J. Sleep Res. (2012) Regular Research Paper

Added value of a mandible movement automated analysis in the screening of obstructive sleep apnea

GISELE MAURY[1], LAURENT CAMBRON[2], JACQUES JAMART[3], ERIC MARCHAND[1], FRÉDÉRIC SENNY[4] and ROBERT POIRRIER[2]

[1]Department of Pneumology, CHU Mont Godinne, Université catholique de Louvain, Yvoir, Belgium, [2]Department of Neurology, Faculty of Medicine, B35 ULg (Sart-Tilman), Liège, Belgium, [3]Scientific Support Unit, CHU Mont Godinne, Université catholique de Louvain, Yvoir, Belgium and [4]Department of Electronic, Montefiore Institute, B28 ULg (Sart-Tilman), Liège, Belgium

Keywords
automated analysis, hypopnea, mandible, salient mandible movement, screening method, sleep apnea syndrome

Correspondence
Gisèle Maury, MD, Pneumology, CHU Mont Godinne, Université catholique de Louvain, Yvoir, Belgium, 5530.
Tel.: 003281423351;
fax: 003281423352;
e-mail: gisele.maury@uclouvain.be

Accepted in revised form 15 June 2012; received 20 January 2012

DOI: 10.1111/j.1365-2869.2012.01035.x

SUMMARY

In-laboratory polysomnography is the 'gold standard' for diagnosing obstructive sleep apnea syndrome, but is time consuming and costly, with long waiting lists in many sleep laboratories. Therefore, the search for alternative methods to detect respiratory events is growing. In this prospective study, we compared attended polysomnography with two other methods, with or without mandible movement automated analysis provided by a distance-meter and added to airflow and oxygen saturation analysis for the detection of respiratory events. The mandible movement automated analysis allows for the detection of salient mandible movement, which is a surrogate for arousal. All parameters were recorded simultaneously in 570 consecutive patients (M/F: 381/189; age: 50 ± 14 years; body mass index: 29 ± 7 kg m^{-2}) visiting a sleep laboratory. The most frequent main diagnoses were: obstructive sleep apnea (344; 60%); insomnia/anxiety/depression (75; 13%); and upper airway resistance syndrome (25; 4%). The correlation between polysomnography and the method with mandible movement automated analysis was excellent (r: 0.95; $P < 0.001$). Accuracy characteristics of the methods showed a statistical improvement in sensitivity and negative predictive value with the addition of mandible movement automated analysis. This was true for different diagnostic thresholds of obstructive sleep severity, with an excellent efficiency for moderate to severe index (apnea–hypopnea index ≥15 h^{-1}). A Bland & Altman plot corroborated the analysis. The addition of mandible movement automated analysis significantly improves the respiratory index calculation accuracy compared with an airflow and oxygen saturation analysis. This is an attractive method for the screening of obstructive sleep apnea syndrome, increasing the ability to detect hypopnea thanks to the salient mandible movement as a marker of arousals.

INTRODUCTION

Obstructive sleep apnea (OSA) syndrome is a frequent disease associated with multiple co-morbidities. Assessing the severity of the disease relies on the calculation of the apnea–hypopnea index (AHI), which is the number of apneas and hypopneas per hour of sleep time. In the Wisconsin Cohort Study related by Young *et al.* (1993), the prevalence of symptomatic sleep apnea (AHI ≥5 h^{-1} with excessive daytime sleepiness) for men and women was 4 and 2%, respectively. Untreated, OSA has important health and socioeconomic consequences, but efficient therapies are available and consequently its diagnosis is important. However, a large number of patients with OSA remain undiagnosed. Currently, polysomnography (PSG) is the golden standard test for diagnosing OSA. PSG has important drawbacks, including the need for a full night in a sleep laboratory with cumbersome sensors under the supervision of a competent technician, being labor intensive and time consuming. Moreover, access

to PSG is limited and, hence, diagnosis and treatment of OSA are often delayed.

Several less elaborated sleep-monitoring equipments have been developed for the diagnosis of OSA. These are categorized into four types, based on the American Sleep Disorder Association (ASDA) classification (Ferber *et al.*, 1994). The validity of these tests remains disputed because of concerns regarding the lack of rigorous evaluation: lack of face to face comparison with the gold standard, selected and poorly described populations, etc,…(Flemons *et al.*, 2003; Flemons and Littner, 2003). Due to the lack of sleep evaluation, type 3 and 4 equipments allow the calculation of a respiratory disturbance index (RDI; defined as the number of apneas and hypopneas per hour of recording time between lights out and lights on) – which differs from the AHI, where the denominator is the total sleep time (TST).

Respiratory sleep disorders are thought to be associated with typical mouth opening and closing movements. In healthy adults without OSA, Miyamoto *et al.* (1998)described mouth opening that was less than 5 mm for 88.9% TST. Increased mouth opening during obstructive apnea was described in 1991 by Hollowell and Suratt (1991), and later by Miyamoto *et al.* (1999). Accordingly, mouth opening is greater than 5 mm for >69.3% TST in patients with OSA. Miyamoto *et al.* (1998, 1999) have described a common pattern consisting of a gradual opening and a quick closure of the mouth, mostly following an arousal response in normal patients and patients with OSA.

The quick closure of the mouth is hereunder called salient mandible movement (SMM), which could also start as a lowering of the mandible. Sjöholm *et al.* (2000) studied a small group of 21 patients with OSA, and analysed masseter (masticatory muscle and elevator of the mandible) contraction episodes. They found that masseter contraction episodes were associated with the termination of apnea or hypopnea. We postulated that mandible movement automated analysis (MMAA) could improve accuracy in respiratory event recognition, mainly through a better detection of hypopnea thanks to the associated SMM, as described by Senny *et al.* (2008). According to the definition used for hypopnea in this study, SMM is considered as an arousal marker.

MATERIALS AND METHODS

Setting

The setting was the sleep laboratory of the CHU Liège. This laboratory is linked to a neurological unit.

Participants

We prospectively included 629 consecutive subjects referred for diagnostic PSG who underwent recording in one of the rooms of the sleep laboratory of the CHU Liège between January 2006 and March 2010. There was no intentional selection for assignment of the patients in this particular room. The only exclusion criterion was ongoing continuous positive airway pressure treatment. The test has been recorded with additional oxygen (nasal canula) in one patient. The experiments were conducted in accordance with the principles of Declaration of Helsinki for Human Experimentation and after approval of Ethics Review Board; the patients were informed of the aim of the study and gave consent.

Recordings

We compared attended full PSG with two methods combining nasal airflow (NAF) and pulse oximetry (SpO_2) recordings, with or without MMAA.

All subjects underwent attended PSG (S7000 or N7000 polysomnographs; EMBLA Medcare, Denver, USA) and JawSens® (Nomics, Liège, Belgium) recording. The PSG included the following neurophysiology signals: a three-channel electroencephalography (EEG; C3-A2, C4-A1, FZ-CZ), left and right electrooculography, submental electromyography (chin EMG) for sleep staging and arousal scoring, two (left and right) tibial EMG for periodic leg movement evaluation; cardiorespiratory signals, including electrocardiogram, nasal cannula/pressure transducer (NAF; Protech, Mukilteo, USA), chest and abdominal inductance plethysmography belts, plethypulse, SpO_2 (Nonin, Plymouth, USA), snoring detection (piezoelectric sensor from EMBLA), and body position (body position sensor Protech) recordings. Oximetry sampling was done at 2 Hz.

MMAA was performed by means of a distance-meter called JawSens® (Jaw sensor, the recorded signal is called 'Jawac' for jaw activity) based on the principle of magnetometry. The sensors were composed of two coils and capacitors, each embedded in a small cylinder (7 mm diameter; 25 mm main axis). They were disposed, parallel to each other, perpendicular to the midline of the face, and fixed with plasters, one in the dimple above the chin and the other on the forehead. They were connected to an electronic circuit by two cables. The electronic circuit converted distance into voltage. The signal was digitalized with a sampling frequency of 10 Hz. Physical calibration was done by asking the patient to first close his/her mouth and then to open it fully.

Analysis

The number of respiratory events (apneas and hypopneas) per hour was calculated using as denominator the time between lights out and lights on, called the total dark time (TDT) for the two methods, and TST [total non-rapid eye movement (NREM) stages 1–4 and rapid eye movement (REM) sleep] providing AHI for PSG. The primary measures were the RDI for the two automated methods that were compared with the visual analysis for PSG. The first RDI was defined as respiratory events provided by the NAF and SpO_2 analysis divided by the TDT. The second RDI was called SMM-RDI, and was defined as the number of respiratory

events recognized by the NAF, SpO_2 and MMAA analysis divided by the TDT. Recordings of PSG and the two alternative methods were performed simultaneously, but on separate acquisition channels in each patient.

All sleep studies (PSG) were manually analysed either by a competent certified PSG technician, or two sleep physicians that were blind to the mandibular movement signal. The sleep analysis was performed by visual analysis of successive 30-s epochs. The sleep stages were defined according to the scoring rules of Rechtschaffen and Kales (1968). Arousals were reported as recommended by the ASDA (1992), and defined as an acceleration of the EEG frequencies in NREM sleep, associated with an increase of chin EMG amplitude in REM sleep, occurring for at least 3 s with at least 10 s of stable sleep preceding the change. An apnea was defined as a cessation of airflow amplitude, below 20% of the reference value, for 10 s or more.

For the three related methods, a hypopnea was scored each time a reduction of airflow (10 s or more) occurred below 70% of the reference value, provided that it was associated with oxygen desaturation ≥4%. Furthermore, a reduction of flow as defined above associated with a cortical arousal was also defined as a hypopnea for PSG recordings, and a reduction of airflow as defined above was defined as a hypopnea when associated with a SMM on automated analysis.

To perform the statistical analysis, we selected three cut-offs for the AHI. An abnormal AHI was considered when ≥5 h^{-1}. Moderate to severe and severe OSA corresponded, respectively, to AHI ≥15 h^{-1} and ≥30 h^{-1}. An upper airway resistance syndrome (UARS) was diagnosed if a clear drop in inspiratory airflow (remaining above 70% of the reference value) occurred concurrently with increased respiratory effort on the belts, with associated respiratory effort-related arousals (RERAs; a brief change in sleep state – arousal) as proposed in the International Classification of Sleep Disorders (American Academy of Sleep Medicine, 2005). UARS was diagnosed if >15 RERAs h^{-1} and/or >20% TST with flattening of the inspiratory flow curve were recorded, associated to symptoms.

Apnea, hypopnea or central and obstructive respiratory events were not differentiated for the present study.

The number of patients with OSA syndrome (clinical diagnosis) mentioned in Table 1 'main final diagnoses' and allowing population characterization was not exactly the same as sleep apnea (pure respiratory event analysis, including AHI ≥5 h^{-1} disregarding the symptoms).

The main relevant diagnosis was determined based on the complaints and history of the patient, and established by the specialists in charge of the final results (Table 1).

Signal processing of the mandible movement signal for the recognition of sleep-disordered breathing comprises the following stages. First, NAF and SpO_2 were automatically analysed based on the above-mentioned definition. SMMs were labeled as respiratory arousals if they were associated with a reduction of airflow (10 s or more) below 70% of the reference value in order to respect the hypopnea definition used. SMMs were identified by the continuous wavelet transform, as described by Senny *et al.* (2008). Each mandibular movement was characterized by a value that depends on its amplitude and its sharpness (the more discontinuous or sharp the movement, the greater the value). A SMM was considered as a relevant arousal for hypopnea definition if its maximal value was greater or equal to 1 (which corresponds to a movement of about 1 mm), and twice the average of the movement values during the related event. Based on this sequential analysis, considered SMMs were always associated with a hypopnea not associated with oxygen desaturation. The SMM-RDI was defined as the total apneas and hypopneas recorded by the automated analysis of NAF, SpO_2 and the MMAA divided by the time (expressed in hours) between lights out and lights on.

These automated analyses (RDI and SMM-RDI) were performed randomly and collected by an engineer that was blind to the results of the PSG.

In order to avoid a bias linked to technical divergences of sensors, NAF and SpO_2 signals were common for PSG and the two automated methods. Each patient had one NAF and one digital SpO_2. NAF and SpO_2 signals were electronically derived to feed both the automated analysis and the PSG.

In order to assess how SMM was linked to cortical arousals associated with hypopnea not defined by the presence of desaturation, we performed a *post hoc* analysis. For each patient, the number of hypopneas associated with arousal (visual analysis) but without any arterial oxygen desaturation was measured and compared with the number of hypopneas (without desaturation) associated with SMM, detected by the automatic system. The only requirement was that the SMM

Table 1 Final clinical diagnoses

Main Final diagnosis	*Main*
OSA	344 (60)
Insomnia/anxiety/depression	75 (13)
UARS/RERA	25 (4.4)
Circadian rhythm sleep disorder	20 (3.5)
Restless legs syndrome	17 (3)
Periodic limb movement syndrome	14 (2.5)
Inadequate sleep hygiene, caffeine	13 (2.3)
Hypersomnia	7 (1.3)
Narcolepsy	2 (0.35)
Primary snoring	7 (1.3)
REM sleep disorder	2 (0.35)
Unknown	24 (4.2)
Normal	10 (1.8)
Other medical disease	8 (1.4)

OSA, obstructive sleep apnea; REM, rapid eye movement; UARS/RERA, upper airway resistance syndrome/respiratory effort-related arousal. Results are provided in *n* (%). Other medical diseases known before sleep investigation and maintained as final diagnosis: fibromyalgia, pain, gastroesophageal reflux.

had to be in a time interval of −5 and +15 s of the end of the respiratory event detected by the NAF.

The PSG scorer was blind to jaw activity. The SMM-RDI analysis was fully automated. The number of hypopnea linked to arousal or SMM detected was divided by the TST to calculate a hypopnea index (HI).

Statistical Analysis

Sensitivity (Se) and specificity (Sp), positive and negative predictive value (PPV and NPV), and likelihood ratio (LR) compared with the reference (PSG) were determined for patients with mild (AHI $\geq 5\ h^{-1}$), moderate to severe (AHI $\geq 15\ h^{-1}$) and severe (AHI $\geq 30\ h^{-1}$) sleep apnea. Exact confidence intervals were computed by *F*-distribution. Se and Sp were compared by binomial test.

Paired Student's *t*-test was applied for comparison of the difference between the two methods compared with standard diagnosis.

Linear regression and Pearson correlation between the two indexes were also calculated on the entire studied population for assessing association between parameters. Correlations were compared by the Hotelling test, described by Cohen (1989).

Bland & Altman graphs were performed with MedCalc (Mariakerke, Belgium) on the entire studied population. For the Bland–Altman plot, each data point represents the difference between the two measurements (AHI − RDI) plotted against the mean for each patient ([AHI + RDI]/2). Mean difference and limits of agreement (Bland & Altman) are reported, where the limits of agreement have been defined as ±1.96 SD of the difference. The same was done with SMM-RDI.

All statistical tests are two-tailed. Statistical significance was set at $P < 0.05$.

Analyses were done with NCSS version 2007 (Kaysville, Utah, USA) and SPSS 15.0 (SPSS, Chicago, IL, USA).

RESULTS

Data were collected in 629 patients visiting the sleep laboratory, and 570 were considered for the analysis. Fifty-nine patients were excluded from the analysis. Lost data occurred in 22 patients due to technical problems: loss of NAF signal ($n = 4$); loss of jaw activity signal during the night ($n = 4$); no jaw activity signal ($n = 12$); and loss of SpO_2 ($n = 2$). This means that 3.5% of the recordings were excluded due to technical problem. Thirty-seven patients had a history of OSA diagnosed by PSG. To avoid impact on predictive parameters, they were excluded from this analysis. The final study population (570 patients; M/F: 381/189) had a mean age: 50 ± 14 years; min: 13 years; max: 85 years; body mass index (BMI): 29 ± 7 kg m^{-2}; min: 14 kg m^{-2}; max: 57 kg m^{-2}. The main complaints are listed in Table 2; in one-third of the cases, these complaints were snoring and/or sleepiness, sometimes in association. The mean Epworth Sleepiness Score was 11 ± 7/24.

Table 2 Patients' characteristics

N	570
M/F [*n* (%)]	381 (67)/189 (33)
Age (years)	50 ± 14
BMI (kg m^{-2})	29 ± 7
Complaints	
Snoring	204 (36)
Sleepiness	175 (31)
Apnea (witness)	111 (19.5)
Tiredness	100 (17.5)
Lack of sleep	53 (9)
Not refreshing sleep	24 (4)
Frequent awakenings	12 (2)
Gasping	9 (1.6)
Epworth Sleepiness Scale	11 ± 7
Shift work	14 (2.5)
Known smoking habits (Current/former/non-smoker/unknown)	134 (23.5)/302 (53) /103 (18)/31 (5.4)
Current BzD use	90 (16)
HTA	192 (66.7)
Chronic respiratory disease:	58 (10)
Asthma/COPD/others*	25/24/4
Cardiovascular disease (cardiopathy and peripheral vascular disease)	64 (11.3)
Psychiatric disorder	94 (16.5)
Neurological disease:	113 (19.8)
Ischemic event/MS/ epilepsy/others†	21/7/9/76
Transplantation (solid organ)/RA	4 (0.7)/3 (0.5)
Nose and throat disease	111 (19.5)
Uvulo-palatal surgery	15 (2.6)
Surgical mandible treatment	9 (1.6)

BMI, body mass index; BzD, benzodiazepine; COPD, chronic obstructive pulmonary disease; HTA, hypertension; MS: multiple sclerosis; RA, rheumatoid arthritis. *n* (%); M, male; F, female. *Bronchiectasis, pulmonary fibrosis, thoraco-pulmonary sequellae. †Muscular dystrophy, traumatic injury, headache. Complaints are those related by the patients, sometimes in association, which is not mentioned.

The most frequent final diagnoses were OSA (344; 60%), insomnia/anxiety/depression (75; 13%) and UARS (25; 4%) (Table 1). These were mostly dyssomnia, as described in the International Classification of Sleep Disorders (American Academy of Sleep Medicine, 2005).

The mean total analysed period was 553.6 ± 89.7 min for both RDI calculation, and 419 ± 98.4 for the AHI (TST).

Using the AHI cut-off of 5 h^{-1} combined with symptoms, OSA syndrome was the main diagnosis in 60% ($n = 344$) of this population. For the whole studied population, the mean PSG AHI was 24.3 ± 23.3 events h^{-1}, while the mean SMM-RDI was 19.2 ± 18.6 events h^{-1} and 14.3 ± 18.3 for RDI. Simple regression analysis was used to study the correlation between AHI and the two RDI. The relation was linear and direct. Correlations were strong between PSG AHI and

SMM-RDI (r: 0.95) accounting for 90% variance ($P < 0.001$), and between PSG AHI and RDI (r: 0.93; $P < 0.001$; Figs 1 and 2). Correlations were significantly different ($P < 0.001$).

A Bland & Altman plot corroborated the analysis. The mean difference between AHI and SMM-RDI was 5.1 ± 8 ($P < 0.001$) versus 10 ± 9.23 ($P < 0.001$) for RDI. With MMAA, agreement was good with few points outside the limits of agreement (Figs 3 and 4). Limits of agreement were [−8.1; 28.1] for AHI-RDI and [−10.7; 20.9] for AHI-SMM-RDI.

Respectively, 462 (81%), 307 (54%) and 169 (30%) patients had an AHI ≥5 h^{-1}, ≥15 h^{-1} and ≥30 h^{-1}. Accuracy characteristics of the method with or without MMAA for the different AHI cut-offs are described in Table 3. We found a statistically significant improvement in Se and NPV with the addition of MMAA (SMM-RDI) for each AHI cut-off: ≥5 h^{-1} (93 and 72%), ≥15 h^{-1} (75 and 77%), compared with RDI with the same thresholds: ≥5 h^{-1} (69 and 43%), ≥15 h^{-1} (54 and 65%). There was no significant loss of PPV, and the values of Sp for moderate to severe and severe AHI were excellent. For severe patients (≥30 h^{-1}), Se and NPV improved by adding MMAA (respectively, 52 and 83% for RDI; 68 and 88% for SMM-RDI). Although not formally significant, there was an improvement in NPV in this severe group.

Positive and negative LRs were comparable for the two methods for moderate to severe index (AHI ≥15 h^{-1}).

In order to assess how hypopneas related to arousal are the hypopneas linked to SMM, we performed a *post hoc* analysis on the entire population. The results showed that 62.7% of hypopneas associated with cortical arousals but without desaturation on PSG were associated with a SMM. The HIs calculated for PSG and for MMAA added to airflow analysis were, respectively, 6.0 ± 6.0 h^{-1} and 3.8 ± 4.2 h^{-1}. The correlation coefficient between these two indices was excellent (r: 0.92). The mean difference from Bland & Altman plots was 2.2 ± 2.7 h^{-1}.

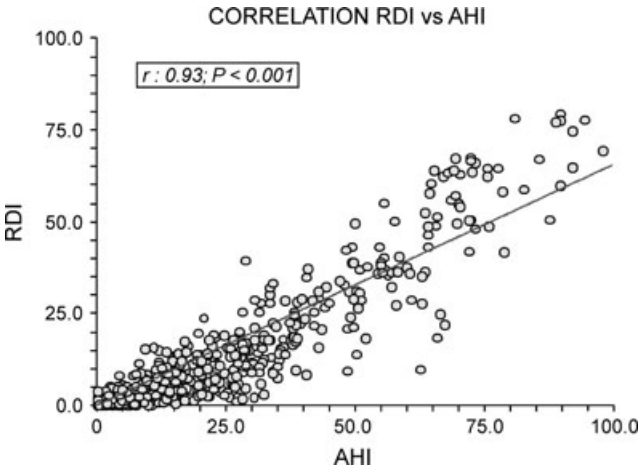

Figure 1. Correlation graph between RDI and AHI. AHI, apnea–hypopnea index (PSG); RDI, respiratory disturbance index (NAF and SpO_2).

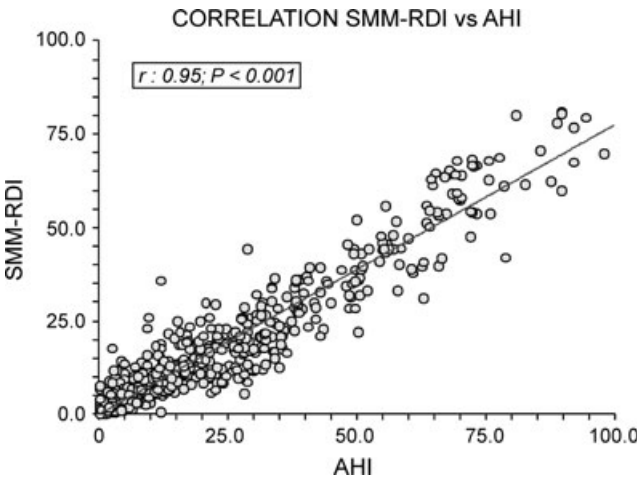

Figure 2. Correlation graph between SMM-RDI and AHI. AHI, apnea–hypopnea index (PSG); SMM-RDI, respiratory disturbance index based on NAF, SpO_2 and MMAA. The estimated model for regression is (0.8) + (0.76 × PSG AHI) for correlation between AHI and SMM-RDI.

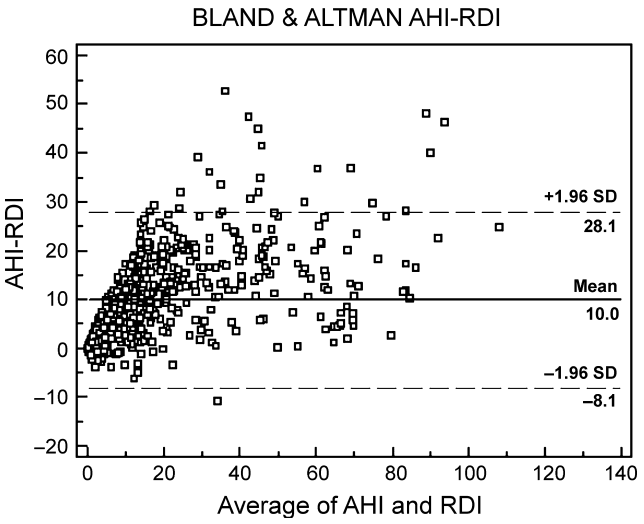

Figure 3. Bland & Altman graph for AHI and RDI. AHI, apnea–hypopnea index (PSG); RDI, respiratory disturbance index (NAF and SpO_2).

DISCUSSION

In the present study, we found that adding MMAA to NAF and SpO_2 automated analysis allowed a significant increase in Se and NPV for the diagnosis of moderate to severe OSA (AHI > 15 h^{-1}) without a significant drop in PPV.

Globally, correlation was excellent between AHI and SMM-RDI. There was a bias underestimating respiratory events index (Bland & Altman mean difference is 5.1 ± 8). This bias was significantly smaller than with RDI ($P < 0.001$). As seen on the Bland & Altman graphs (Figs 3 and 4), there was a linear trend between the bias (AHI-RDI; for positive and

Table 3 Se, Sp, PPV, NPV and LR derived from the two methods and compared with simultaneous PSG

	≥5 h^{-1} N = *462 (81%)*			≥15 h^{-1} N = *307 (54%)*			≥30 h^{-1} N = *169 (30%)*		
AHI	*RDI*	*SMM-RDI*	P	*RDI*	*SMM-RDI*	P	*RDI*	*SMM-RDI*	P
Se	69 [65.13; 73.74]	93 [90.57; 95.38]	<0.001	54 [48.64; 60.06]	75 [70.02; 79.97]	<0.001	52 [44.27; 59.81]	68 [60.45; 75.00]	<0.001
Sp	96 [90.95; 99.00]	72 [62.44; 79.99]	<0.001	99 [96.7; 99.76]	95 [92.16; 97.62]	<0.05	99.7 [98.62; 99.99]	99 [97.83; 99.85]	NS
PPV	99 [96.87; 99.66]	93 [90.57; 95.38]		98 [94.93; 99.63]	95 [91.53; 97.42]		99 [93.90; 99.97]	97 [92.75; 99.47]	
NPV	43 [36.92; 49.53]	72 [62.44; 79.99]		65 [60.10; 69.67]	77 [71.79; 81.23]		83 [79.51; 86.40]	88 [84.70; 90.89]	
LR +	19.13 [7.30; 50.16]	3.32 [2.47; 4.54]		47.69 [15.41; 147]	16.49 [9.45; 28.77]		208 [29.33; 1486]	90.96 [29.32; 282.16]	
LR −	0.32 [0.27; 0.36]	0.1 [0.07; 0.137]		0.46 [0.41; 0.52]	0.26 [0.21; 0.32]		0.48 [0.41; 0.56]	0.32 [0.26; 0.40]	

AHI, apnea–hypopnea index (PSG); LR+, positive likelihood ratio; LR−, negative likelihood ratio; NPV, negative predictive value; NS, not significant; PPV, positive predictive value; RDI, respiratory disturbance index (NAF and SpO_2); Se, sensitivity; SMM-RDI: respiratory disturbance index based on NAF, SpO_2 and MMAA; Sp, specificity. Exact confidence intervals are based on *F*-distribution.

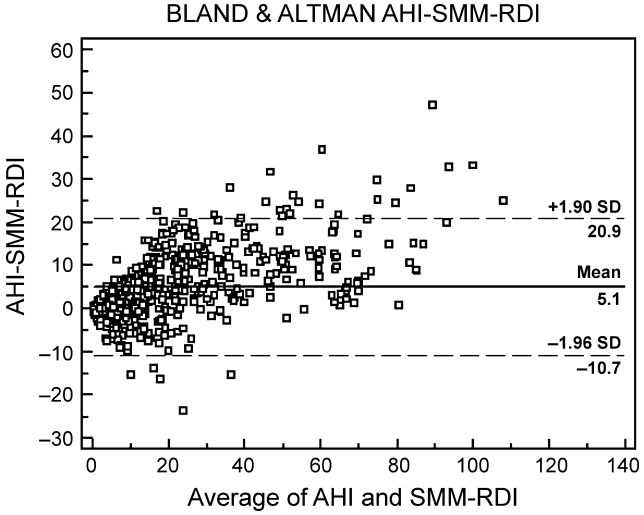

Figure 4. Bland & Altman graph for AHI and SMM-RDI. AHI, apnea–hypopnea index (PSG); SMM-RDI, respiratory disturbance index based on NAF, SpO_2 and MMAA.

negative differences) and AHI below 30 h^{-1} for both methods. Another large cohort related by Masa *et al.* (2011) showed the same trends. Moreover, limits of agreement are narrower for SMM-RDI compared with RDI, showing more accuracy in the concordance.

Thanks to the SMM, used as an arousal marker, the accuracy in respiratory events detection improved by increasing the ability to detect hypopnea.

For defining hypopneas in the present study, we used the criteria of Meoli *et al.* (2001), which are currently the approved hypopnea definition for the Centers for Medicare and Medicaid Services in the USA (2008) but also considering arousals. A hypopnea occurred if an airflow reduction occurred with either a desaturation or an arousal. A *post hoc* analysis between all the hypopneas with cortical arousals (without desaturation) and all the hypopneas with SMM (without desaturation) showed that 62.7% of this type of hypopnea could be automatically detected using MMAA. The fact that only 62.7% were recognized is probably due to the automatic thresholds that were imposed. This mainly concerns the minimal SMM amplitude for the automatic recognition. A SMM was considered as a relevant arousal for hypopnea definition if its maximal value was greater or equal to 1 (which corresponds to a movement of about 1 mm) and twice the average of the movement values during the related event.

The two components of the SMM definition are important for the recognition of SMM as a marker of respiratory arousal. This explains why the HI derived from MMAA is strongly correlated with HI derived from PSG (*r*: 0.92).

As shown by Ruehland *et al.* (2009), all differences in hypopnea criteria would induce variable final AHI, leading to misclassified OSA. By using SMM to detect arousals included in our hypopnea definition, MMAA significantly improved predictive parameters, suggesting that the combination of SMM with SpO_2 and NAF analysis in a

simple device could be an attractive solution for the screening of OSA.

Lost data occurred in 22 (3.5%) patients, which is acceptable. This was due to variable technical problems detailed in the Results.

The difference in respiratory events index could be explained by several factors. First, there is variability in scoring respiratory events between automatic and manual (human perception) analysis. This is the case for very short events, but considered by the reviewer as sufficient to reach 10 s, or for very long events (mainly hypopnea) considered as one or multiple events. Moreover, a drop greater than 30% in airflow to detect hypopnea is not so easy to assess for human eyes, but automatic analysis will always measure it and stick to the definition. This could explain a part of the variability. Second, a comparison was made between AHI and RDI: the analysed period is different (mean difference RDI-AHI in the present study: 134 min). The denominator used for RDI and SMM-RDI is the TDT, which includes sleep latency; wake periods during the night; and short onset awakening before lights on. Different denominators explain variability in the respiratory events index. This is an important point, explaining higher AHI than RDI and SMM-RDI.

Third, the maximal value of MMAA must reach >1 and be twice the average of the movement values during the related event to be considered as significant SMM. Some body position could influence the amplitude of the mandible movement and affect hypopnea recognition with MMAA, whilst respiratory arousals are more consistently identified on hypnogram by the PSG scorer.

It is important to recognize a severe disease that is accessible to an efficient and easy treatment. Therefore, the improvement in Se by the added MMAA is interesting. Moreover, in this population (AHI > 30 h^{-1}), positive LRs were excellent for both studied methods with an overlap in confidence intervals. Despite this improvement by adding MMAA, Se was lower than 80%. This lack of Se could be overcome by associating an anthropomorphic questionnaire with high Se. The high number of false-positive cases associated with this first step could potentially be balanced by the high Sp (99%) of the portable monitoring (PM) with added MMAA. Moreover, the selection of patients with sleep apnea in the recruitment process, frequently occurring in the sleep laboratory, would improve the pre-test probability of having OSA. Compared with other studies with PM, moderate to severe sleep apnea prevalence is not very high in our population (AHI ≥15 h^{-1}: n = 307; 54%), which could have interacted with predictive values, a higher prevalence leading to higher percentage of positive test.

The present study has several limitations. First, the software version used did not comprise a sleep/wake assessment as it was not available at the time of the recordings.

The important difference between TDT and TST is probably due to the fact that our population is composed of patients with complaints of bad sleep and submitted to 'first night effect'. In the future, MMAA has the potential to be used as an actigraphy, and to give an indirect sleep/wake assessment as described by Senny *et al.* (2011). Another prospective and multicenter study would give more powerful data to the PM, showing a wider spectrum of analysis.

Second, the addition of the MMAA to a simple PM resulted in a poorer Sp, albeit an improved Se when considering a cut-off of AHI >5 h^{-1}. Therefore, as recommended in guidelines on portable monitoring (Flemons and Littner, 2003; Flemons *et al.*, 2003), screening device has to be confronted to the pre-test clinical complaints (pre-test probabilities of having the disease).

A third limitation is the restricted in-laboratory assessment, particularly for a screening method dedicated to in-home use.

The population of the study included nine adolescents (between 13 and 16 years old). They were referred to the sleep study to exclude narcolepsy. Sleep and respiratory events definitions used were identical as those for the entire population. None of them was outliers of the general Bland & Altman plots.

The occurrence of abnormal repetitive mandible movement such as bruxism, chewing, has not been studied, but seemed to occur rather rarely in our sample. It was registered in four patients (0.7%), as based on the visual PSG scoring. In our study, this was not the subject and we did not process the analysis by searching bruxism. Nevertheless, bruxism has another pattern to SMM, occurs with closed mouth and is repetitive. It has been described in association with sleep respiratory disorders by Sjöholm *et al.* (2000) with a high incidence which is questioning for us. We think that masseter contraction should not be regarded as bruxism, unless rhythmic and repetitive. This kind of behavior does not correspond to the definition of SMM used in the automated analysis.

CONCLUSION

In conclusion, in a population with a high prevalence of respiratory disorder, the addition of MMAA significantly improves the accuracy for the detection of respiratory events with a mean underestimation RDI of 5 h^{-1}. This portable monitoring is an interesting screening method for moderate to severe sleep apnea. Even if, in this study, the method is less accurate for very low AHI, this is a promising way to detect severe disease requiring a treatment that is accessible.

CONFLICT OF INTEREST

For each author, there is no conflict of interests.

REFERENCES

American Academy of Sleep Medicine. *The International Classification of Sleep Disorders Revised – Diagnostic and Coding Manual* (2nd edn.) American Academy of Sleep Medicine, Westchester, IL, 2005.

ASDA (American Sleep Disorders Association, Sleep Research Society) Atlas Task Force. EEG arousals: scoring rules and examples. *Sleep*, 1992, 15: 173–184.

Centers for Medicare and Medicaid Services. National coverage determination for continuous positive airway pressure (CPAP) therapy for obstructive sleep apnea (OSA). NCD#240.4.2005 (cited 22 September 2008); available from: http://www.cms.hhs.gov, 2008.

Cohen, A. Comparison of correlated correlations. *Stat. Med.*, 1989, 8: 1485–1495.

Ferber, R. A., Millman, R. P., Coppola, M. P. *et al.* Portable monitoring in the assessment of obstructive sleep apnea. ASDA standards of practice. *Sleep*, 1994, 17: 378–392.

Flemons, W. W. and Littner, M. R. Measuring agreement between diagnostic devices. *Chest*, 2003, 124: 1535–1542.

Flemons, W. W., Littner, M. R., Rowley, J. A. *et al.* Home diagnosis of sleep apnea: a systematic review of the literature. *Chest*, 2003, 124: 1543–1579.

Hollowell, D. E. and Suratt, P. M. Mandible position and activation of submental and masseter muscles during sleep. *J. Appl. Physiol.*, 1991, 71: 2267–2273.

Masa, J. F., Corral, J., Pereira, R. *et al.* Effectiveness of home respiratory polygraphy for the diagnosis of sleep apnoea and hypopnoea syndrome. *Thorax*, 2011, 66: 567–573.

Meoli, A. L., Casey, K. R., Clark, R. W. *et al.* Clinical Practice Review Committee. Hypopnea in sleep disordered breathing in adults. *Sleep*, 2001, 24: 469–470.

Miyamoto, K., Özbek, M. M., Lowe, A. A. *et al.* Mandibular posture during sleep in healthy adults. *Arch. Oral Biol.*, 1998, 43: 269–275.

Miyamoto, K., Özbek, M. M., Lowe, A. A. *et al.* Mandibular posture during sleep in patients with obstructive sleep apnoea. *Arch. Oral Biol.*, 1999, 44: 657–664.

Rechtschaffen, A. and Kales, A. *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*. National Institutes of Health, Washington, DC, 1968.

Ruehland, W. R., Rochford, P. R., O'Donoghue, F. J., Pierce, R. L., Singh, P. and Thornton, A. T. The new AASM criteria for scoring hypopneas: impact on the apnea hypopnea index. *Sleep*, 2009, 32: 150–157.

Senny, F., Destiné, J. and Poirrier, R. Midsagittal jaw movement analysis for the scoring of sleep apneas and hypopneas. *IEEE Trans. Biomed. Engng.*, 2008, 55: 87–95.

Senny, F., Maury, G., Cambron, L., Leroux, A., Destiné, J. and Poirrier, R. The sleep/wake state scoring from mandible movement signal. *Sleep Breath.*, 2011, 16: 535–542.

Sjöholm, T. T., Lowe, A. A., Miyamoto, K., Fleetham, J. A. and Ryan, C. F. Sleep bruxism in patients with sleep-disordered breathing. *Arch. Oral Biol.*, 2000, 45: 889–896.

Young, T., Palta, M., Dempsey, J., Skatrud, J., Weber, S. and Badr, S. The occurrence of sleep-disordered breathing among middle-aged adults. *N. Engl. J. Med.*, 1993, 328: 1230–1235.