, and oximetry (ORO) as a diagnostic tool for mild sleep-disordered breathing (SDB), as compared with overnight polysomnography (PSG).

Patients and methods: Seventy-nine patients evaluated for mild SDB by PSG and simultaneously by oesophageal pressure (Pes) measurement, oximetry, respiratory flow and respiratory movement on a single night. An oesophageal event (OE) was defined as irregular respiration with crescendo in Pes and rapid return to baseline with a minimal increase in the negative Pes at the end of the OE of at least 5 cm H₂O or more than 50% of the baseline level. SDB was defined by ORO when oesophageal events were > 5/h, and by PSG when the respiratory disturbance index was > 5/h. The diagnostic accuracy and cost-effectiveness of ORO were compared with PSG.

Results: Although the ability of ORO to detect SDB was poor: sensitivity 64%, specificity 78%, use of ORO for screening prior to PSG would have saved 5000 EUR per 100 patients compared to initial PSG.

Conclusion: Using the combination of oesophageal pressure, respiratory flow and movement and oximetry for the diagnosis of mild SDB is not cost-effective, because of its poor diagnostic accuracy. New devices having alternative means to predict arousal and respiratory effort variation should be evaluated for cost-effectiveness.

Keywords  
mild sleep-disordered breathing; cost-effectiveness; oesophageal pressure.

INTRODUCTION

Sleep-disordered breathing (SDB), a disorder syndrome associated with repetitive apnoeas and hypopnoeas during sleep with oxygen desaturations, sleep fragmentation, and daytime sleepiness, has been estimated to affect 2–4% of the adult middle-aged population (1). Untreated SDB has several consequences, such as an increased risk for motor vehicle accidents (2). Even in subjects with mild forms of SDB, with the respiratory disturbance index (RDI) at > 5/h, hypertension, sleepiness, and motor vehicle accidents may occur (3). Meanwhile, in a recent cross-sectional study (4), SDB showed a modest-to-moderate effect on various manifestations of cardiovascular diseases, within a range of apnoea and hypopnoea index values considered normal or only mildly elevated. In addition, patients with undiagnosed SDB had considerably higher medical costs than did age- and sex-matched individuals without SDB (5), and treatment with continuous positive airway pressure improve the patients’ symptoms even in mild cases (6).

As sleep monitoring is expensive and resource intensive, several studies have considered alternative methods and revealed that the diagnosis of moderate and severe SDB can be made by nocturnal oximetry alone (7–9), but the diagnosis of mild forms requires overnight polysomnography (PSG) that includes monitoring of sleep as well as of breathing. Yamashiro et al. (10), in fact, found that 30% of patients considered normal by oximetry alone had a mild form of SDB on PSG. In addition, several other studies (11,12) using nocturnal oximetry for the diagnosis of SDB and showing low specificity values did not take into account milder forms of the disease, and concluded that all abnormal oximetry results required PSG for confirmation.

Patients with a mild form of SDB usually present with variable degrees of oxygen desaturation episodes, few apnoeas or hypopnoeas and several sequences of breaths characterised by increasing respiratory effort leading to arousal (RER-A). Measurement of oesophageal pressure (PES) with continuous overnight monitoring is the reference standard for measuring respiratory effort. No other technique has demonstrated sufficient accuracy.
and precision, or has shown a correlation with clinically important outcomes (13).

As sleep monitoring is the most expensive and resource-intensive part of overnight polysomnography, we selected oesophageal pressure, respiratory movements and airflow, and oximetry (ORO) and studied this hypothetical system in the diagnosis of mild sleep-disordered breathing (SDB). We report here the diagnostic accuracy and cost-effectiveness of ORO for the diagnosis of mild SDB compared with those of PSG.

PATIENTS

We reviewed the sleep studies of consecutive patients referred for evaluation of possible mild SDB who underwent PSG with measurement of $P_{es}$ in our sleep laboratory from 1998 to 2000. Mild SDB was clinically suspected in patients with mild-to-moderate excessive daytime sleepiness and a history of snoring. A recording with the oesophageal catheter was started in 88 patients; 79 slept more than 2 h with the oesophageal catheter and were included in this study. We studied 51 men and 28 women, mean (SD) age 49 (10) years, neck circumference 39 (4) cm, and body mass index (BMI) 28 (5) kg/m² (Table 1).

METHODS

The severity of daytime sleepiness was evaluated by the Epworth sleepiness scale (ESS) and classified clinically according to AASM recommendations (13). Polysomnographies were performed with a computerised 24-channel polygraph (Alice 3, Healthdyne Technologies, Marletta, GA, U.S.A.). This included a four-channel electroencephalogram (C3A2, C4A1, O1A2 and O2A1), electro-oculogram, and submental and leg electromyogram. Heart rate was monitored through standard leads. Airflow was monitored by a nasal and oral thermistor. Thoracic and abdominal belts (Healthdyne piezo effort sensor) were used for respiratory movement detection. Pulse oximetry and nasal expired CO$_2$ were monitored by (BCI Capnocheck plus, BCI International, Waukesha, WI, U.S.A.). Oesophageal pressure was monitored by a piezoresistor pressure sensor (Synectics FTC catheter, French size F8/2.7 mm. Synectics Medical AB, Stockholm, Sweden) with a sampling rate at 40 Hz. The oesophageal catheter was calibrated regularly, inserted through the nose after local anaesthesia and fixed about 38 cm from the nostril. This manoeuvre was abandoned if the patient complained of discomfort. The catheter was removed during the recording when requested by the patient. A nurse trained in sleep medicine attended the polysomnography.

Sleep was manually scored in epochs of 30 s according to Rechtschaffen and Kales criteria (15). An arousal was defined by an EEG frequency shift into the alpha range for at least 3 s (16). An apnoeic event was defined as cessation of nasal/oral airflow for at least 10 s without regard to either arousal or oxygen desaturation, while a hypopnoea was defined as a $>$50% decrement in nasal/oral airflow for at least 10 s, associated with either an arousal or oxygen desaturation $\geq$3%.

Respiratory effort-related arousal (RERA) events were scored according to AASM recommendations (13). An RERA event was scored when it was not caused by apnoea or hypopnoea, it lasted 10 s or longer with a pattern of progressively more negative oesophageal pressure, and was terminated by a sudden change in pressure to a less negative level and an arousal (Fig. 1).

The respiratory disturbance index RDI was calculated by the following formula: number of apnoeas+hypopnoeas+RERAs/total sleep time (h).

An oesophageal event was registered when a pattern of progressively more negative oesophageal pressure lasted for 10 s or longer and ended with a sudden return to the baseline, regardless of whether it was associated with apnoea or hypopnoea (Fig. 2). The minimal increase in the negative oesophageal pressure at the end of the oesophageal event was at least 5 cm H$_2$O or more than

| Table 1. Patients’ characteristics and data from polysomnography and from ORO |
|---------------------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
|                                | BMI (kg/m²)   | Age (year)    | ESS            | TIB (min)      | TST (min)      | $S_{eff}$ (%)  | Delta (%)      | REM (%)        | ODI$_4$ (%/h)  |
| Mean                            | 28            | 49            | 9              | 458            | 345            | 75             | 16             | 18             | 10             | 5.8            | 33             | 3.4            | 15             | 8.2            | 5.6            | 3.0            |
| SD                              | 5             | 10            | 4              | 50             | 87             | 15             | 9              | 8              | 10             | 0.5            | 30             | 2.3            | 7              | 5               | 6.0            | 2.5            |
| Maximum                         | 45            | 73            | 17             | 542            | 507            | 98             | 34             | 35             | 44             | 6.8            | 95             | 10.6           | 38             | 26.5           | 30             | 20.0           |
| Minimum                         | 19            | 25            | 2              | 268            | 138            | 33             | 0              | 0              | 0              | 44             | 60             | 0.0            | 5              | 0.4            | 0.0            |

Note: ORO=Oesophageal pressure, respiratory flow and movement and oximetry; BMI=body mass index; ESS=Epworth sleepiness scale; TIB=time in bed with oesophageal catheter; TST=total sleep time; $S_{eff}$=sleep efficiency; delta=% of delta sleep, REM=% of rapid eye movement sleep; ODI$_4$=oxygen desaturation index of 4%; ETCO$_2$=mean of end tidal CO$_2$ during TIB. CO$_2$ >6 kPa=% of time in bed when ETCO$_2$ >6 kPa; RERA=respiratory effort-related arousal per sleep hour; Arousal=arousal index, RDI=respiratory disturbance index; OI=oesophageal index; SD=standard deviation.
FIG. 1. Example of a respiratory effort-related arousal (RERA): a pattern of progressively more negative oesophageal pressure, terminated by a sudden change in pressure to a less negative level and arousal. Abbreviations: light = light intensity (from 0 to 50), LEOG = left oculogram; REOG = right oculogram; Chin = submental electromyogram; Esophag = oesophageal pressure; Nose = nasal airflow; Mouth = mouth airflow; ABD = abdominal respiratory movement. THO = thoracic respiratory movement; Micro = snoring; \( \text{G} \text{SaO}_2 \) = \( \text{SpO}_2 \) graphical curve; \( \text{EtCO}_2 \) = end-tidal \( \text{CO}_2 \); \( \text{GCO}_2 \) = graphical curve of \( \text{ETCO}_2 \); Body = body position; S = supine.
50% of the baseline level. Oximetry was used to guide the observer to respiratory events but no minimal drop in oxygen saturation was used for scoring oesophageal event. Oesophageal index (OI) was calculated by the following formula: number of oesophageal events/total time in bed (h).

Sleep efficiency ($S_{eff}$) was calculated by the following formula: total sleep time $\times$ 100/time in bed. During time in bed (TIB), the following parameters were calculated automatically and validated by a chest physician trained in sleep medicine: mean end tidal CO$_2$ (ETCO$_2$), percentage of recording time spend when ETCO$_2$ > 6 kPa, and oxygen desaturation episodes of 4%/h (ODI$_4$).

**PROTOCOL**

**Data analysis**

PSG recordings were scored manually by a chest physician trained in sleep medicine. The diagnosis of sleep-disordered breathing (SDB) was made when the RDI on PSG was > 5/h. From the same polysomnography recording only oesophageal pressure, respiratory flow and movement and oximetry (ORO) were made available for manual scoring to another chest physician trained in sleep medicine, who was blinded to the PSG results but did have access to the patient’s file. The diagnosis of SDB was made when the oesophageal index (OI) for the ORO system was > 5/h.

**Cost analysis**

Two diagnostic algorithms were compared by use of the data from the initial PSGs and OROs. In the first algorithm, ORO served as the initial screening test, and all patients with abnormal ORO (OI > 5/h) results went on to PSG for definitive diagnosis. For the second algorithm, all patients had an initial PSG. In both algorithms, all patients diagnosed as having SDB underwent a one-night CPAP titration for treatment (Fig. 3). During PSG, the additional cost of monitoring oesophageal pressure was estimated to be equivalent to the additional cost caused by CPAP titration. The cost reported for ORO was 250 EUR; the costs for interpreted PSG with CPAP titration or with oesophageal pressure monitoring amounted to 400 EUR, total.
Statistical analysis
For the evaluation of the diagnostic accuracy of ORO we used overnight PSG as the gold standard. We evaluated the sensitivity (true positive/[true positive+false negative]), the specificity (true negative/[true negative+false positive]), the positive predictive value (true positive/[true positive+false positive]), and the negative predictive value (true negative/[true negative+false negative]). We also drew a receiver-operating curve (ROC) for ORO in the detection of SDB with a value OI between 0.2 and 30. When needed, we used the Student's t-test to compare a subgroup with all the subjects.

RESULTS
In 47 of the 79 patients, a diagnosis of SDB was made by PSG, with a mean RDI (sd) of 8.2 (5.6) (Table 1). The ability of ORO to detect SDB varied according to the diagnostic criteria used. For an OI of 5/h, sensitivity was 64%, specificity 78%, positive predictive value 81%, and negative predictive value 60%. This accuracy did not differ significantly between obese and non-obese subjects (Table 2). With a variable diagnostic criterion of OI ranging from 0.2 to 30, the sensitivity decreased from 96 to 0%, and the specificity increased from 6 to 100% (Fig. 4).

All patients slept more than 120 min, with a mean time in bed TIB (sd) of 458 (50) min, total sleep time TST 345 (87) min, sleep efficiency 75 (15)%, delta sleep 16 (9)% of TST, REM sleep 18 (8)% of TST, ETCO2 5.8 (0.5)kpa and ETCO2 > 6 kPa 33 (30)% of TIB (Table I).

Cost analysis
In the first hypothetical algorithm, ORO served as the initial screening test. Of the 79 subjects, 37 had abnormal ORO (OI > 5/h) results and went on to PSG for definitive diagnosis. Of these 37, 30 had SDB and underwent a CPAP titration trial. For the second hypothetical algorithm, all patients had an initial PSG; 47 had mild SDB and underwent a CPAP titration trial (Fig. 3).
The cost of ORO as a screening test for the diagnosis of mild SDB was 59.000 EUR and for the algorithm with PSG as the initial test 64.000 EUR. Use of ORO as a screening test prior to PSG, rather than initial PSG testing, would have saved 5.000 EUR per 100 patients evaluated, if all patients with mild SDB had required a second-night study for CPAP titration.

Although this method reduced costs, 17 of the 42 patients with normal ORO results had mild SDB missed by screening ORO alone; their age was a mean 51.8, BMI 29.3, RDI 9.0, RERA 5.2, and AI 1.5. We found no significant difference ($P > 0.05$) in these variables between these patients and the whole population.

**DISCUSSION**

Our study confirms the findings of previous studies that PSG is still necessary for the diagnosis of mild SDB (8–10, 12). Applying a diagnostic algorithm utilising ORO as a screening tool to guide decisions on which patients should undergo PSG did result in small cost savings compared to use of PSG alone. However, the consequence of ORO for screening was that a significant number of patients with sleep disorders that cause excessive sleepiness would have remained undiagnosed and untreated. The high cost of sleepiness-related motor vehicle, work-related, and home-based accidents could negate these small cost savings (17).

We report the cost-effectiveness according to the cost in our university hospital. Prices in Finland could vary between public health centres up to 30% because of a decentralised health policy. In France, the estimated real cost for an overnight polysomnography in a university hospital in Paris, is at 500 EUR, meanwhile the public health insurance considers the cost at only 150 EUR. In addition, prices of health-care services within Europe could vary significantly. Nevertheless, this variation would not change the conclusion of this study, as the increase in the saving, induced by the use of ORO, in an expensive health-care system would be counterbalanced by the increase in the cost induced by the sleepiness-related motor vehicle accidents.

There are some likely explanations for the failure of ORO to recognise SDB in some patients. Poor $P_{es}$ monitoring may disturb the scoring of an oesophageal event. In fact, it is known that the quality of $P_{es}$ monitoring is related to the position of the pressure sensor in the oesophageal lumen and also to the patient’s position. Poor quality $P_{es}$ data during a portion of the recording have been found in 18 out of 40 study patients (18). In addition, the respiratory effort in response to upper airway occlusion in SDB patients is reported to be lower in REM than in non-REM sleep and decreases with increasing age (19). Meanwhile, a global reduction in $P_{es}$ swings is usually present in obese patients. Our patients had a wide range of age and of BMI in addition to the varying amount of REM sleep. Although obesity may reduce the $P_{es}$ swings, we found that the accuracy of ORO in detecting SDB was low in both obese and non-obese subjects.

Difficulties in scoring oesophageal events and RERAs may lead to underestimation. This risk of underestimation has little effect on RDI, as apnoeas and hypopnoeas are scored regardless of $P_{es}$ swings. It is known that $P_{es}$ swings are more prominent during obstructive apnoea than during a respiratory event with partial upper airway obstruction. Patients with mild SDB may have more partial obstructions than real obstructive apnoeas. Meanwhile, the presence of an arousal at the end of a respiratory event with partial obstruction may alert the physician to the importance of this respiratory event, therefore, such small $P_{es}$ swings are scored as RERAs in polysomnography but probably missed in ORO.

As it is logical to expect that OI is higher than RERA index, we found a few discrepancies in some patients between OI and RERA. This is due in part to the difference in their sleep efficiency. We used TIB as a denominator to calculate OI whereas TST was used for RDI. This will lead to a decrease in OI for the same number of respiratory events. In addition, as there is no general agreement for the exact definition of an OE we may have missed a small number of respiratory events with a mild variation in oesophageal pressure but this variation may have been sufficient to induce an arousal.

We did not correct OI in relation to sleep efficiency. This correction will increase the threshold of normality of OI to more than 5/h. Nevertheless, even with a higher OI threshold, the sensitivity and specificity of ORO in detecting OSA would still be below acceptable values for a screening test.

We used thermistors for measuring airflow. Thermistors are known to give a measure more qualitative than quantitative (20); Nevertheless, for detecting apnoeas,
they have good validity, whereas problems mount in detecting hypopnoeas (I3). Because we monitored $P_{es}$ at the same time, the hypopnoea potentially missed by thermistors would be scored as RERA in PSG, therefore having the final value of RDI unaffected.

We found a high rate of RERAs in relation to respiratory events, whereas Cracowski et al. (21), recently reported a low rate of RERAs in mild-to-moderate sleep disordered breathing syndrome. They used a pneumotachograph for measuring respiratory flow and considered a threshold of 30% as a minimal flow reduction for scoring a hypopnoea. This will lead to an increase in hypopnoeas and a decrease in RERAs.

A recent study (22) reported an interobserver variability of 29.7% in the detection of respiratory effort by $P_{es}$ monitoring. In our study, we did not measure interobserver variability. We calculated the sensitivity of ORO in detecting mild SDB at 64%. This implies that factors other than interobserver variability are likely to account for the results observed. In addition, as it is usual to find a relatively low concordance between different scorers when scoring only one channel (23), we made oximetry available also to the observer of ORO, as desaturation is easy to recognise.

Oximetry is known as an easy and low-cost tool for screening SDB. Nevertheless, a minimal oxygen drop was not included in the diagnostic criteria of OE, as we aimed to increase the sensitivity in scoring OE by recognising respiratory events that are not associated with a drop in oxygen saturation. These respiratory events without a drop in oxygen saturation are usually frequent in young subjects with short respiratory events and healthy lungs.

We found that ETCO$_2$ had a tendency to rise in patients with mild SDB and also during simple snoring. This may rise the question of the effects of such rise in this group of patients.

We tried to find diagnostic method to serve as an alternative to polysomnography in patients with a suspicion of mild SDB. In this particular group of patients, diagnosis requires the monitoring of respiratory effort with an accurate device. Several studies that did not include $P_{es}$ monitoring failed to recognise mild SDB (I12,24). Our results agree with Ross et al. (25) that alternative methods of diagnosis of OSA have provided insufficient evidence to allow their recommendation instead of full polysomnography. Nevertheless, the use of pulse transit time for measuring respiratory effort seems to be promising (22).

$P_{es}$ monitoring, although regarded as the gold standard for measuring respiratory effort (I3), is, however, not without drawbacks. It is invasive, often uncomfortable for the patient, and may not be tolerated. In addition, evidence exists that an oesophageal catheter may modify pharyngeal airway dynamics (26), and it has been suggested that its presence may itself impair the quality of sleep (27), though this is still disputed (28). We used a solid-state piezoresistor pressure sensor with a high sampling rate. The catheter is easy to calibrate and has good reproducibility (29). To the best of our knowledge, we found no study in patients with mild SBD comparing the use of a catheter with piezoresistor pressure sensor to the use of the classical rubber balloon sensor. Nevertheless, several sleep centres use similar solid-state sensor catheter (26,30,31), as well as, medical centres using ambulatory oesophageal manometry (29).

Our study of patients with excessive daytime sleepiness and a suspicion of mild SDB, only about half had SDB by PSG. This agrees with Rowley's report (32) that a clinician's subjective impression or the use of prediction models is not sufficiently accurate to discriminate between patients with or without SDB.

We conclude that the combination of oesophageal pressure, respiratory flow and movement, and oximetry for the diagnosis of mild SDB is not cost-effective because of its poor diagnostic accuracy. Our results emphasise the role of overnight polysomnography in such diagnosis. New devices with alternative means of predicting arousal and respiratory effort variation, like pulse transit time and nasal pressure measurement, should be evaluated for cost-effectiveness in screening for mild SDB.

REFERENCES


