ORIGINAL ARTICLE

Scoring criteria for portable monitor recordings: a comparison of four hypopnoea definitions in a population-based cohort

Sopharat Vat,[1,2] Jose Haba-Rubio,[1] Mehdi Tafti,[1,3] Nadia Tobback,[1] Daniela Andries,[1] Raphael Heinzer[1,4]

[1]Centre for Investigation and Research in Sleep (CIRS), Lausanne University Hospital (CHUV), Lausanne, Switzerland
[2]Pulmonary Medicine Department, University Hospital of Montreal (CHUM), Montreal, Quebec, Canada
[3]Centre for Integrative Genomics, University of Lausanne (UNIL), Lausanne, Switzerland
[4]Pulmonary Department (CHUV), Lausanne University Hospital, Lausanne, Vaud, Switzerland

Correspondence to
Dr Raphaël Heinzer, Center for Investigation and Research in Sleep (CIRS), University Hospital of Lausanne, Lausanne 1011, Switzerland; Raphael.Heinzer@chuv.ch

SV and JH-R are joint first authors.

ABSTRACT

Rationale Limited-channel portable monitors (PMs) are increasingly used as an alternative to polysomnography (PSG) for the diagnosis of obstructive sleep apnoea (OSA). However, recommendations for the scoring of PM recordings are still lacking. Pulse-wave amplitude (PWA) drops, considered as surrogates for EEG arousals, may increase the detection sensitivity for respiratory events in PM recordings.

Objectives To investigate the performance of four different hypopnoea scoring criteria, using 3% or 4% oxygen desaturation levels, including or not PWA drops as surrogates for EEG arousals, and to determine the impact of measured versus reported sleep time on OSA diagnosis.

Methods Subjects drawn from a population-based cohort underwent a complete home PSG. The PSG recordings were scored using the 2012 American Academy of Sleep Medicine criteria to determine the apnoea–hypopnoea index (AHI). Recordings were then rescored using only parameters available on type 3 PM devices according to different hypopnoea criteria and patients-reported sleep duration to determine the 'portable monitor AHIs' (PM-AHIs).

Main results 312 subjects were included. Overall, PM-AHIs showed a good concordance with the PSG-based AHI although it tended to slightly underestimate it. The PM-AHI using 3% desaturation without PWA drops showed the best diagnostic accuracy for AHI thresholds of ≥5/h and ≥15/h (correctly classifying 94.55% and 93.27% of subjects, respectively, vs 80.13% and 87.50% with PWA drops). There was a significant but modest correlation between PWA drops and EEG arousals (r=0.20, p=0.0004).

Conclusion Interpretation of PM recordings using hypopnoea criteria which include 3% desaturation without PWA drops as EEG arousal surrogate showed the best diagnosis accuracy compared with full PSG.

Key messages

What is the key question?
► How should sleep portable monitor (PM) recordings be scored in the absence of EEG-defined arousals and sleep duration?

What is the bottom line?
► Hypopnoea criteria using 3% oxygen desaturation level without pulse-wave amplitude (PWA) drops as EEG arousal surrogate accurately classified mild-to-moderate obstructive sleep apnoea (OSA), whereas including PWA drops may provide a higher accuracy only for diagnosis of severe OSA.

Why read on?
► This study allows direct comparison between complete polysomnography and type 3 PM recorders, provides new insights on the scoring of respiratory events with PM recordings and evaluates a novel way to define respiratory events with the use of PWA drops as surrogates for EEG arousals.

INTRODUCTION

Obstructive sleep apnoea (OSA) is a highly prevalent disease associated with neurocognitive[1] impairment and cardiovascular morbidity.[2][3] Despite an increased awareness, studies have shown that >75% of the patients remain undiagnosed and untreated.[4] To date, attended in-laboratory polysomnography (PSG) is considered the gold standard for OSA diagnosis.[5] However, considering the cost, technical complexity and human resources required for PSG, unattended portable monitor (PM) devices have been proposed as an alternative technique for the diagnosis of OSA. Recent studies suggest that the diagnosis of OSA with PM devices can be as accurate as with PSG in selected populations.[6–11] Current guidelines from the American Academy of Sleep Medicine (AASM) recommend the use of unattended PMs for the diagnosis of OSA only in combination with a comprehensive clinical assessment in patients with a high pre-test probability of moderate-to-severe OSA and without comorbid sleep disorder or major comorbid medical disorders.[6][12] The AASM recommends type 3 PM devices, which include an average of 4–7 channels with at least a measurement of airflow, respiratory effort and blood oxygenation.[6][13]

There are however concerns regarding the interpretation of PM recordings because of the lack of an EEG signal for the determination of sleep periods and arousals. Frequency of respiratory events identified with PM devices (PM-apnoea–hypopnoea index (AHI)) is calculated based on the

number of apnoeas and hypopnoeas divided by the total recording time (TRT), while the PSG-AHI is calculated based on the number of apnoeas and hypopnoeas divided by the total sleep time (TST). Moreover, PM recordings cannot incorporate the criteria of arousals used for the identification of hypopnoeas as recommended by the AASM.[5] [14] Therefore, PM devices may underestimate the severity of OSA because of a longer TRT compared with TST (increased denominator) and because of a lower number of hypopnoeas due to the absence of EEG arousal detection (decreased numerator). On the other hand, possible scoring of respiratory events during nocturnal wake periods with PM devices (increased numerator) may result in an overestimation of the PM-AHI compared with the PSG-determined AHI.

Currently, there are no specific recommendations for the scoring of respiratory events in PM recordings. Even though type 3 PMs cannot detect arousals because of the absence of EEG, indirect detection of arousals using surrogate signals have been elaborated. For instance, pulse-wave amplitude (PWA) drops have been shown to be a sensitive surrogate marker for EEG arousals.[15] PWA drops is a non-invasive measure obtained from finger photoplethysmography provided by most pulse oximeters. It is believed to be a marker of peripheral vasoconstriction induced by the autonomic response associated with arousals from sleep (autonomic arousals). PWA drops could therefore be considered as a surrogate for EEG arousals and used as a secondary criterion in the definition of hypopnoeas in PM recordings. Figure 1 shows an example of a hypopnoea followed by an arousal and a PWA drop.

Even though this marker has not been validated in the general population, PWA drops sensitivity of 89.1% and 70.9% for EEG arousals were reported in patients with non-invasive ventilation and in patients with OSA respectively.[15] [16] Moreover the magnitude of PWA drops is associated with the EEG arousal intensity and the use of PWA drops as surrogate for EEG arousal was shown to improve inter-scorer reliability for PSG.[17] [18]

We hypothesised that the use of PWA drops as surrogates for EEG arousals may improve the sensitivity of the PM-AHI and that the use of a higher threshold for oxygen desaturation (4% instead of 3%) could improve the specificity of the PM-AHI for the diagnosis of OSA. The objective of our study was to determine the performance of different hypopnoea scoring criteria, including or not PWA drops and using 3% or 4% oxygen desaturation thresholds, for the scoring of PM recordings in comparison with complete home PSG.

METHODS
Selection and description of the participants
For the present study, 312 subjects were drawn from the HypnoLaus cohort (previously described[19]) and stratified according to AHI severity (<5/h, 5–15/h, 15–30/h and >30/h), age (≤60 and >60 years of age) and gender. We also identified a subpopulation with a high pre-test risk of OSA defined as an Epworth Sleepiness Scale (ESS)[20] score >10 and a STOP-BANG score[21] of at least 3/8. The latter score included the following items: *S*noring, daytime *T*iredness, *O*bserved apnoea, high blood *P*ressure, *B*ody mass index ≥35 kg/m^2, *A*ge ≥50 years, *N*eck circumference ≥40 cm and male *G*ender. Subjects on β-blocker therapy were excluded from the analysis as this treatment could potentially blunt autonomic response and therefore alter PWA drop detection.

Polysomnography
Complete PSGs were performed unattended at home and were scored according to the AASM 2007 recommendation.[14] Certified PSG technicians equipped the subjects with the PSG recording system (Titanium, Embla Flaga, Reykjavik, Iceland) in late afternoon at the Lausanne University Sleep Centre (Centre for Investigation and Research in Sleep). PSG recording and interpretation techniques used in HypnoLaus was previously described.[19] Respiratory events were scored manually according to the updated AASM 2012 recommendations[5] with hypopnoea defined as: (1) a drop of nasal pressure excursion of at least 30% from pre-event baseline, (2) lasting at least 10 s and (3) accompanied with either a 3% oxygen desaturation or an EEG arousal. Frequencies of respiratory events with PSG were reported as AHI, that is, the number of apnoeas and hypopnoeas divided by the TST in hours. The arousal index was calculated as the number of EEG arousals divided by the TST.

Type 3 portable monitor devices
The PSG recordings were reinterpreted by a trained pulmonologist (SV) after removing channels available only in PSG (EEG, electro-oculography and muscle activity) in order to simulate a type 3 PM device. The remaining channels were: nasal pressure, pulse oximetry, thoracic and abdominal respiratory effort belts and body position. Recording duration was determined according to the subject's self-reported 'lights-off' and 'lights-on' times. Respiratory events were scored manually. Hypopnoea was defined as: (1) a drop of nasal pressure excursion of at least 30% from pre-event baseline, (2) lasting at least 10 s and (3) four sets of secondary criteria described in table 1. An automated algorithm identified PWA drops on the photoplethysmography channel based on drop of at least 30% of baseline PWA, a PWA drop duration of at least 3 s and ≤50 s and a minimum interval between two PWA drops of at least 10 s. To be associated with a respiratory event, PWA drops had to occur at least 5 s after the beginning of the respiratory event and no later than 10 s after its end. Frequencies of respiratory events with type 3 PM recordings were reported as PM-AHI values, that is, the number of apnoea and hypopnoea divided by the TRT. The PWA drop index corresponded to the number of PWA drops divided by the TRT.

Statistical analysis
Agreement between PM-AHIs and the AHI, between arousal index and PWA drop index and between TST and TRT were assessed using Bland–Altman plots in which the mean value of both measures are plotted on the abscissa and their difference value on the ordinate. The limit of agreement between the two measurements was defined as the mean difference value ±1.96×SD of the individual differences (see online supplementary figure S1). The accuracy of the PM-AHI was also determined for different AHI thresholds (≥5/h, ≥15/h and ≥30/h) and was expressed as sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio (+LR), negative likelihood ratio (−LR), positive post-test probability (number of true-positive divided by all positive results), negative post-test probability (number of false-negative divided by all negative results) and percentage of correctly classified subjects (the number of true-positives and -negatives results divided by all positive and negative results). Receiver operating characteristic (ROC) curves were used to illustrate the discriminatory abilities of the PM-AHI for the same established AHI thresholds. The statistical programmes used were Stata V.11 (StataCorp,
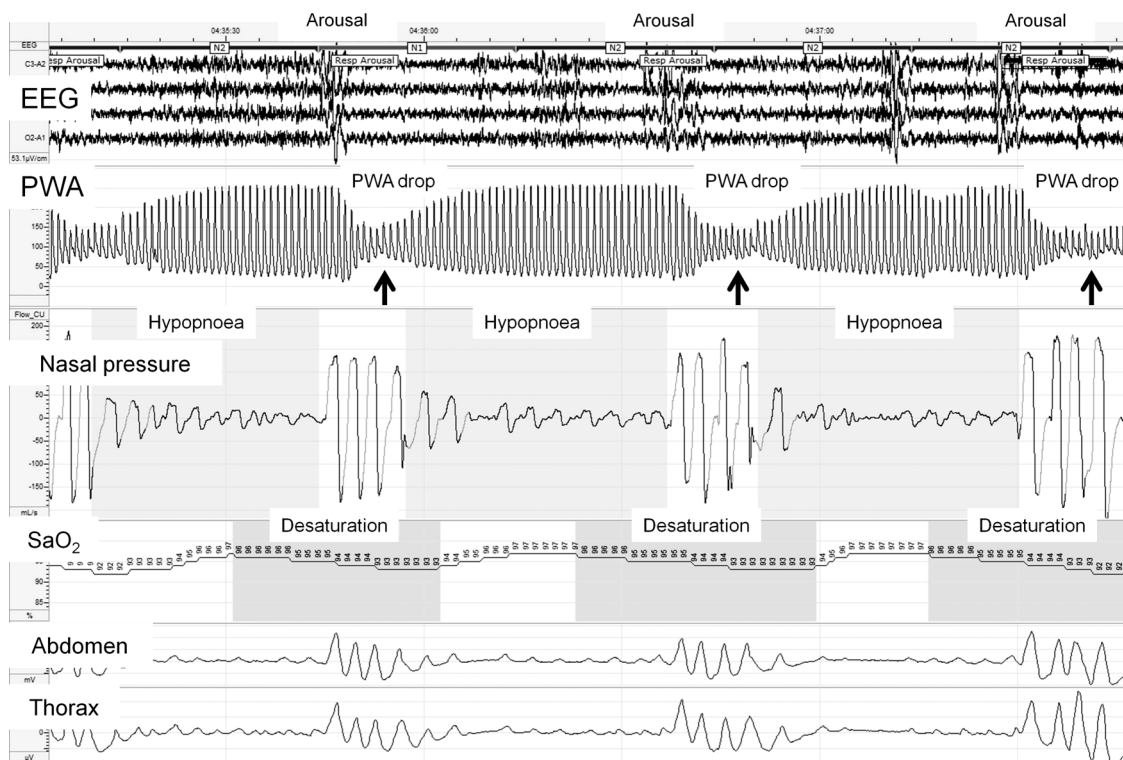
Figure 1 Example of hypopnoea followed by EEG arousal and a pulse-wave amplitude (PWA) drops (arrows). SaO$_2$, blood oxygen saturation level; Abdomen, abdominal movements; Thorax, thoracic movements.

College Station, Texas, USA) and Medcalc for Bland–Altman plots (V.12.7.7.0 Medcalc software, Ostend, Belgium).

RESULTS

Table 2 summarises anthropometric characteristics and main PSG data. Included subjects were evenly distributed among OSA severity categories defined by AHI thresholds (<5/h, 5–15/h, 15–30/h and ≥30/h), between age categories (≤60 and >60 years of age) and between genders. A total of 116 subjects (37.18%) were identified as having 'high pre-test-risk' of OSA.

We observed a significant but modest correlation between the EEG arousal index determined by PSG and the PWA drop index obtained with type 3 PM devices (r=0.20, p=0.0004). The latter tended to overestimate the arousal index by a mean of 15.61 ±17.53 events/h. Almost a third of hypopnoeas (median (95% CI): 31.3% (27.8 to 33.9)) scored on the PSG recordings were not associated with a 3% oxygen desaturation. There was also a

correlation between TST and TRT (r=0.565, p<0.0001) but TST was overestimated by a mean of 80.74±64.11 min by TRT.

Table 3 summarises the results of the Bland–Altman analysis for the global and 'high-risk' populations, respectively (the four detailed Bland–Altman plots are shown on online supplementary table S1 in repository). This table shows that the evaluated PM-AHIs tend to underestimate the PSG-determined AHI (mean difference PM-AHI$_1$<PM-AHI$_4$<PM-AHI$_3$: −1.3±4.8/ h< −2.8±7.4/h< −7.6±7.5/h, respectively) except for PM-AHI$_2$ that resulted in an overestimation (mean difference +3.5±5.4/h). Differences between the AHI and PM-AHIs and

Table 1 Hypopnoea scoring criteria

	Hypopnoea scoring criteria*
	a. Drop of nasal pressure excursion of at least 30% from baseline *and* b. Duration of at least 10 s *and*
PM-AHI$_1$	c. Oxygen desaturation of at least 3%
PM-AHI$_2$	c. Oxygen desaturation of at least 3% *and/or* PWA drop of 30% from baseline
PM-AHI$_3$	c. Oxygen desaturation of at least 4%
PM-AHI$_4$	c. Oxygen desaturation of at least 4% *and/or* PWA drop of 30% from baseline

*Scoring of a hypopnoea requires the presence of criteria (a) and (b) in association with either (c) and/or (d).
PM-AHI, portable monitor apnoea–hypopnoea index; PWA, pulse-wave amplitude.

Table 2 Characteristics of the studied population

	Total population n=312	'High-risk' subgroup n=116
Male; %	50.32	72.41
Age (years); mean±SD	61.16±10.71	61.55±10.81
40–60 years; %	50.00	48.28
60–85 years; %	50.00	51.72
BMI, kg/m^2; mean±SD	26.19±4.60	27.49±5.10
AHI, n/h; median (IQR)	13.85 (4.50–27.10)	16.85 (6.95–30.95)
AHI categories; n (%)		
AHI<5/h	80 (25.6)	23 (19.8)
AHI 5–15/h	80 (25.6)	30 (25.9)
AHI 15–30/h	78 (25.0)	31 (26.7)
AHI ≥30/h	74 (23.7)	32 (27.6)
ODI 3%, n/h; median (IQR)	13.1 (4.6–24.6)	17.65 (6.80–27.95)
ODI 4%, n/h; median (IQR)	5.75 (1.35–14.7)	9.00 (2.65–18.40)

AHI, apnoea–hypopnoea index; BMI, body mass index; 'high-risk' for OSA, subjects with an Epworth Sleepiness Scale score of >10 and/or a STOPBANG score of ≥3/8; ODI, oxygen desaturation index; OSA, obstructive sleep apnoea.

Table 3 Bland–Altman plots summary

	Total population		High-risk population	
	Δ AHI—PM-AHI	95% CI	Δ AHI—PM-AHI	95% CI
PM-AHI$_1$	1.3	−8.2 to 10.8	1.7	−8.6 to 11.9
PM-AHI$_2$	−3.5	−14.0 to 7.0	−3.0	−14.9 to 8.8
PM-AHI$_3$	7.6	−7.0 to 22.2	8.7	−7.5 to 24.9
PM-AHI$_4$	2.8	−11.7 to 17.4	4.0	−12.4 to 20.4

AHI, the polysomnography apnoea–hypopnoea index; PM-AHI$_x$, the portable monitor apnoea–hypopnoea indices; PSG, polysomnography; ΔAHI—PM-AHI, the difference between the polysomnography and the portable monitor apnoea–hypopnoea index (positive numbers show an underestimation of PSG-AHI by PM-AHI, negative numbers the opposite).

limits of agreement (1.96×SD) were the lowest with PM-AHI$_1$ (mean difference 1.3±4.8/h for the total population and 1.7 ±5.2/h for the 'high-risk').

The discriminatory abilities of the PM-AHIs were assessed in terms of sensitivity and specificity using ROC curve analysis for AHI thresholds of ≥5/h, ≥15/h and ≥30/h. The area under the curve (AUC) values are summarised in table 4. For all subjects, the best operating characteristics were obtained mostly with PM-AHI$_1$ and increased with higher AHI thresholds.

Tables 5 and 6 summarise the diagnostic accuracy of the PM-AHIs. These tables show that for any given AHI thresholds, inclusion of PWA drops in hypopnoea criteria increased sensitivity, whereas using an oxygen desaturation of 4% increased specificity. In our population sample, prevalence of OSA (pre-test probability) for AHI thresholds of ≥5/h and ≥15/h were 74.4% and 48.7%, respectively, according to the PSG results. In the high-risk population, the prevalence for the same AHI thresholds was 80.2% and 54.3%, respectively. For these AHI thresholds, PM-AHI$_1$ resulted in the best diagnostic accuracy in terms of combined sensitivity (≥90.13%), specificity (≥81.25%), +LR (≥5.24), −LR (≤0.10) and thus correctly classified 93.3% of subjects. For an AHI threshold of ≥30/h, PM-AHI$_2$ was superior to PM-AHI$_1$ regarding combined sensitivity (≥89.19%), specificity (≥95.24%), +LR (≥20.34), −LR (≤0.11) and therefore correctly classified 94.2% subjects. The accuracy of the four PM-AHIs was mildly higher in the 'high-risk group' compared with all subjects, probably due to a slightly higher sleep

disordered breathing (SDB) prevalence, reducing the risk of false-negative results in this group.

DISCUSSION

To our knowledge, this is the first study comparing different scoring criteria for the interpretation of PM recordings using traditional and alternative signals such as PWA drops. According to our results, interpretation of PM recordings using hypopnoea criteria, which include 3% desaturation level without PWA drops, showed the best diagnosis accuracy for mild and moderate OSA. Incorporating PWA drops as surrogates for EEG arousals only adds accuracy in detecting severe OSA.

Overall, all tested PM-AHIs showed a high concordance with PSG-AHI and slightly underestimated the number of respiratory events for most of them (except PM-AHI$_2$). This seems to be mainly due to a difference between TRT based on self-reported 'lights off' and 'lights on' times compared with EEG-measured TST, which was about 80 min shorter. This increase in the denominator of the PM-AHI (ie, the TRT) systematically decreased the severity of PM-AHI values in almost all of the subjects. There is thus a need for a reliable non-EEG-based algorithm that could predict sleep time using PM physiological signals in order to improve PM accuracy.

When PWA drops were included in hypopnoea definitions along with 3% desaturation levels (PM-AHI$_2$), the increase in the numerator (ie, the number of respiratory events) led to a mild overestimation of the index. However, including PWA drops did not improve the number of correctly classified subjects with mild-to-moderate SDB, probably due to the poor correlation between PWA drops and EEG arousals.

Most previous studies in which PM devices were used evaluated 'high-risk' populations referred to sleep clinics for suspicion of OSA. For that reason, AASM 2007 clinical guidelines[6] recommended the use of type 3 PM devices only in patients with a 'high pre-test probability' for OSA. One of the strengths of our study was the inclusion of a sample of subjects drawn from the general population and stratified to obtain a sample evenly distributed in terms of age, gender and OSA severity. This allowed testing the effect of the different scoring criteria in a population with higher and lower pre-test probability of OSA. When we specifically studied a subgroup of subjects identified as 'high-risk' for OSA based on clinical symptoms and parameters (ESS[20] >10/24 and STOP-BANG[21] >2/8) to simulate the

Table 4 Details of receiver operating characteristic curves for PM-AHI$_1$, PM-AHI$_2$, PM-AHI$_3$ and PM-AHI$_4$ according to polysomnographic AHI thresholds of ≥5/h, ≥15/h and ≥30/h for (1) all the subjects and (2) 'high-risk' subjects

	AHI ≥5/h		AHI ≥15/h		AHI ≥30/h	
	AUC	95% CI	AUC	95% CI	AUC	95% CI
All subjects						
PM-AHI$_1$	0.986	0.966 to 0.996	0.987	0.968 to 0.997	0.991	0.973 to 0.998
PM-AHI$_2$	0.971	0.946 to 0.987	0.979	0.957 to 0.992	0.984	0.963 to 0.995
PM-AHI$_3$	0.981	0.959 to 0.993	0.981	0.958 to 0.993	0.983	0.962 to 0.994
PM-AHI$_4$	0.947	0.916 to 0.969	0.968	0.941 to 0.984	0.979	0.957 to 0.992
High-risk subjects						
PM-AHI$_1$	0.996	0.961 to 1.000	0.988	0.947 to 0.999	0.994	0.957 to 1.000
PM-AHI$_2$	0.984	0.940 to 0.998	0.984	0.940 to 0.998	0.984	0.941 to 0.998
PM-AHI$_3$	0.998	0.965 to 1.000	0.985	0.943 to 0.999	0.985	0.942 to 0.999
PM-AHI$_4$	0.957	0.903 to 0.986	0.972	0.924 to 0.994	0.981	0.936 to 0.997

AHI, apnoea–hypopnoea index; AUC, area under the curve; 'high-risk' subjects, subjects with an Epworth Sleepiness Scale score of >10 and/or a STOP-BANG score of ≥3/8; PM-AHI, portable monitor apnoea–hypopnoea index.

Table 5 Diagnostic accuracy of PM-AHIs for AHI thresholds of ≥5/h, ≥15/h and ≥30/h in all subjects

	Sens (%)	Spe (%)	PPV (%)	NPV (%)	+LR	+PTP (%)	−LR	−PTP (%)	CC (%)
AHI ≥5/h (n=232; pre-test probability=74.36%)									
PM-AHI$_1$	99.14	81.25	93.88	97.01	5.29	93.88	0.01	2.82	94.55
PM-AHI$_2$	100.00	22.50	78.91	100.00	1.29	78.91	0.00	0.00	80.13
PM-AHI$_3$	77.16	98.75	99.44	59.85	61.72	99.44	0.23	40.01	82.69
PM-AHI$_4$	98.71	51.25	85.45	93.18	2.02	85.41	0.03	8.00	86.54
AHI ≥15/h (n=152; pre-test probability=48.72%)									
PM-AHI$_1$	90.13	96.25	95.80	91.12	24.04	95.81	0.10	8.68	93.27
PM-AHI$_2$	99.34	76.25	79.89	99.19	4.18	79.88	0.01	0.94	87.50
PM-AHI$_3$	53.95	100.00	100.00	69.57	*	100.00	0.46	30.41	77.56
PM-AHI$_4$	77.63	96.88	95.93	82.01	24.84	95.93	0.23	17.93	87.50
AHI ≥30/h (n=74; pre-test probability=23.72%)									
PM-AHI$_1$	70.27	100.00	100.00	91.54	*	100.00	0.30	8.53	92.95
PM-AHI$_2$	89.19	95.80	86.84	96.61	21.23	86.85	0.11	3.31	94.23
PM-AHI$_3$	29.73	100.00	100.00	82.07	*	100.00	0.70	17.88	83.33
PM-AHI$_4$	50.00	99.58	97.37	86.50	119.00	97.37	0.50	13.46	87.82

*Infinite value.
AHI, apnoea–hypopnoea index; CC, correctly classified; −LR, negative likelihood ratio; +LR, positive likelihood ratio; NPV, negative predictive value; PM-AHI, portable monitors apnoea–hypopnoea index; PPV, positive predictive value; −PTP, negative post-test probability; +PTP, positive post-test probability; Sens, sensitivity; Spe, specificity.

population seen in clinical practice, we obtained results comparable to our stratified general population sample. Overall, we found a higher prevalence of sleep disordered breathing in our study sample than previously reported. As discussed in another article,[19] we believe that this is due to the use of more sensitive sensors (nasal pressure instead of thermistor and modern oximeters) and new scoring criteria compared with previous cohorts.

Our results show that the strongest correlation between PSG-AHI and type 3 PM was obtained using the PM-AHI$_1$ criteria, which define hypopnoea using 3% oxygen desaturation without PWA drops. According to the ROC curve analysis, PM-AHI$_1$ also showed the best discriminatory abilities for the diagnosis of OSA in almost all AHI categories. In the global population and in the 'high-risk' subgroup, PM-AHI$_1$ correctly classified most subjects when AHI thresholds of ≥5/h and ≥15/h were used. These criteria (PM-AHI$_1$) seem to be the best compromise between a high +LR and a low −LR and thus seem to

be the most reliable criteria to score PM recordings in a clinical setting. However, we cannot exclude that a more stringent definition of PWA drops (eg, 50% instead of 30% drop) or another surrogate for arousals that shows better correlation to EEG arousals could potentially yield a better agreement between PM and full PSG scoring.

Previous studies tried to use surrogate to EEG arousal to score PM recording. Masa et al[22] used breathing amplitude increases to estimate the presence of an arousal but found that it did not substantially increase the agreement between PM and PSG. In our study, PWA drop signal was used as a surrogate for EEG-defined arousals in the PM-AHI$_2$ and PM-AHI$_4$. Overall, the PWA drop index overestimated the EEG arousal index by a mean of 15.6 events/h. This can probably be explained by the important variations in PWA signal occurring during nocturnal wake periods included in PM-based PWA drop index but not in PSG-based arousal index. Moreover, non-respiratory stimulations of the autonomic system such as arousals due to noise,

Table 6 Diagnostic accuracy of PM-AHIs for AHI thresholds of ≥5/h, ≥15/h and ≥30/h in high-risk subjects

	Sens (%)	Spe (%)	PPV (%)	NPV (%)	LR+	PPTP (%)	LR−	NPTP (%)	CC (%)
AHI ≥5/h (n=93; pre-test probability=80.17%)									
PM-AHI$_1$	98.92	86.96	96.84	95.24	7.58	96.84	0.01	3.89	96.55
PM-AHI$_2$	100.00	26.09	84.55	100.00	1.35	84.52	0.00	0.00	85.34
PM-AHI$_3$	81.72	100.00	100.00	57.50	*	100.00	0.18	42.12	85.34
PM-AHI$_4$	100.00	47.83	88.57	100.00	1.92	88.59	0.00	0.00	89.66
AHI ≥15/h (n=63; pre-test probability=54.31%)									
PM-AHI$_1$	95.24	92.45	93.75	94.23	12.62	93.75	0.05	5.61	93.97
PM-AHI$_2$	100.00	64.15	76.83	100.00	2.79	76.83	0.00	0.00	83.62
PM-AHI$_3$	58.73	100.00	100.00	67.09	*	100.00	0.41	32.77	77.59
PM-AHI$_4$	80.95	98.11	98.08	81.25	42.90	98.08	0.19	18.42	88.79
AHI ≥30/h (n=32; pre-test probability=27.59%)									
PM-AHI$_1$	81.25	100.00	100.00	93.33	*	100.00	0.19	6.75	94.83
PM-AHI$_2$	96.88	95.24	88.57	98.77	20.34	88.57	0.03	1.13	95.69
PM-AHI$_3$	37.50	100.00	100.00	80.77	*	100.00	0.63	19.36	82.76
PM-AHI$_4$	62.50	98.81	95.24	87.37	52.50	95.24	0.38	12.65	88.79

*Infinite value.
AHI, apnoea–hypopnoea index; CC, correctly classified; −LR, negative likelihood ratio; +LR, positive likelihood ratio; NPTP, negative post-test probability; NPV, negative predictive value; PM-AHI, portable monitors apnoea–hypopnoea index; PPTP, positive post-test probability; PPV, positive predictive value; Sens, sensitivity; Spe, specificity.

pain or periodic leg movements can also generate PWA drops without an EEG arousal and without a SaO_2 drop

No single index derived from PSG recordings have shown to be a clear predictor of incident hypertension,[23] but previous studies have suggested that PWA drops are associated with increased blood pressure[24] and reflect subtle cortical activation even in the absence of EEG-defined arousals.[15] Even though adding PWA drops does not substantially increase the overall accuracy of PM, we cannot exclude that including this signal in the definition of respiratory events such as $PM-AHI_2$ (or $PM-AHI_4$) could be a more reliable predictor of incident cardiovascular and metabolic outcomes than standard AHI. However, only the long-term follow-up of this group of subjects will provide the answer to this question.

Several studies[25–32] analysed the correlation between PSG-AHI and AHIs obtained from different PM devices and showed good diagnostic agreements, but the comparison between these studies among them and with ours is challenging because of differences in the sensors used or the settings in which the studies took place (unattended at home or attended in-laboratory, simultaneous or separate recordings). In our study, we tried to limit the influence of these factors through a study design including a single PSG night recording at the subject's home using one single portable recorder device. We cannot exclude, however, that using a PM with a less-sensitive oximeter could yield a lower agreement rate with the PSG recording.[33]

There are also limitations in our study that need to be considered. First, the HypnoLaus cohort study, from which these subjects were drawn, included only subjects between 40 and 85 years of age. Our results can thus only apply to this age range and cannot be generalised to younger subjects. We believe, however, that the 40- to 85-year-old population is representative of standard patients referred sleep clinic patients for a suspicion of OSA. Second, only a subset of HypnoLaus cohort subjects (312) were re-analysed with the four sets of PM-AHI criteria. Still, since the concordance between the PM-AHI and the AHI was already highly significant, we do not believe that adding more subjects would change our conclusion. Third, our analysis is based on unattended home PSG with the recorder and sensors attached in the sleep laboratory, which may slightly differ from standard type 3 PMs that can be attached in the sleep laboratory or at home. Fourth, we did not perform an event-by-event comparison between PSG and PM. It is thus possible that not the same events were identified with the two scoring methods. Despite these possible differences, the overall clinical classifications showed a good agreement. Last, we excluded subjects using β-blockers as these medications are believed to influence the PWA signal. Consequently, our results cannot be generalised to this population.

Overall this study suggests that, despite differences in the measured parameters, PM can provide a reliable evaluation of OSA severity in most patients. Recent studies further evaluated the place of PM in diagnostic algorithms[9 34–37] and suggested that PM devices could be used in current practice.

CONCLUSION

These results show that, compared with home PSG-AHI, PM-derived AHI values using standard criteria (3% desaturation without PWA drops) and self-reported TRT can correctly classify patients with OSA with an accuracy of >93%. Criteria including PWA drops as a surrogate for EEG arousal for the scoring of hypopnoea showed a higher sensitivity only for the identification of severe OSA subjects but do not seem to substantially increase the overall accuracy of PM.

REFERENCES

1 Verstraeten E. Neurocognitive effects of obstructive sleep apnea syndrome. *Curr Neurol Neurosci Rep* 2007;7:161–6.
2 Bradley TD, Floras JS. Obstructive sleep apnoea and its cardiovascular consequences. *Lancet* 2009;373:82–93.
3 Levy P, Bonsignore MR, Eckel J. Sleep, sleep-disordered breathing and metabolic consequences. *Eur Respir J* 2009;34:243–60.
4 Young T, Evans L, Finn L, et al. Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. *Sleep* 1997;20:705–6.
5 Berry RB, Budhiraja R, Gottlieb DJ, et al. 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.
6 Collop NA, Anderson WM, Boehlecke B, et al. Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients. Portable Monitoring Task Force of the American Academy of Sleep Medicine. *J Clin Sleep Med* 2007;3:737–47.
7 Kuna ST, Gurubhagavatula I, Maislin G, et al. Noninferiority of functional outcome in ambulatory management of obstructive sleep apnea. *Am J Respir Crit Care Med* 2011;183:1238–44.
8 Masa JF, Corral J, Pereira R, et al. Effectiveness of home respiratory polygraphy for the diagnosis of sleep apnoea and hypopnoea syndrome. *Thorax* 2011;66:567–73.
9 Berry RB, Hill G, Thompson L, et al. Portable monitoring and autotitration versus polysomnography for the diagnosis and treatment of sleep apnea. *Sleep* 2008;31:1423–31.
10 Santos-Silva R, Sartori DE, Truksinas V, et al. Validation of a portable monitoring system for the diagnosis of obstructive sleep apnea syndrome. *Sleep* 2009;32:629–36.
11 Corral-Penafiel J, Pepin JL, Barbe F. Ambulatory monitoring in the diagnosis and management of obstructive sleep apnoea syndrome. *Eur Respir Rev* 2013;22:312–24.
12 Epstein LJ, Kristo D, Strollo PJ Jr, et al. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. *J Clin Sleep Med* 2009;5:263–76.
13 Collop NA, Tracy SL, Kapur V, et al. Obstructive sleep apnea devices for out-of-center (OOC) testing: technology evaluation. *J Clin Sleep Med* 2011;7:531–48.
14 Iber C, American Academy of Sleep Medicine. *The AASM manual for the scoring of sleep and associated events: rules, terminology, and technical specifications*. Westchester, IL: American Academy of Sleep Medicine, 2007. 59pp.
15 Delessert A, Espa F, Rossetti A, et al. Pulse wave amplitude drops during sleep are reliable surrogate markers of changes in cortical activity. *Sleep* 2010;33:1687–92.
16 Adler D, Bridevaux PO, Contal O, et al. Pulse wave amplitude reduction: A surrogate marker of micro-arousals associated with respiratory events occurring under non-invasive ventilation? *Respir Med* 2013;107:2053–60.
17 Zacharia A, Haba-Rubio J, Simon R, et al. Sleep apnea syndrome: improved detection of respiratory events and cortical arousals using oxymetry pulse wave amplitude during polysomnography. *Sleep Breath* 2008;12:33–8.
18 Haba-Rubio J, Darbellay G, Herrmann FR, et al. Obstructive sleep apnea syndrome: effect of respiratory events and arousal on pulse wave amplitude measured by photoplethysmography in NREM sleep. *Sleep Breath* 2005;9:73–81.
19 Heinzer R, Vat S, Marques-Vidal P, et al. Prevalence of sleep-disordered breathing in the general population: the HypnoLaus study. *Lancet Respir Med* 2015;3:310–18.
20 Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 1991;14:540–5.
21 Farney RJ, Walker BS, Farney RM, et al. The STOP-Bang equivalent model and prediction of severity of obstructive sleep apnea: relation to polysomnographic measurements of the apnea/hypopnea index. *J Clin Sleep Med* 2011;7:459–65B.

22 Masa JF, Corral J, Sanchez de Cos J, *et al*. Effectiveness of three sleep apnea management alternatives. *Sleep* 2013;36:1799–807.

23 Redline S, Min NI, Shahar E, *et al*. Polysomnographic predictors of blood pressure and hypertension: is one index best? *Sleep* 2005;28:1122–30.

24 Zou D, Grote L, Radlinski J, *et al*. Nocturnal pulse wave attenuation is associated with office blood pressure in a population based cohort. *Sleep Med* 2009;10:836–43.

25 El Shayeb M, Topfer LA, Stafinski T, *et al*. Diagnostic accuracy of level 3 portable sleep tests versus level 1 polysomnography for sleep-disordered breathing: a systematic review and meta-analysis. *CMAJ* 2014;186:E25–51.

26 Ferber R, Millman R, Coppola M, *et al*. Portable recording in the assessment of obstructive sleep apnea. ASDA standards of practice. *Sleep* 1994;17:378–92.

27 Ahmed M, Patel NP, Rosen I. Portable monitors in the diagnosis of obstructive sleep apnea. *Chest* 2007;132:1672–7.

28 Flemons WW, Littner MR, Rowley JA, *et al*. Home diagnosis of sleep apnea: a systematic review of the literature. An evidence review cosponsored by the American Academy of Sleep Medicine, the American College of Chest Physicians, and the American Thoracic Society. *Chest* 2003;124:1543–79.

29 Chesson AL Jr., Berry RB, Pack A, American Academy of Sleep M, American Thoracic S, American College of Chest P. Practice parameters for the use of portable monitoring devices in the investigation of suspected obstructive sleep apnea in adults. *Sleep* 2003;26:907–13.

30 Ayappa I, Norman RG, Seelall V, *et al*. Validation of a self-applied unattended monitor for sleep disordered breathing. *J Clin Sleep Med* 2008;4:26–37.

31 Tonelli de Oliveira AC, Martinez D, Vasconcelos LF, *et al*. Diagnosis of obstructive sleep apnea syndrome and its outcomes with home portable monitoring. *Chest* 2009;135:330–6.

32 Driver HS, Pereira EJ, Bjerring K, *et al*. Validation of the MediByte(R) type 3 portable monitor compared with polysomnography for screening of obstructive sleep apnea. *Can Respir J* 2011;18:137–43.

33 Zafar S, Ayappa I, Norman RG, *et al*. Choice of oximeter affects apnea-hypopnea index. *Chest* 2005;127:80–8.

34 Mulgrew AT, Fox N, Ayas NT, *et al*. Diagnosis and initial management of obstructive sleep apnea without polysomnography: a randomized validation study. *Ann Intern Med* 2007;146:157–66.

35 Antic NA, Buchan C, Esterman A, *et al*. A randomized controlled trial of nurse-led care for symptomatic moderate-severe obstructive sleep apnea. *Am J Respir Crit Care Med* 2009;179:501–8.

36 Skomro RP, Gjevre EJ, Reid J, *et al*. Outcomes of home-based diagnosis and treatment of obstructive sleep apnea. *Chest* 2010;138:257–63.

37 Rosen CL, Auckley D, Benca R, *et al*. A multisite randomized trial of portable sleep studies and positive airway pressure autotitration versus laboratory-based polysomnography for the diagnosis and treatment of obstructive sleep apnea: the HomePAP study. *Sleep* 2012;35:757–67.

Figure S1