Official Journal of the Asian Pacific Society of Respirology
Respirology

ORIGINAL ARTICLE

Mandibular position and movements: Suitability for diagnosis of sleep apnoea

JEAN-BENOIT MARTINOT,[1] JEAN-CHRISTIAN BOREL,[2,3] VALÉRIE CUTHBERT,[1] HERVÉ JEAN-PIERRE GUÉNARD,[4] STÉPHANE DENISON,[1] PHILIP E. SILKOFF,[5] DAVID GOZAL[6] AND JEAN-LOUIS PEPIN[3,7]

[1]*Laboratory of Sleep Medicine, CHU UCL Namur, Namur, Belgium,* [2]*'AGIR à dom' Association, La Tronche,* [3]*University Grenoble Alpes,* [7]*Sleep and Exercise Laboratory, Thorax and Vessels Division, Grenoble Alpes University Hospital, Grenoble,* [4]*Laboratory of Physiology, Bordeaux University, Bordeaux, France,* [5]*Department of Physiology, Temple University, Philadelphia, Pennsylvania, and* [6]*Department of Medicine, University of Chicago, Chicago, Illinois, USA*

ABSTRACT

Background and objective: **Mandibular movements (MMs) and position during sleep reflect respiratory efforts related to increases in upper airway resistance and micro-arousals. The study objective was to assess whether MM identifies sleep-disordered breathing (SDB) in patients with moderate to high pre-test probability.**

Methods: **This was a prospective study of 87 consecutive patients referred for an in-laboratory sleep test. Magnetometer-derived MM signals were incorporated into standard polysomnography (PSG). Respiratory events detected with MM analysis were compared with PSG for respiratory disturbance index (RDI) with a blinded scoring. All records were scored manually according to American Academy of Sleep Medicine rules. Primary outcome was to rule-in obstructive sleep apnoea syndrome (OSAS) defined as RDI cut-off value ≥5 or 15/h total sleep time (TST).**

Results: **High concordance emerged between MM and PSG-derived RDI with high temporal coincidence between events ($R^2 = 0.906$; $P < 0.001$). The mean diagnostic accuracy of MM for OSAS using RDI MM cut-off values of 5.9 and 13.5 was 0.935 (0.86–0.97) and 0.913 (0.84–0.95), with a mean positive likelihood ratio (LLR+) of 3.73 (2.7–20.4) and 8.46 (2.3–31.5), respectively. Receiver operating characteristic (ROC) curves at PSG cut-off values of 5 and 15/h TST had areas under the curve (AUC) of 0.96 (95% CI: 0.89–0.99) and 0.97 (95% CI: 0.91–0.99) ($P < 0.001$), respectively. MM analysis accurately identified SDB at different levels of severity.**

Conclusion: **RDI assessed by MM is highly concordant with PSG, suggesting a role of ambulatory MM recordings to screen for SDB in patients with moderate to high pre-test probability.**

SUMMARY AT A GLANCE

The study aims to score blindly the mandibular movements (MMs) compared with the respiratory disordered events during an in-laboratory polysomnography in patients with obstructive sleep apnoea. High agreement was found between both scorings. A simple recording of the MMs could be explored for home diagnosis of sleep-disordered breathing.

Key words: obstructive sleep apnoea syndrome, portable monitor device, respiratory effort, sleep mandibular movements.

Abbreviations: AASM, American Academy of Sleep Medicine; AHI, apnoea/hypopnoea index; AUC, area under the curve; BMI, body mass index; Co.Sen, coincidence sensitivity; Co.Spe, coincidence specificity; ESS, Epworth Sleepiness Scale; ICSD-3, International Classification of Sleep Disorders, Third Edition; LLR, likelihood ratio; MM, mandibular movement; MML, large MM; MMO, mouth opening MM; MMS, sudden MM; ODI, hourly rate of >3% oxygen desaturation; OSA, obstructive sleep apnoea; OSAS, OSA syndrome; PSG, polysomnography; RDI, respiratory disturbance index; RE, respiratory effort; RERA, RE-related arousal; ROC, receiver operating characteristic; SDB, sleep-disordered breathing; TST, total sleep time.

INTRODUCTION

Laboratory-based polysomnography (PSG) is the gold standard for diagnosing obstructive sleep apnoea (OSA) syndrome (OSAS). However, this method is expensive, time-consuming and cannot keep pace with demand.[1] In-laboratory testing is also inconvenient for patients, and does not assess their normal sleeping environment.

There are several unstandardized type 4 diagnostic devices consisting of one or two recording channels comprising oximetry, but their utility for diagnosing OSAS in different populations has yielded mixed results.[2] Indeed, nocturnal oximetry as a stand-alone signal is not currently recommended because the obstructive nature of identified events cannot be

Correspondence: Jean-Benoit Martinot, Centre du Sommeil et de la Vigilance, CHU UCL Namur Site Ste Elisabeth, 15, Place Louise Godin, 5000 Namur, Belgium. E-mail: martinot.j@scarlet.be

Received 3 June 2016; invited to revise 3 August 2016; revised 16 August 2016; accepted 16 August 2016 (Associate Editor: David Barnes).

documented and not all events cause oxygen desaturation.[3] Recently, the use of a single-channel recording of nasal flow was proposed as a first-line diagnostic for OSAS, despite the fact that not all recordings were interpretable.[4] Furthermore, nasal pressure measurements do not detect oral airflow, leading to overscoring of apnoeic events.[5] Finally, none of the type 4 devices accurately record sleep time, instead relying on total recording time. Consequently, these devices routinely over or underestimate the respiratory disturbance index (RDI, hourly rate of respiratory disturbances) compared with PSG, impacting therapeutic decisions.[6]

Recordings of mandibular position with a high-resolution magnetometer can accurately identify cortical arousals, respiratory effort (RE) and RE-related arousals (RERAs).[7,8] Additionally, analysis of mandibular movements (MMs) accurately estimates sleep duration and detects mouth opening as a surrogate for oral breathing.[9] Importantly, MMs are not influenced by head position, and thus the signal is reliably and consistently detected throughout sleep.[10]

We hypothesized that MM analysis would compare favourably to PSG. We therefore compared MM recordings to standardized in-laboratory PSG for the diagnosis of OSAS (ICSD-3, International Classification of Sleep Disorders, Third Edition) and assessed if MM analysis was able to determinate the severity of the disease based on apnoea/hypopnoea index (AHI, hourly rate of apnoea and hypopnoea) classification.

METHODS

Study design

This was a prospective study of consenting adult patients scheduled for a single overnight in-laboratory PSG. The patients eligible to participate were consecutive subjects, 18 years and older with symptoms suggestive of sleep-disordered breathing (SDB) undergoing a single PSG. The study was approved by the 'Comité d'Ethique Hospitalo-Facultaire-Universitaire de Liège' (*IRB-00004890-N°B707201523388*) and all participants provided informed consent.

Sample size calculation

At least 90 subjects were deemed necessary to obtain a significant value of 0.9 for both intraclass correlation coefficients (2.1) and area under the curve (AUC) of MM-RDI against the null-hypothesis, at $\alpha = 0.05$ and power = 0.8. After adjusting for a dropout rate of 10%, with a negative/positive ratio of 0.1, a total sample size of 100 subjects was planned for recruitment. To obtain a large enough number of true negatives (as some type 4 diagnostic devices have been reported to perform less in patients with mild or no SDB), 13 volunteers who had no specific sleep complaints were recruited by word of mouth.[11]

Study evaluations and measurements

Demographic characteristics, medical history, physical examination and a standardized survey of seven symptoms reported in ICSD-3 as being related to SDB using a Likert-scale with four frequency levels (0: never, 1: sometimes, 2: frequently and 3: every night) were obtained. The seven symptoms included nocturnal gasping and/or choking, witnessed apnoeas, morning headaches, non-refreshing sleep, daytime sleepiness, nocturia and reported habitual snoring. In addition, the Epworth excessive daytime sleepiness questionnaire scale (Epworth Sleepiness Scale, ESS) was administered.[12]

Routine laboratory-based PSG was recorded with a SomnoscreenPlus (Somnomedics, Randersacken, Germany). For further details, see Appendix S1 (Supplementary Information)

A mid-sagittal MM magnetic sensor (Brizzy, Nomics, Liege, Belgium) measured the distance in mm between two parallel, coupled, resonant circuits placed on the forehead and on the chin (Fig. 1). The transmitter generates a pulsed magnetic wave of low energy at 10 Hz. The change in the magnetic field is inversely related to the cube of the distance (d) between the chin and forehead probes. The probes were connected to an electronic module, and the distance was computed with a resolution of 0.1 mm before transmission to the PSG through a wired connection. The more negative the signal, the lower the mandibular position and the greater the mouth opening. Mandibular-derived variables are described in the Abbreviation list, and depicted in Figures 1 and S1 (Supplementary Information).

Scoring and quality control

Individual recordings were judged acceptable if >4 h duration, with flow, O_2 saturation and artefact-free MM data present for >90% of sleep time. Scoring for MM was performed by two blinded independent readers who had been trained to read MM tracings, while a different experienced reader analysed the standard PSG, after de-identification of records. The accuracy of MM scoring of each of the two MM readers was audited by a sleep specialist on a random sample of 15 records. American Academy of Sleep Medicine (AASM) rules were used to score all respiratory events and to calculate AHI and RDI (RDI = AHI + RERA index) as recommended by the ICSD-3 for OSAS diagnosis.[13,14]

Statistical analysis

Data were analysed using SAS 9.2 (SAS Institute Inc., Cary, NC, USA). The analysis focused on two main issues, which are described below.

Agreement and relationship between MM and PSG scorings

- Stepwise linear regression analysis was performed to evaluate the relationship between MM-RDI and PSG-RDI.
- The agreement between PSG-RDI and MM-RDI was evaluated using Bland–Altman plot[15] (based on absolute scoring) and Cohen's linear weighted kappa coefficient[16] (based upon the classification of patients into four severity groups).
- Inter-observer agreement for scoring MM and PSG was evaluated for all studies by intra-class

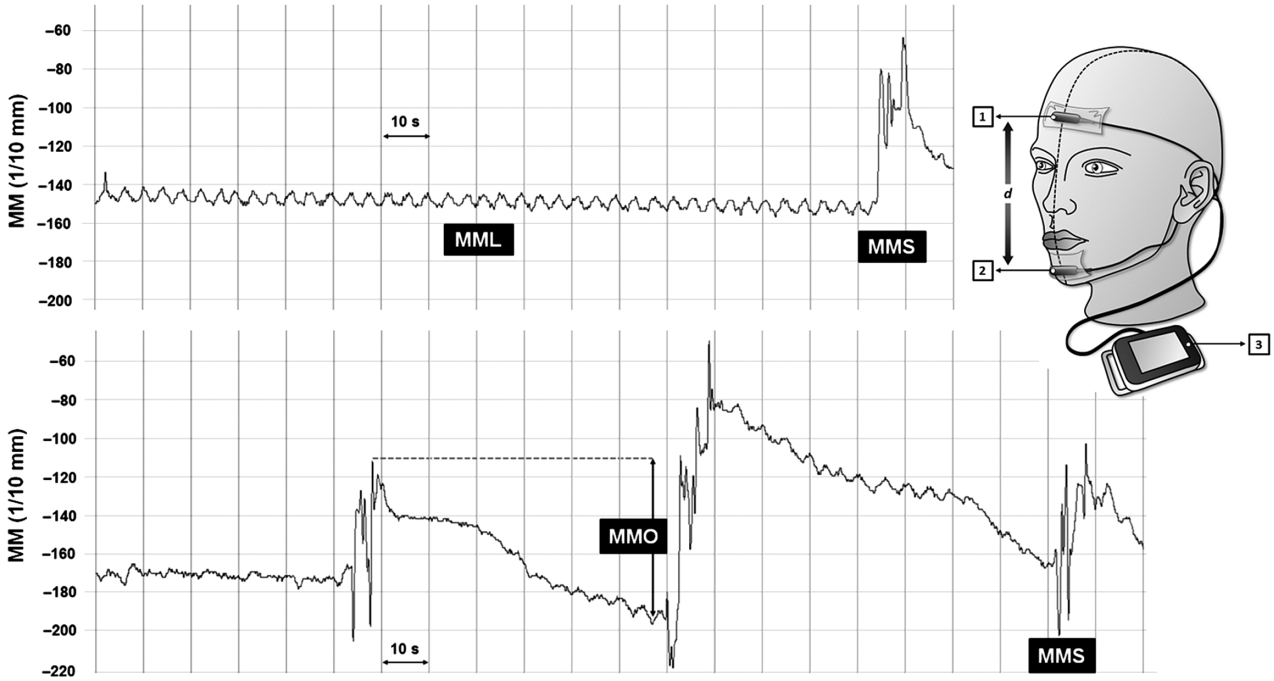

Figure 1 Classification of mandibular movements (MMs). Large MM (MML): periods of ≥10 s during which MM amplitudes were ≥0.3 mm from peak to peak were reported as MML. Sudden MM (MMS): a single sudden very MML (amplitude > 1 mm) during a respiratory cycle with a change in motion slope, disrupting the previous breathing frequency and the envelope around the previous peak to peak mandibular displacement. Note: MMS is closing the mouth accompanied or not with an MML and/or MMO to calculate MM-RDI (respiratory disturbance index). Mouth opening MM (MMO): periods of mouth opening > 0.3 mm with a duration of at least two respiratory cycles or at least 10 s, in the presence or absence of MML or MMS. 1, Forehead sensor; 2, chin sensor; 3, signal processing unit.

correlation coefficient[17] using two-way random model for single measure (ICC 2.1).

- Temporal coincidence between MM and PSG events was determined as follows: (i) the end of the PSG event was marked on the time line and (ii) a concurrent MM event was defined as that occurring within 15 s around the marked event. The results were described as positive when co-occurring, or negative when not. The sensitivity and the specificity of the coincidence (Co.Sen and Co.Spe, respectively) were determined and analysed as a function of the PSG AHI, with and without adjustment for total sleep time (TST).

Performance of MM scoring against the gold standard

The performance of MM-RDI for detecting OSAS patients based on two PSG-RDI thresholds was evaluated by receiver operating characteristic (ROC) curve analysis.[18] The outcome variable related to the diagnostic of the disease was based on a sensitivity/specificity analysis of MM device with the two different polysomnographic pre-specified cut-off values of RDI recommended in ICSD-3 (PSG-RDI ≥ 5 and ≥15/h TST). OSAS severity was evaluated from AHI, with <5, 5–15, 15–30 and >30/h TST representing the four severity categories. The post-test probability for each cut-off point was calculated as previously reported by Collop *et al.*[19]

Descriptive data are presented as mean ± SD and 95% CI. Statistical significance was defined at the 5% level.

RESULTS

Characteristics of study population

A total of 100 patients were enrolled over a period of 3 weeks. The performance of the integrated MM and PSG was acceptable with a failure rate <10% in keeping with the recommendations of Collop *et al.*[19] for an unattended PSG. Specifically, eight recordings were technically unacceptable, in four subjects due to failure to capture the MM signal into the PSG, three subjects due to poor oximetry recording and in one subject due to the loss of belts signal.

The final dataset included 79 patients with suspected OSAS and 13 healthy subjects. The characteristics of the participants are presented in Tables 1 and S1 (Supplementary Information). Patients were predominantly males, aged 18–80 years. The majority were either overweight or obese. Enrolled patients represented a typical referral population in which the pre-test probability (after bootstrapping) was 81.5% (95% CI: 73.9–88.0%) for an AHI > 5/h TST and 46.7% (95% CI: 38.0–55.4%) for an AHI > 15/h TST. The distribution of the clinical scores (0–3) among 79 OSAS suspected patients is presented for each of the seven symptoms in Figure S2 (Supplementary Information).

Table 1 Characteristics of the study population

Characteristics	Patients (n = 79)		Healthy controls (n = 13)	
	Mean ± SD	95% CI	Mean ± SD	95% CI
Age (years)	48.8 ± 14.6	45.5–51.7	27.4 ± 12.3	21.9–33.9
Weight (kg)	95.1 ± 25.8	89.8–100.9	57.5 ± 7.2	54.2–61.2
Height (cm)	172.6 ± 10.2	170.4–175.0	168.6 ± 6.9	165.5–171.9
BMI (kg/m^2)	31.8 ± 7.7	30.2–33.4	20.2 ± 2.1	19.0–21.6
ESS score	10.0 ± 5.9	8.6–11.2	4.7 ± 3.4	3.2–6.4
PSG-TST (min)	429.2 ± 9.9	408.9–449.1	363.3 ± 88.2	317.2–403.9
PSG-AHI (n/h)[†]	27.1 ± 27.2	21.2–33.2	3.9 ± 3.0	2.5–5.6
PSG-RERA (n/h)	8.4 ± 6.5	6.9–10.0	3.1 ± 2.7	1.7–4.7
PSG-RDI (n/h)[‡]	35.5 ± 24.4	30.4–40.7	7.0 ± 4.1	5.0–9.5
ODI (n/h)[§]	21.8 ± 2.5	16.5–26.9	1.9 ± 1.6	1.2–2.7

[†]Number per hour of apnoea and hypopnoea events.
[‡]Number per hour of respiratory disturbance events.
[§]Number per hour of 3% oxygen desaturation events.
AHI, apnoea/hypopnoea index; CI, confidence interval; ESS, Epworth Sleepiness Scale; PSG, polysomnography; RDI, respiratory disturbance index; RERA, respiratory effort-related arousal; SD, standard deviation; TST, total sleep time.

Agreement and relationship between MM and PSG

RDI and TST differences between MM and PSG

Scoring for MM showed a strong agreement between the two readers: ICC (2.1) = 0.967 (95% CI: 0.901–0.992; $P < 0.001$). The differences in RDI and the TST between both scoring methods, PSG and MM, are presented in Tables S1, S2 and S8 (Supplementary Information).

Agreement between the two scoring methods

The MM-RDI underestimated PSG-RDI by a mean of 2.88 events and such underestimate remained constant across all ranges of PSG-RDI (Fig. 2). MM-RDI agreed also with PSG-AHI and hourly rate of >3% oxygen desaturation (ODI). The linear kappa coefficient was good to very good (Table 3). Both AHI and RERA contributed significantly to MM-RDI variance (Table S3, Supplementary Information). Moreover, the adjusted MM-RDI for the mean difference between MM-RDI and PSG-RDI significantly differed by the severity of OSA patients.

Linear relationship between MM and PSG scorings for RDI

In the pooled sample of patients and controls (n = 92), MM-RDI highly significantly correlated with PSG-RDI (Fig. S3, Supplementary Information) with a coefficient of 1.005 (95% CI: 0.96–1.05; $P = 0.001$) and a small residual error (0.026). A simple linear model: PSG-RDI = 2.743 + 1.005 × MM-RDI allowed accurate prediction of up to 90.6% of the PSG-RDI variance (Fisher test: $P < 0.001$). Detailed information about the model is presented in Table S4 and Figure S4 (Supplementary Information). The relationship between MM-RDI and PSG-RDI scores (n/h) was significantly influenced by OSAS severity based on AHI criteria ($P < 0.001$), that is being stronger in more severe OSAS patients (see Table S5, Supplementary Information). MM-RDI was significantly correlated with the hourly rate of cortical

Table 2 Receiver operator characteristics of MM-RDI for detecting PSG-RDI at the diagnostic levels reported in ICSD-3

	For detecting PSG-RDI ≥ 5	For detecting PSG-RDI ≥ 15	For detecting PSG-RDI ≥ 30
Best cut-off point[†]	>5.9	>13.5	>32.5
AUC	0.96 (0.89–0.99)	0.97 (0.91–0.99)	0.91 (0.84–0.96)
Youden's J index[‡]	0.93 (0.83–0.96)	0.92 (0.83–0.97)	0.71 (0.50–0.80)
Sensitivity	93.2 (86.77–97.04)	89.0 (79.83–94.31)	74.29 (57.67–85.92)
Specificity	100.0 (51.01–100.0)	100.0 (83.18–100.0)	96.49 (87.23–99.65)
False positive rate (%)	6.82 (3.16–14.09)	9.59 (4.72–18.49)	25.71 (14.16–42.07)
False negative rate (%)	0.0 (0.0–48.99)	0.0 (0.0–16.82)	3.51 (0.97–11.92)
LLR+	3.73 (2.7–20.4)	8.46 (2.3–31.5)	21.17 (5.35–83.76)
LLR–	0.07 (0.03–0.15)	0.11 (0.06–0.21)	0.27 (0.15–0.47)
Accuracy	0.93 (0.86–0.97)	0.91 (0.84–0.95)	0.88 (0.80–0.93)

[†]Best cut-off point was determined using Youden's J index. The 95% CIs were determined by bootstrapping.
[‡]Youden's J index = sensitivity + specificity − 1.
AUC, area under the curve; CI, confidence interval; ICSD-3, International Classification of Sleep Disorders, Third Edition; LLR+, positive likelihood ratio; LLR–, negative LLR; MM, mandibular movement; PSG, polysomnography; RDI, respiratory disturbance index.

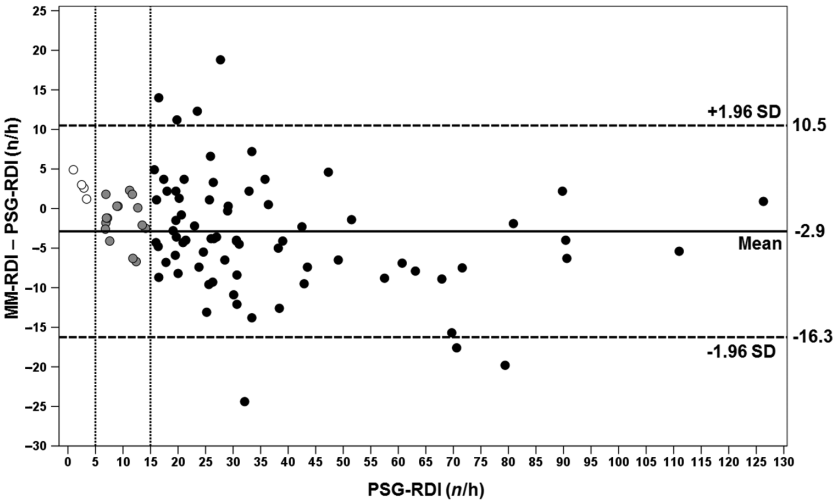

Figure 2 Bland–Altman plot displaying the concordance between MM-RDI (mandibular movement-respiratory disturbance index) and polysomnography (PSG)-RDI. Bland–Altman plot in which the difference between MM-RDI and PSG-RDI (y-coordinate) is plotted against the PSG-RDI (reference method, x-coordinate). The mean of difference (MM-RDI – PSG-RDI = –2.9) is presented as a continuous horizontal line. Two additional horizontal lines are plotted above and below this line at a distance of 1.96 × standard deviation (SD) of the differences and represent the upper and lower limits of agreement (with their SD in dotted lines) at +10.5 and –15.0, respectively. The shaded columns correspond to the critical intervals of mild obstructive sleep apnoea (OSA) (PSG-RDI from 5 to 15) and moderate OSA (PSG-RDI from 15 to 30).

arousals from PSG (Pearson's r = 0.741; 95% CI: 0.536–0.873; $P < 0.001$).

Temporal coincidence

Mean Co.Sen and Co.Spe were 0.83 (95% CI: 0.79–0.87) and 0.74 (95% CI: 0.70–0.78), respectively. Co.Sen was similar across the patients of different severities while Co.Spe was proportional to the severity of OSAS (see Tables S6, S7, Supplementary Information).

Diagnostic performance of MM-RDI against the PSG-based criteria

ROC curve analyses were performed to evaluate the ability of MM-RDI to detect PSG-defined OSAS at three pre-specified selected cut-off points (Fig. 3). The characteristics as well as the best cut-off point of these three classifications are given in Table 2. The post-test probabilities for a positive diagnosis with PSG-RDI ≥ 5/h TST at a cut-off value of MM-RDI = 5.9 and ≥15/h TST at a cut-off value of MM-RDI = 13.5 were 94.2% (95% CI: 92.2–98.9%) and 88.1% (95% CI: 66.8–96.5%), respectively.

Table 3 Agreement between MM and PSG by Cohen's kappa coefficient

	Linear weighted kappa coefficient	95% CI		*P*-value
MM-RDI vs PSG-RDI†	0.679	0.573	0.784	<0.001
MM-RDI vs PSG-AHI	0.520	0.418	0.622	<0.001
MM-RDI vs ODI	0.389	0.295	0.452	0.014
ODI vs PSG-RDI	0.324	0.228	0.420	0.012

The estimated value and 95% CI of the kappa coefficient for classifying the severity of the OSAS patients by three RDI thresholds (5, 15 and 30 events/h).

†PSG-RDI and MM-RDI mean respiratory disturbances hourly index, measured by polysomnography and by MM record, respectively.

AHI, apnoea/hypopnoea index; CI, confidence interval; MM, mandibular movement; ODI, hourly rate of >3% oxygen desaturation; OSAS, obstructive sleep apnoea syndrome; PSG, polysomnography; RDI, respiratory disturbance index.

DISCUSSION

We compared MM analysis with the findings of an in-laboratory unattended conventional PSG for assessment of OSAS in a mid-to-high pre-test probability cohort. The PSG-derived RDI and the MM-RDI were highly correlated in a population spanning a large spectrum of OSAS severity and normal individuals (adjusted $R^2 = 0.906$). Furthermore, MM recordings

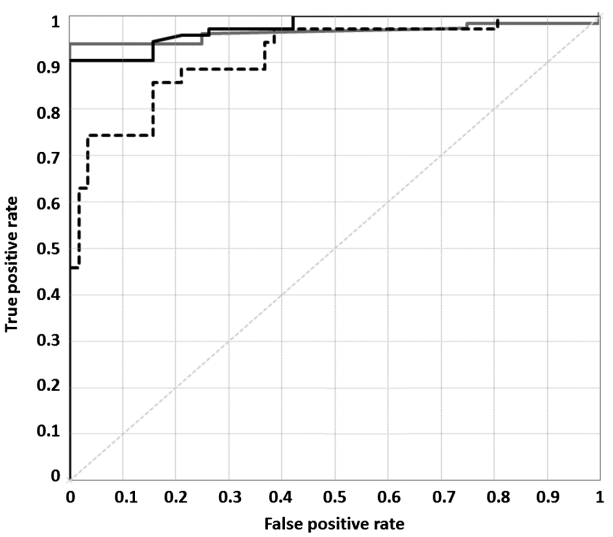

Figure 3 Receiver operating characteristic (ROC) curve analysis. Three ROC curves show the ability of MM-RDI (mandibular movement-respiratory disturbance index) to detect polysomnography (PSG)-defined obstructive sleep apnoea syndrome (OSAS) at three cut-off points: PSG-RDI ≥ 5 (grey), PSG-RDI ≥ 15 (black) and PSG-RDI ≥ 30 (dashed). Their area under the curve (AUC) of ROC values were 0.957, 0.976 and 0.914, respectively.

accurately diagnosed OSAS with an RDI ≥ 5/h TST in the presence of suggestive symptoms, or an RDI ≥ 15/h TST with or without symptoms, independent of the presence of co-morbidities (ICSD-3).

At the cut-off value of ≥15/h TST, the accuracy of type 4 devices has been reported as low.[4,20–28] In contrast, here, for PSG-RDI of ≥5 and ≥15/h TST, the mean positive likelihood ratios (LLRs+) were 3.73 and 8.46, and the post-test probabilities of obtaining a true positive diagnosis were 94.2% and 88.1%, respectively. Our results are aligned with the recommendations of Collop *et al.* regarding the mandatory performance characteristics of portable devices for assessing potential OSA patients in out-of-centre settings.[29] The coincidence analysis provided an event-by-event unique overview of the mandibular behaviour in relationship with the traditional variables recorded during PSG.

Bland–Altman analysis between MM-RDI and PSG-RDI showed an acceptable mean (range) difference of −2.9 (+10.5 to −16.3), and the level of agreement measured with Cohen's kappa coefficient was good to very good, and remarkably better than the agreement levels shown in other studies exploring single-channel devices.[19,27,29] The agreement between MM-RDI and PSG-AHI was observed but to a lesser degree than with PSG-RDI. This is due to the contributions of RERAs to determine MM-RDI as shown in Table S3 (Supplementary Information). Furthermore, the level of agreement was high regardless of severity of the disease, even when considering RDI instead of AHI (detection of RERAs is optional in the AASM scoring rules), an approach that could have amplified the differences between the two methods.

MM recordings detected SDB with high sensitivity in all spectrums of event severity (i.e. apnoeas, hypopnoeas or RERAs) with a low false negative rate. To increase the number of subjects in the low RDI range, in which single-channel approaches might perform poorly, we added 13 young healthy individuals without sleep complaints. This facilitated a better estimate of the specificity when considering the performance of MM and an RDI cut-off value of >5/h TST. The population explored in our centre also included subjects without a high pre-test probability of OSAS, thus encompassing the entire spectrum of the clinical OSAS presentations commonly seen in clinical practice. This prevented overestimation of true positives for MM analysis in the context of a high pre-test probability for OSAS, making our analyses more relevant to clinical practice, where all disease severities occur. This is clinically relevant as specific screening strategies are desirable in asymptomatic cardiovascular and metabolic populations. Thus, the significance of the relationship between MM-RDI and PSG-RDI improved with OSA severity, while the risk of underestimation of RDI by MM was minimal even at low level of OSAS severity.

The high agreement encountered herein confirms that MM recording performs extremely well across all OSAS severities, with only four outliers being identified in the PSG-RDI range of 15–30/h TST. In these cases, MM-RDI overestimated RDI compared with PSG-RDI. This contrasts to other type 4 portable devices which underestimate RDI.[28,30] This overestimation may be due to increased sensitivity of MM to detect REs, or the recording of events during wakefulness, with MM recordings failing to identify such wakefulness periods. However, both the present study and a previous one[10] indicate that MM analysis provides reliable estimates of TST (which is used in this study to compute MM-RDI, Tables S3, S4 (Supplementary Information)). The reliability of TST provided by MM analysis did contribute to the performance of MM-RDI.

Recording of nasal pressure (nasal cannula connected to a pressure transducer) alone or in combination with oximetry using a portable device has been documented to confidently rule in OSAS in a high pre-test sleep clinic population, and is sensitive to even subtle changes in airflow.[4] The inspiratory portion of the nasal pressure waveform can display flattening, a surrogate of airflow limitation when using appropriate filter settings.[31–33] Nevertheless, nasal flow remains an unstable signal that is susceptible to confounding factors such as nasal congestion, unstable positioning of the nasal cannula and diminution of the signal due to mouth breathing, while in contrast MM analysis detects mouth opening and is insensitive to the other confounders.[34] Compared with a nasal pressure-based type 4 device, superior diagnostic performance was observed with MM because hypopnoeas, RERAs and cortical arousals are reliably identified, in marked contrast with other portable devices.[10]

Common pitfalls related to portable devices type 3 and type 4 when compared with type 1 or type 2 for estimating RDI have been controlled for in this study, as there were no effects of different study nights (night to night variability) and the environment was standardized.

Our study was one of the first assessing event-by-event temporal coincidences and not only comparing RDI between two methods. We should emphasize that the temporal coincidence between PSG and MM respiratory events was high. In terms of sensitivity, the coincidence was unchanged across levels of severities supporting the use of MM in ambulatory device for diagnosing SDB.

The limitations of the study are the following: The respiratory events identified by MM analysis alone cannot differentiate between apnoea, hypopnoea or RERAs as the flow is not measured. However, MM is a reliable indicator for respiratory events and can differentiate readily between obstructive or central respiratory events.[9]

The high level of in-laboratory agreement between MM-RDI and PSG-RDI will need to be confirmed with a home-based study. On the other hand, video surveillance during PSG could inform about sleep MM.

The accuracy to rule in OSA has to be measured in a group of patients with low (<0.5) pre-test probability to ascertain whether the LLR for a positive test provides a post-test probability of at least 0.95 for a true positive diagnosis. Our study did not address the question of ruling out OSAS. However, the false negative rate was low, thereby contrasting with other unattended portable studies, such that the risk of wrongly categorizing mild subjects below the diagnostic threshold for OSAS remains very low.[27]

In summary, MM emerges as a useful tool in the diagnosis of OSAS. MM recordings during one night estimate a similar RDI compared with traditional PSG in a group of symptomatic patients of all ages with or without co-morbidities. Moreover, the risk of misidentifying mild OSAS subjects as being normal remains very low. Thus, a type 4 device based on MM analysis could offer a reliable and valid alternative to the conventional in-laboratory recording when its use is aimed at a clearly defined and carefully selected population.

Acknowledgements

Many thanks to Sandrine Lamy and Liesbeth Orij who kindly spent innumerable hours providing secretarial assistance.

REFERENCES

1 Heinzer R, Vat S, Marques-Vidal P, Marti-Soler H, Andries D, Tobback N, Mooser V, Preisig M, Malhotra A, Waeber G *et al.* Prevalence of sleep-disordered breathing in the general population: the HypnoLaus study. *Lancet Respir. Med.* 2015; **3**: 310–8.
2 Corral-Peñafiel J, Pepin JL, Barbe F. Ambulatory monitoring in the diagnosis and management of obstructive sleep apnoea syndrome. *Eur. Respir. Rev.* 2013; **22**: 312–24.
3 Park JG, Gay PC. Home sleep testing: it's not how you play the game, it's whether you win or lose. *J. Clin. Sleep Med.* 2015; **11**: 411–2.
4 Masa JF, Duran-Cantolla J, Capote F, Cabello M, Abad J, Garcia-Rio F, Ferrer A, Mayos M, Gonzalez-Mangado N, de la Peña M *et al.*; Spanish Sleep Network. Effectiveness of home single-channel nasal pressure for sleep apnea diagnosis. *Sleep* 2014; **37:** 1953–61.
5 Thornton AT, Singh P, Ruehland WR, Rochford PD. AASM criteria for scoring respiratory events: interaction between apnea sensor and hypopnea definition. *Sleep* 2012; **35**: 425–32.
6 Masa JF, Corral J, Pereira R, Duran-Cantolla J, Cabello M, Hernández-Blasco L, Monasterio C, Alonso A, Chiner E, Zamorano J *et al.*; Spanish Sleep Network. Therapeutic decision-making for sleep apnea and hypopnea syndrome using home respiratory polygraphy: a large multicentric study. *Am. J. Respir. Crit. Care Med.* 2011; **184:** 964–71.
7 Martinot JB, Senny F, Denison S, Cuthbert V, Gueulette E, Guénard H, Pépin JL. Mandibular movements identify respiratory effort in pediatric obstructive sleep apnea. *J. Clin. Sleep Med.* 2015; **11**: 567–74.
8 Martinot JB, Denison SJ, Senny FH, Robillard TA, Khatwa U, Guenard H. Mandibular movements and respiratory efforts against upper airway obstruction during sleep of a pre-school child. *Open Sleep J.* 2014; **7**: 1–5.
9 Maury G, Senny F, Cambron L, Albert A, Seidel L, Poirrier R. Mandible behaviour interpretation during wakefulness, sleep and sleep-disordered breathing. *J. Sleep Res.* 2014; **23**: 709–16.
10 Cheliout-Heraut F, Senny F, Djouadi F, Ouayoun M, Bour F. Obstructive sleep apnoea syndrome: comparison between polysomnography and portable sleep monitoring based on jaw recordings. *Neurophysiol. Clin.* 2011; **41**: 191–8.
11 Escourrou P, Grote L, Penzel T, Mcnicholas WT, Verbraecken J, Tkacova R, Riha RL, Hedner J; ESADA Study Group. The diagnostic method has a strong influence on classification of obstructive sleep apnea. *J. Sleep Res.* 2015; **24**: 730–8.
12 Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. *Sleep* 1991; **14**: 50–5.
13 Berry RB, Budhiraja R, Gottlieb DJ, Gozal D, Iber C, Kapur VK, Marcus CL, Mehra R, Parthasarathy S, Quan SF *et al.*; American Academy of Sleep Medicine. Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. *J. Clin. Sleep Med.* 2012; **8:** 597–619.
14 American Academy of Sleep Medicine. *International Classification of Sleep Disorders*, 3rd edn. Darien, IL, American Academy of Sleep Medicine, 2014.
15 Krouwer JS. Why Bland-Altman plots should use X, not (Y+X)/2 when X is a reference method. *Stat. Med.* 2008; **27**: 778–80.
16 Jacob C. Weighted kappa: nominal scale agreement among many raters. *Psychol. Bull.* 1968; **76**: 378–82.
17 RN ML. Interrater Reliability with SPSS for Windows 5.0. *Am. Stat.* 1993; **47**: 292–6.
18 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; **44**: 837–45.
19 Collop NA, Tracy SL, Kapur V, Mehra R, Kuhlmann D, Fleishman SA, Ojile JM. Obstructive sleep apnea devices for out-of-center (OOC) testing: technology evaluation. *J. Clin. Sleep Med.* 2011; **7:** 531–48.
20 Harrell FE. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York, Springer-Verlag, 2001.
21 Bar A, Pillar G, Dvir I, Sheffy J, Schnall RP, Lavie P. Evaluation of a portable device based on peripheral arterial tone for unattended home sleep studies. *Chest* 2003; **123**: 695–703.
22 Grover SS, Pittman SD. Automated detection of sleep disordered breathing using a nasal pressure monitoring device. *Sleep Breath.* 2008; **12**: 339–45.
23 Heneghan C, Chua CP, Garvey JF, de Chazal P, Shouldice R, Boyle P, McNicholas WT. A portable automated assessment tool for sleep apnea using a combined Holter-oximeter. *Sleep* 2008; **31**: 1432–9.
24 Levy P, Pepin JL, Deschaux Blanc C, Paramelle B, Brambilla C. Accuracy of oximetry for detection of respiratory disturbances in sleep apnea syndrome. *Chest* 1996; **109**: 395–9.
25 Epstein LJ, Dorlac GR. Cost-effectiveness analysis of nocturnal oximetry as a method of screening for sleep apnea-hypopnea syndrome. *Chest* 1998; **113**: 97–103.
26 Pittman SD, Ayas NT, MacDonald MM, Malhotra A, Fogel RB, White DP. Using a wrist-worn device based on peripheral arterial tonometry to diagnose obstructive sleep apnea: in-laboratory and ambulatory validation. *Sleep* 2004; **27**: 923–33.
27 Ward KL, McArdle N, James A, Bremner AP, Simpson L, Cooper MN, Palmer LJ, Fedson AC, Mukherjee S, Hillman DR. A comprehensive evaluation of a two-channel portable monitor to "rule in" obstructive sleep apnea. *J. Clin. Sleep Med.* 2015; **11**: 433–44.
28 Fredheim JM, Røislien J, Hjelmesæth J. Validation of a portable monitor for the diagnosis of obstructive sleep apnea in morbidly obese patients. *J. Clin. Sleep Med.* 2014; **10**: 751–7.
29 Erman MK, Stewart D, Einhorn D, Gordon N, Casal E. Validation of the ApneaLink for the screening of sleep apnea: a novel and simple single-channel recording device. *J. Clin. Sleep Med.* 2007; **3**: 387–92.
30 Ayappa I, Norman RG, Seelall V, Rapoport DM. Validation of a self-applied unattended monitor for sleep disordered breathing. *J. Clin. Sleep Med.* 2008; **4**: 26–37.
31 Kirsch DB. PRO: sliding into home: portable sleep testing is effective for diagnosis of obstructive sleep apnea. *J. Clin. Sleep Med.* 2013; **9**: 5–7.
32 Farré R, Montserrat JM, Navajas D. Noninvasive monitoring of respiratory mechanics during sleep. *Eur. Respir. J.* 2004; **24**: 1052–60.
33 Hosselet JJ, Norman RG, Ayappa I, Rapoport DM. Detection of flow limitation with a nasal cannula/pressure transducer system. *Am. J. Respir. Crit. Care Med.* 1998; **157**(5 Pt 1): 1461–7.
34 Norman RG, Ahmed MM, Walsleben JA, Rapoport DM. Detection of respiratory events during NPSG: nasal cannula/pressure sensor versus thermistor. *Sleep* 1997; **20**: 1175–84.

Supplementary Information

Additional supplementary information can be accessed via the *html* version of this article at the publisher's website.

Appendix S1 Methods.
Figure S1 Temporal coincidence analysis between mandibular movement signal and polysomnography recording.
Figure S2 Frequency and intensity of individual obstructive sleep apnoea symptoms among OSA syndrome suspected patients in the cohort.
Figure S3 Linear relationship between respiratory disturbance index obtained by the two scoring methods (polysomnography vs mandibular movement).
Figure S4 MM-RDI results by four levels of obstructive sleep apnoea syndrome severity defined by polysomnography.
Table S1 PSG-RDI and mandibular movement RDI scores by categories of obstructive sleep apnoea syndrome severity.
Table S2 Differences between the two scoring methods (mandibular movement analysis and conventional polysomnography).
Table S3 Multivariate linear model for evaluating the partial effects of PSG-RERA and PSG-apnoea/hypopnoea index on MM-RDI in the studied population.
Table S4 Simple linear model for evaluating the relationship between PSG-RDI and mandibular movement RDI in the studied population.
Table S5 Moderation analysis.
Table S6 Linear regression analysis between the coincidence sensitivity and the coincidence specificity and PSG-AHI.
Table S7 Sensitivity and specificity of the temporal coincidence across four groups of obstructive sleep apnoea syndrome severity.
Table S8 The relationship between PSG-TST and mandibular movement TST analysed by a linear regression analysis.

