



Original Investigation | Pulmonary Medicine

Assessment of Mandibular Movement Monitoring With Machine Learning Analysis for the Diagnosis of Obstructive Sleep Apnea

Jean-Louis Pépin, MD, PhD; Clément Letesson, PhD; Nhat Nam Le-Dong, MD, PhD; Antoine Dedave, MSc; Stéphane Denison, MSc; Valérie Cuthbert, MSc; Jean-Benoît Martinot, MD; David Gozal, MD, MBA

Abstract

IMPORTANCE Given the high prevalence of obstructive sleep apnea (OSA), there is a need for simpler and automated diagnostic approaches.

OBJECTIVE To evaluate whether mandibular movement (MM) monitoring during sleep coupled with an automated analysis by machine learning is appropriate for OSA diagnosis.

DESIGN, SETTING, AND PARTICIPANTS Diagnostic study of adults undergoing overnight in-laboratory polysomnography (PSG) as the reference method compared with simultaneous MM monitoring at a sleep clinic in an academic institution (Sleep Laboratory, Centre Hospitalier Universitaire Université Catholique de Louvain Namur Site Sainte-Elisabeth, Namur, Belgium). Patients with suspected OSA were enrolled from July 5, 2017, to October 31, 2018.

MAIN OUTCOMES AND MEASURES Obstructive sleep apnea diagnosis required either evoking signs or symptoms or related medical or psychiatric comorbidities coupled with a PSG-derived respiratory disturbance index (PSG-RDI) of at least 5 events/h. A PSG-RDI of at least 15 events/h satisfied the diagnosis criteria even in the absence of associated symptoms or comorbidities. Patients who did not meet these criteria were classified as not having OSA. Agreement analysis and diagnostic performance were assessed by Bland-Altman plot comparing PSG-RDI and the Sunrise system RDI (Sr-RDI) with diagnosis threshold optimization via receiver operating characteristic curves, allowing for evaluation of the device sensitivity and specificity in detecting OSA at 5 events/h and 15 events/h.

RESULTS Among 376 consecutive adults with suspected OSA, the mean (SD) age was 49.7 (13.2) years, the mean (SD) body mass index was 31.0 (7.1), and 207 (55.1%) were men. Reliable agreement was found between PSG-RDI and Sr-RDI in patients without OSA ($n = 46$; mean difference, 1.31; 95% CI, -1.05 to 3.66 events/h) and in patients with OSA with a PSG-RDI of at least 5 events/h with symptoms ($n = 107$; mean difference, -0.69; 95% CI, -3.77 to 2.38 events/h). An Sr-RDI underestimation of -11.74 (95% CI, -20.83 to -2.67) events/h in patients with OSA with a PSG-RDI of at least 15 events/h was detected and corrected by optimization of the Sunrise system diagnostic threshold. The Sr-RDI showed diagnostic capability, with areas under the receiver operating characteristic curve of 0.95 (95% CI, 0.92-0.96) and 0.93 (95% CI, 0.90-0.93) for corresponding PSG-RDIs of 5 events/h and 15 events/h, respectively. At the 2 optimal cutoffs of 7.63 events/h and 12.65 events/h, Sr-RDI had accuracy of 0.92 (95% CI, 0.90-0.94) and 0.88 (95% CI, 0.86-0.90) as well as posttest probabilities of 0.99 (95% CI, 0.99-0.99) and 0.89 (95% CI, 0.88-0.91) at PSG-RDIs of at least 5 events/h and at least 15 events/h, respectively, corresponding to positive likelihood ratios of 14.86 (95% CI, 9.86-30.12) and 5.63 (95% CI, 4.92-7.27), respectively.

(continued)

Key Points

Question How does the performance of an automated mandibular movement analysis for the diagnosis of obstructive sleep apnea compare with that of polysomnography?

Findings In this diagnostic study of 376 adults with suspected obstructive sleep apnea, the mandibular movement-derived respiratory disturbance index identified patients with a polysomnography respiratory disturbance index of at least 5 events/h or at least 15 events/h with accuracy of 0.92 and 0.88, respectively.

Meaning Automatic analysis of mandibular movement patterns reliably calculated respiratory disturbance index, and the use of this approach in obstructive sleep apnea diagnosis appears to be promising.

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

Open Access. This is an open access article distributed under the terms of the CC-BY-NC-ND License.

Abstract (continued)

CONCLUSIONS AND RELEVANCE Automatic analysis of MM patterns provided reliable performance in RDI calculation. The use of this index in OSA diagnosis appears to be promising.

JAMA Network Open. 2020;3(1):e1919657. doi:10.1001/jamanetworkopen.2019.19657

Introduction

Obstructive sleep apnea (OSA) affects almost 1 billion people worldwide, resulting in high socioeconomic and health care burden.¹ Excessive daytime sleepiness and fatigue, the chief problems reported by patients with OSA, may have negative consequences on neurocognitive function, mood, and productivity at work, leading to decreased quality of life and increased risk of occupational injuries and motor vehicle crashes.²⁻⁵ Obstructive sleep apnea is also a major risk factor for a variety of medical conditions, increasing the risk and severity of cardiometabolic diseases, including hypertension, arrhythmias, stroke, coronary heart disease, type 1 and type 2 diabetes, and metabolic dysfunction,⁶⁻⁸ ultimately resulting in increased overall mortality.⁹ Continuous positive airway pressure, the first-line therapy for OSA, is effective in alleviating symptoms, restoring neurocognitive function, and improving quality of life.^{10,11} Although OSA is one of the most prevalent chronic diseases associated with a wide range of disabilities, it remains an underdiagnosed health problem.^{1,12-14}

Polysomnography (PSG), the reference method and criterion-standard diagnostic tool for OSA, is unsuitable for the widespread use required to address the sleep apnea epidemic. The performance of PSG is onerous because of the complexity of implementation and the time-consuming and laborious scoring of multichannel recordings, including electroencephalogram, electrocardiogram, and respiratory signals.^{15,16} The manual scoring of sleep stages, microarousals, and respiratory events in a specialized clinic is tedious, technically challenging, and not suitable to the growing population in need of OSA evaluation.^{17,18}

There is consensus among experts on the need to develop advanced diagnostic approaches that incorporate novel technologies providing valid surrogates for sleep staging and respiratory pattern evaluation.^{19,20} Therefore, artificial intelligence-driven sensors that automatically score sleep and respiratory events might be a future direction for OSA diagnosis.²¹⁻²³

In this context, the objective of this study was to assess the diagnostic capabilities of a novel technology predicated on mandibular movement (MM) analysis (Sunrise) compared with PSG in a large population of consecutive patients with suspected OSA. The Sunrise system represents a technological advancement that combines MM recordings with an automated analysis that is supported by machine learning. Mandibular movement analysis has the advantage of being a reliable marker of sleep fragmentation and respiratory effort, providing information on the obstructive nature of respiratory events during sleep.²⁴⁻²⁷ We hypothesized that the Sunrise system-derived respiratory disturbance index (Sr-RDI) would compare favorably with the PSG-derived RDI (PSG-RDI).

Methods

Study Design

In this prospective, diagnostic study of adult patients who were referred for a single overnight in-laboratory PSG, the PSG was used as the reference method and, with blinding, was compared with simultaneous MM recordings using the Sunrise system (a full description is given in the MM Recordings and Description of the Sunrise System subsection). The study was conducted at a sleep clinic at an academic institution (Sleep Laboratory, Centre Hospitalier Universitaire Université Catholique de Louvain Namur Site Sainte-Elisabeth, Namur, Belgium). The Comité d'Éthique

Hospitalo-Facultaire-Universitaire de Liège approved the study, and each participant provided written informed consent. This study followed the Standards for Reporting of Diagnostic Accuracy (STARD) reporting guideline.

Overnight Sleep Study

In-laboratory PSG was recorded with a digital acquisition system (Somnoscreen Plus; Somnomedics). The parameters monitored included electroencephalogram (Fz–A+, Cz–A+, and Pz–A+), right and left electro-oculogram, submental electromyogram, tibial electromyogram, chest and abdominal wall motion by respiratory inductance plethysmography (SleepSense; S.L.P. Inc), and nasal and oral flows with a pressure transducer and a thermistor, respectively, as well as oxygen saturation by digital oximeter displaying pulse waveform (Nonin; Nonin Medical).

The PSG data were manually scored by 2 experienced investigators (V.C. and a nonauthor) who were blinded to the identity of the patients. All sleep stages, electroencephalogram arousals, and sleep-related respiratory events were visually scored in accordance with the recommended criteria established by the American Academy of Sleep Medicine (AASM) *Manual for the Scoring of Sleep and Associated Events*.²⁸ Hypopneas were scored using the AASM-recommended hypopnea 1A definition, requiring at least a 30% decrement in airflow lasting 10 seconds or longer and associated with a decrease of at least 3% in oxygen saturation as measured by pulse oximetry or an arousal.^{28,29} Obstructive sleep apnea diagnosis was established according to the third edition of the *International Classification of Sleep Disorders*³⁰ (ICSD-3) and required either evoking signs or symptoms or related medical or psychiatric comorbidities coupled with at least 5 predominantly obstructive respiratory events (ie, obstructive and mixed apneas, hypopneas, or respiratory effort-related arousals [hereafter referred to as PSG-RDI]) per hour of sleep during PSG. Alternatively, a frequency of obstructive respiratory events of at least 15 events/h satisfied the diagnosis criteria even in the absence of associated symptoms or comorbidities.³¹ Interobserver agreement for scoring PSG was evaluated for all studies by intraclass correlation coefficient (ICC) using a 2-way random model for single measures (ICC, 2,1). Interobserver agreement (ICC, 2,1) for PSG double-scoring was 92% (95% CI, 89%-94%; $P < .001$) in our laboratory. A prespecified, clinically relevant RDI difference of 3 events/h (with an upper bound at 5 events/h) between PSG-RDI and Sr-RDI was agreed on a priori by the study investigators (J-L.P., J-B.M., and D.G.).

MM Recordings and Description of the Sunrise System

The Sunrise system is composed of coin-sized hardware attached by the sleep technician to the chin of the patient in the mentolabial sulcus. Its embedded inertial measurement unit enables MM sensing and communicates with a smartphone application for external control. The collected MM data were automatically transferred to a cloud-based infrastructure at the end of the night, and data analysis was conducted with a dedicated machine learning algorithm (**Figure 1**). Details on the algorithm and the independent sample in which it was developed are provided in eFigure 2, eTable 1, and eTable 2 in the [Supplement](#).

The Sunrise algorithm allowed for automatic identification of obstructive and mixed apneas and hypopneas or respiratory effort-related arousals through stereotypical MM patterns (eFigure 1 in the [Supplement](#)). For that purpose, the algorithm automatically processed MM signal components and assessed whether MM patterns could be classified as wake, arousal, respiratory effort, or quiet sleep. To identify wake, the algorithm tested whether MM signals were fast, irregular, and nonpredictable.^{25,26} For the identification of arousal movements, the algorithm detected brisk MM of large amplitude, indicating abrupt closure of the mouth characteristic of arousals.²⁷ Respiratory effort was identified through oscillating MM at the breathing frequency.²⁴ The Sunrise algorithm identifies respiratory disturbances as a period of respiratory effort ended by an arousal or an awakening. The Sr-RDI consists of the total number of respiratory disturbances accompanied by respiratory effort divided by the total sleep time (TST), which is estimated from the Sunrise analytics.

Statistical Analysis

Data analysis was conducted using R statistical programming language (R Project for Statistical Computing).³² The analysis focused on evaluating the agreement between MM-derived parameters (Sr-RDI, Sr-TST, and arousal index [Sr-ArI]) and their PSG counterparts (PSG-RDI, PSG-TST, and PSG-ArI) as well as optimizing the clinical performance of the RDI derived from Sunrise system analysis in ruling in a diagnosis of OSA at the 2 reference thresholds of PSG of at least 5 events/h or at least 15 events/h, leading to the classification of participants in the following 3 clinical groups: non-OSA (OSA ruled out), PSG-RDI of at least 5 events/h (OSA ruled in with PSG-RDI $\geq$ 5 events/h and the patient had either signs or symptoms or a comorbid medical or psychiatric disorder), and PSG-RDI of at least 15 events/h (OSA is ruled in with PSG-RDI $\geq$ 15 events/h even in the absence of associated symptoms or comorbid disorders).

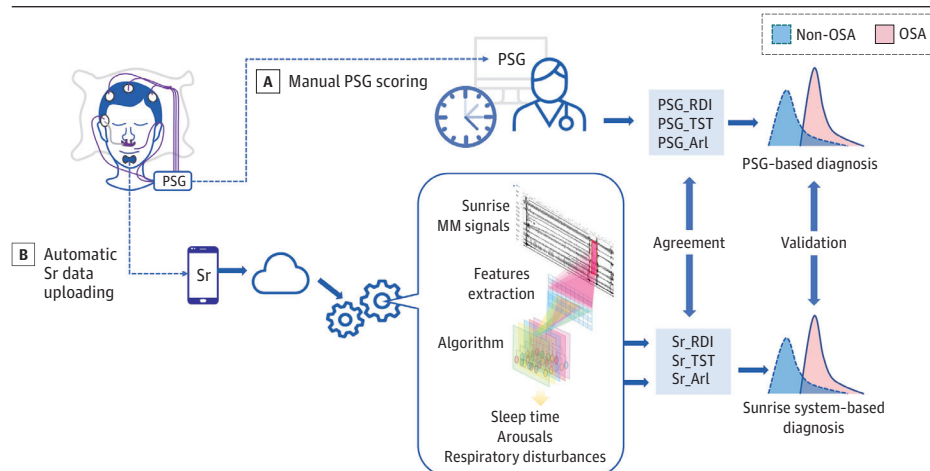
A complete and groupwise Bland-Altman plot³³ was prespecified to estimate the 95% limits of agreement and the systematic bias of MM-derived indexes compared with their PSG counterparts. Estimations of the mean differences and 95% limits of agreement were based on the individual random errors extracted from a mixed model.³⁴ Receiver operating characteristic (ROC) curve analysis was used to evaluate the overall clinical effectiveness of the new diagnostic tool via area under the curve (AUC), and a post hoc analysis was performed to optimize the cutoff points of Sr-RDI for diagnostic decisions compared with the criterion-standard cutoff values of obstructive PSG-RDI recommended in *ICSD-3* (5 events/h and 15 events/h). The optimal MM cutoffs were assessed at the highest value of the Youden index (sensitivity plus specificity minus 1). The metrics of clinical utility and accuracy were also calculated for the optimal detection thresholds. The posttest probability for each cutoff point was also calculated as recommended by Collop et al.³⁵ $P < .001$ was considered statistically significant, and all tests were 2-tailed.

Results

In total, 376 consecutive adults with suspected OSA were enrolled from July 5, 2017, to October 31, 2018. Their mean (SD) age was 49.7 (13.2) years, their mean (SD) body mass index (calculated as weight in kilograms divided by height in meters squared) was 31.0 (7.1), and 207 (55.1%) were men. Details on the sample size calculation are provided in the eAppendix in the [Supplement](#).

The final data set included all 376 patients recruited and was stratified into 3 clinical groups according to *ICSD-3* diagnosis rules. Enrolled patients represented a clinical referral population in which the pretest probabilities were 82.5% and 55.7% for PSG-RDIs of at least 5 events/h and at least 15 events/h, respectively. No technical failures occurred with the use of the Sunrise system.

Figure 1. Flow Diagram of the Study Protocol



Shown is a comparison between polysomnography (PSG) and automated mandibular movement (MM) analysis procedures. Data concomitantly recorded by in-laboratory PSG and the Sunrise system device were analyzed independently. A, The PSG data were manually scored to export a respiratory disturbance index (PSG-RDI) as the reference method for obstructive sleep apnea (OSA) diagnosis. B, The Sunrise system (Sr) data were automatically uploaded into a cloud-based platform without human intervention, where data were handled by a proprietary machine learning algorithm. After algorithm processing, Sr-RDI was automatically derived for agreement analysis and evaluation of diagnosis performance. ArI indicates arousal index; TST, total sleep time.

The characteristics of the cohort are listed in **Table 1**. The cohort had a median PSG-RDI of 18.80 (interquartile range, 7.80-29.80) events/h and a median sleep duration of 7.20 (interquartile range, 6.40-8.00) hours.

Evaluation of the Agreement Between the Sunrise System and PSG for Measuring Sleep Apnea Indexes

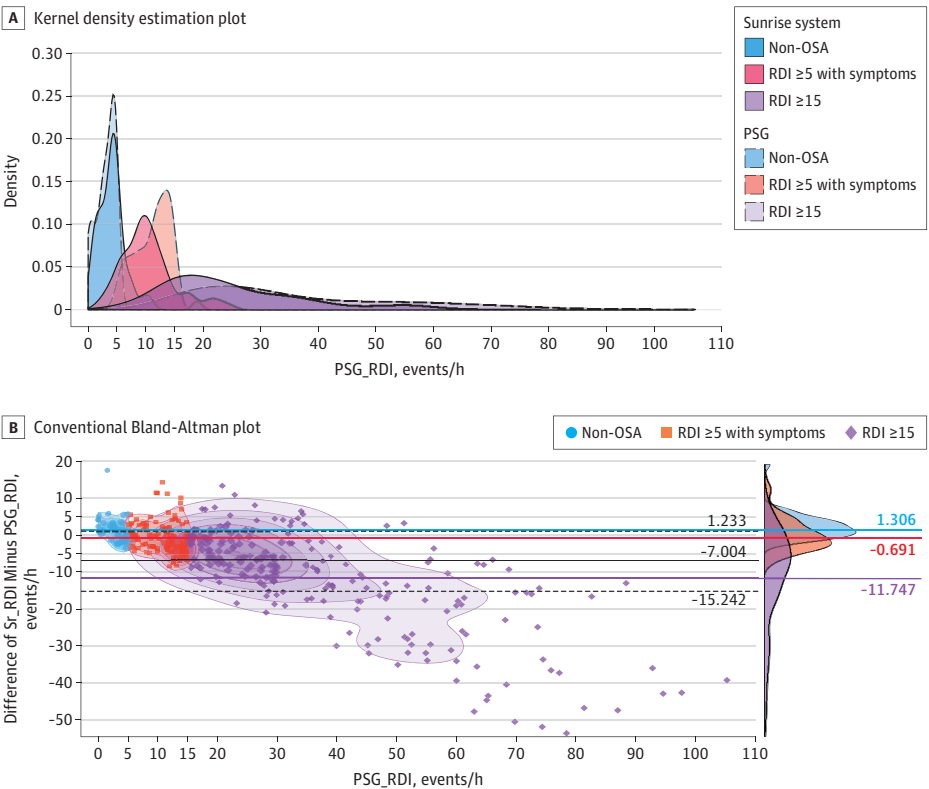
An extended Bland-Altman plot (**Figure 2**) was used to evaluate the agreement level between the Sunrise system-based approach and PSG (reference method) for RDI measurement and to anticipate

Table 1. Characteristics of the Study Population of Adults With Suspected OSA Undergoing Overnight In-Laboratory PSG

Characteristic	Median (Interquartile Range)		
	Non-OSA (n = 46)	PSG-RDI ≥5 events/h With Symptoms (n = 107)	PSG-RDI ≥15 events/h (n = 223)
Age, y	38.3 (33.7-47.5)	45.6 (36.5-55.7)	52.6 (43.6-61.5)
Height, cm	170 (165-175)	172 (163-178)	174 (167-180)
Weight, kg	71 (63-90)	87 (75-103)	95 (82-106)
Neck circumference, cm	37 (35-39)	39 (37-42)	42 (39-45)
BMI	23.5 (21.3-30.1)	28.7 (25.6-36.8)	31.3 (27.4-35.5)
ESS score	11 (8-15)	9 (6-14)	12 (7-15)
PSG-ArI, events/h	7.85 (5.96-10.80)	13.30 (11.00-16.90)	28.50 (20.10-41.50)
PSG-ODI, events/h	2.25 (0.63-4.05)	6.50 (2.85-12.20)	31.10 (16.20-51.60)
PSG-RDI, events/h	3.35 (2.08-4.68)	11.80 (8.79-13.88)	29.70 (21.71-49.46)
PSG-AHI, events/h	3.58 (2.42-4.57)	9.50 (5.68-13.80)	31.20 (21.58-51.40)
PSG-TST, h	7.20 (6.52-7.75)	7.20 (6.45-8.00)	7.10 (6.20-7.95)

Abbreviations: AHI, apnea/hypopnea hourly index; ArI, arousal index; BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); ESS, Epworth Sleepiness Scale; ODI, oxygen desaturation index (3%); OSA, obstructive sleep apnea; PSG, polysomnography; RDI, respiratory disturbance index; TST, total sleep time.

Figure 2. Evaluation of the Agreement Between the 2 Methods of Respiratory Disturbance Index (RDI) Measurement for Obstructive Sleep Apnea (OSA) Diagnosis



The reference method of overnight in-laboratory polysomnography (PSG) is shown on the x-axis. A, Kernel density estimation plot shows the distribution of PSG-derived RDI (PSG-RDI) (discontinued trace) vs the Sunrise system RDI (Sr-RDI) (continuous trace) in the 3 clinical groups. B, Conventional Bland-Altman plot shows the disagreement between PSG-RDI and Sr-RDI (y-axis) as a function of PSG-RDI (x-axis), with individual cases stratified into 3 clinical groups. The horizontal lines indicate the mean difference in the whole sample and within each group. The 2 dashed lines indicate the lower and upper levels (mean, ± 1.96 SD) of the disagreement in the whole sample. Bidimensional kernel density estimation plots are superposed to show the distribution of the disagreement as a function of PSG-RDI. The distribution of the disagreement between the 2 methods, stratified by group, is shown on the right.

the reliability of switching from PSG to the Sunrise system as a diagnostic tool. Overall, the Bland-Altman plot (Figure 2B) showed systematic bias between the 2 methods, with a mean difference of -7.00 events/h (95% CI, -15.24 to 1.23 events/h).

Detailed analysis (Figure 2B) showed that in the non-OSA group ($n = 46$) the disagreement was normally distributed and below a clinically relevant difference (mean difference, 1.31 ; 95% CI, -1.05 to 3.66 events/h, thus including 0). The disagreement was also normally distributed in the OSA group with a PSG-RDI of at least 5 events/h ($n = 107$), and the mean difference was low (-0.69 ; 95% CI, -3.77 to 2.38 events/h) and included 0. In contrast, the distribution of disagreement in the group of patients diagnosed as having a PSG-RDI of at least 15 events/h became positively skewed because of outliers at values exceeding 40 events/h on the PSG-RDI scale. The mean difference was estimated at -11.74 (95% CI, -20.83 to -2.67) events/h in the OSA group with a PSG-RDI of at least 15 events/h.

The RDI distribution curves (Figure 2A) were uniform in the 3 clinical groups irrespective of the measurement method examined. The distribution of Sr-RDI was slightly moved to the left compared with PSG-RDI in the 2 OSA groups but almost overlapped within the non-OSA group, indicating that the systematic bias between PSG-derived and Sunrise system-derived indexes could be corrected by relocating the diagnostic cutoffs on the Sr-RDI scale using ROC curve analysis.

The agreement between PSG-derived and Sunrise system-derived indexes for TST and ArI is shown in Bland-Altman plots in eFigure 3 and eFigure 4 in the [Supplement](#). The difference in TST between the 2 methods contributed to less than 1 unit of the difference between RDIs.

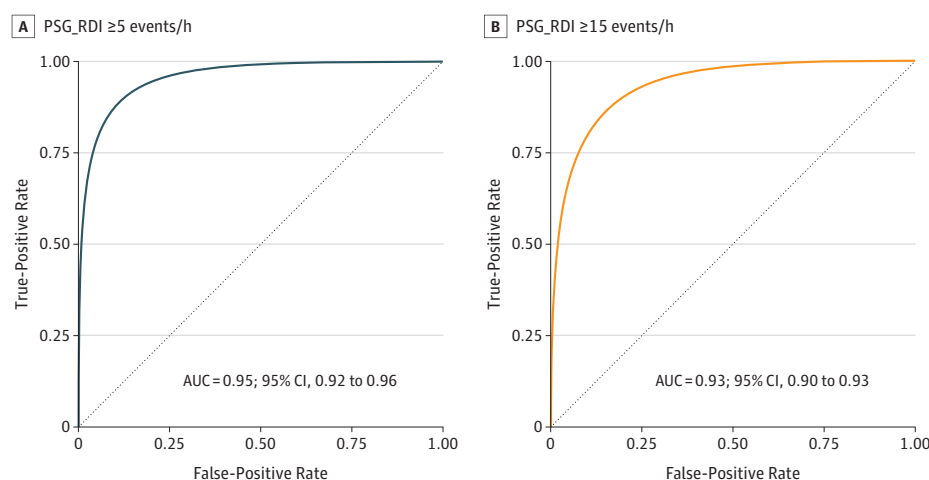
Optimization of Sr-RDI Diagnostic Performance in OSA by ROC Curve Analysis

The variability of Sr-RDI among the 3 clinical groups was explored using pairwise comparisons in a Gardner-Altman plot. The mean differences in Sr-RDI between patients with and without OSA diagnosed at the cutoffs of 5 events/h and 15 events/h were 6.11 (95% CI, 4.59 - 7.61 ; $P < .001$) and 21.20 (95% CI, 18.87 - 23.52 ; $P < .001$), respectively (eFigure 5 in the [Supplement](#)).

An ROC curve analysis was performed to evaluate the clinical performance of Sr-RDI diagnostic rules to detect PSG-defined OSA at the 2 clinical thresholds of 5 events/h and 15 events/h (**Figure 3**). The AUCs of the 2 Sr-RDI binary classification rules targeting PSG-RDI of at least 5 events/h or at least 15 events/h were 0.95 (95% CI, 0.92 - 0.96) and 0.93 (95% CI, 0.90 - 0.93), respectively. The variability of performance metrics over all possible cutoffs on the Sr-RDI scale is shown in eFigure 6 in the [Supplement](#).

At the 2 optimal cutoffs of 7.63 events/h and 12.65 events/h, Sr-RDI detected patients with PSG-RDI of at least 5 events/h or at least 15 events/h with high accuracy (0.92 [95% CI, 0.90 - 0.94]

Figure 3. Receiver Operating Characteristic Curve Analysis for Evaluating the Performance of the Sunrise System Respiratory Disturbance Index (Sr-RDI) in Obstructive Sleep Apnea Diagnosis



Shown are curves of the binary classification rules to detect patients with obstructive sleep apnea with polysomnography-derived respiratory disturbance index (PSG-RDI) of at least 5 events/h (A) and at least 15 events/h (B) using Sr-RDI. The 95% CIs of the area under the curve (AUC) and smoothing effect were obtained by bootstrapping. The diagonal dotted line serves as a reference and shows the performance if obstructive sleep apnea detection was made randomly.

and 0.88 [95% CI, 0.86-0.90], respectively) and good balance between precision and recall (F1 score [the harmonic mean between precision and recall], 0.95 [95% CI, 0.94-0.97] and 0.91 [95% CI, 0.89-0.92], respectively). The Sr-RDI diagnostic rules also showed superior diagnostic performance compared with baseline self-reported symptoms, with high posttest probabilities of obtaining a true-positive diagnosis (0.99 [95% CI, 0.99-0.99] for patients with PSG-RDI ≥ 5 and 0.89 [95% CI, 0.88-0.91] for patients with PSG-RDI ≥ 15 , corresponding to positive likelihood ratios of 14.86 [95% CI, 9.86-30.12] and 5.63 [95% CI, 4.92-7.27], respectively) (**Table 2**).

Discussion

In a large, prospective cohort of patients with and without OSA, we evaluated the agreement between MM-derived Sr-RDI and blindly scored PSG-RDI (Figure 2B). The maximum agreement between PSG-RDI and Sr-RDI was found in patients without OSA and in those with OSA with PSG-RDI of at least 5 events/h, with the latter group representing the most challenging population to be diagnosed when using simplified approaches.³⁵

Diagnostic Capabilities of the Sunrise System vs PSG

The study sample was representative of a typical clinical referral population, including patients with low to high pretest probability of OSA, thus embracing the entire spectrum of OSA. This factor is clinically relevant because simple and less onerous diagnostic strategies are particularly desirable in asymptomatic or specific at-risk populations, such as those with cardiometabolic disease or obesity.³⁶ Similar skewed RDI distributions were found for PSG and the Sunrise system irrespective of the group of patients being considered (Figure 2A). Accordingly, the optimal diagnostic cutoff was adjusted, and the diagnostic PSG-RDI cutoffs of at least 5 events/h and at least 15 events/h could be extrapolated with confidence to Sr-RDI cutoffs of at least 7.63 events/h and at least 12.65 events/h. The Sunrise system diagnostic performance after cutoff optimization compared favorably with PSG, with ROC curves showing high AUCs of 0.95 and 0.93, respectively. At these cutoffs, the mean positive likelihood ratios were 14.86 and 5.63, respectively, and led to posttest probabilities of obtaining a true-positive diagnosis of 99% and 89%, respectively. These results are in line with the recommendations by Collop et al,³⁵ who stated that to efficiently rule in a diagnosis of OSA, portable monitoring devices should improve the pretest probability to a sufficiently high posttest probability.

Sleep apnea is a common chronic disease associated with significant deterioration in quality of life. Continuous positive airway pressure is effective in symptomatic patients and, if adherently used,

Table 2. Performance of the Sr-RDI to Detect Patients With PSG-RDI at the Diagnostic Levels Reported in the Third Edition of the *International Classification of Sleep Disorders*^{30a}

Variable	Value (95% CI)	
	PSG-RDI ≥ 5 Events/h	PSG-RDI ≥ 15 Events/h
Sr-RDI cutoff	7.63	12.65
Youden index	0.84	0.76
AUC	0.95 (0.92-0.96)	0.93 (0.90-0.93)
Balanced accuracy	0.92 (0.90-0.94)	0.88 (0.86-0.90)
Sensitivity	0.91 (0.89-0.92)	0.92 (0.90-0.94)
Specificity	0.94 (0.91-0.97)	0.84 (0.81-0.87)
False-positive rate	0.06 (0.03-0.09)	0.16 (0.13-0.19)
False-negative rate	0.09 (0.08-0.11)	0.08 (0.06-0.10)
Positive likelihood ratio	14.86 (9.86-30.12)	5.63 (4.92-7.27)
Negative likelihood ratio	0.10 (0.08-0.12)	0.10 (0.07-0.12)
Positive predictive value	0.99 (0.99-0.99)	0.89 (0.88-0.91)
Negative predictive value	0.59 (0.55-0.63)	0.88 (0.85-0.91)
F1 score	0.95 (0.94-0.97)	0.91 (0.89-0.92)

Abbreviations: AUC, area under the curve; PSG-RDI, polysomnography-derived respiratory disturbance index; Sr-RDI, Sunrise system RDI.

^a Optimal cutoff points were assessed at the highest value of the Youden index (sensitivity plus specificity minus 1). The F1 score is the harmonic mean between precision and recall. The 95% CIs were obtained by bootstrapping.

may reduce the magnitude of disabling symptoms and restore quality of life, social functioning, and work productivity. Under such circumstances, a diagnostic tool should be sensitive to detect patients across the full spectrum of disease severity. Therefore, our main challenge was to assess the diagnostic thresholds for MM recordings that would differentiate between patients with OSA of varying severities and patients without OSA. To optimize the delineation of such diagnostic thresholds, we used ROC curves and defined the trade-off between true-positive rates and false-positive rates at PSG-RDI of at least 5 events/h and at least 15 events/h.

Sunrise System and High-Performance Sleep Apnea Diagnosis

Mandibular movement analysis has been validated as an accurate measurement modality of respiratory effort during sleep, providing information on the obstructive or central nature of respiratory events.²⁴ Compared with other simplified diagnostic techniques, MM also provides reliable estimation of TST and unbiased identification of microarousals.^{26,27} Correct TST calculation prevents RDI underestimation (relative to the total recording time), and microarousal identification ensures recognition of respiratory effort-related arousals. These capabilities are uncommon among existing portable monitoring devices³⁷ and contribute to the good agreement between PSG-RDI and Sr-RDI.

Persistent Discrepancies in Patients With Severe OSA

We observed systematic underestimation by the Sunrise system compared with PSG for PSG-RDI exceeding 40 events/h. The underestimation might be attributable to the use of the AASM-recommended hypopnea 1A definition. Using this rule, hypopneas can be scored when airflow reduction is followed either by a 3% oxygen desaturation or an arousal from sleep.²⁸ Whereas cortical arousals are clearly detected through the occurrence of brisk and abrupt MM, hypopnea events that were only scored because of the presence of an isolated 3% oxygen desaturation could have been overlooked by the Sunrise system. Ongoing algorithmic developments will address this issue in the future. The fact that the Sunrise system is already reliable in the most difficult populations to characterize (ie, non-OSA and OSA with PSG-RDI ≥ 5 events/h [Figure 2B]) is reassuring and indicates the robustness of this approach. Furthermore, the underestimation occurring for PSG-RDI exceeding 40 events/h will not modify the therapeutic decision.³⁸⁻⁴⁰

Comparison of Sunrise System With Other OSA Diagnostic Systems

The Sunrise system device is designed to be used for ambulatory diagnosis of OSA outside of a sleep center setting and can thus be categorized in the SCOPER (Sleep, Cardiovascular, Oximetry, Position, Effort, and Respiratory parameters) framework⁴¹ as $S_3C_0O_0P_2E_4R_5$. As such, the performance of the device not only favorably compares with but also, in many instances, surpasses the performance of other fully automated diagnostic devices also classified within the SCOPER framework.⁴² A comparison with alternative solutions for the diagnosis of OSA is summarized in eTable 3 in the [Supplement](#). Signal acquisition was successfully completed in all included patients and was not disturbed by motion artifacts. This outcome is a marked advantage compared with existing type III (portable sleep study limited to sleep apnea and recording ≥ 4 signals, including electrocardiogram or heart rate, oxygen saturation, and ≥ 2 channels of respiratory movement and/or airflow) or type IV (sleep study with continuous recording of 1 or 2 signals or any test not fitting into the other categories) systems, which are limited by quality of acquisition in airflow, respiratory movements, or photoplethysmography signals.^{43,44} The performance of the Sunrise system device in identifying OSA can also be positively compared with self-reported measures, such as the Berlin questionnaire,⁴⁵ which varies in accuracy depending on the population studied. Interest in the Sunrise system approach relates to the combination of MM recordings with machine learning analysis. Machine learning approaches are increasingly used for diagnosis, disease risk classification, and therapeutic guidance^{21,46} and are especially promising in sleep medicine.^{22,23,47,48} The proposed automated method for analyzing MM has the potential to decrease setup and scoring time compared

with other OSA diagnostic systems, as well as reducing medical errors by facilitating the complex process of respiratory event scoring.

Limitations

This study has limitations. First, the chosen Sr-RDI cutoff of 7.63 leaves 9% of individuals as having false-negative results. However, a lower cutoff would optimize sensitivity at the expense of a large decrease in specificity. This factor may be addressed by increasing the number of recording nights. Similar to the Sunrise system device, the criterion-standard reference method using 1-night PSG might underestimate positive diagnosis in mild to moderate OSA. However, discomfort caused by electrodes and cables combined with an unfamiliar environment may result in a first-night effect. Repeated PSG has been suggested in sensitive populations (eg, those with insomnia and sleep apnea).⁴⁹ The use of the Sunrise system has greater capability than PSG for repeating nights at lower cost.

Second, it is theoretically possible that measurement of the amplitude of mandibular displacement was limited by localized constraints (ie, friction with the pillow in the lateral or prone position of the head or by excessive adipose tissue around the neck), slightly altering results of the automated Sunrise system analysis. Future studies will address these issues in more depth.

Third, the data were obtained in a sleep laboratory and will thus require confirmation and validation via home-based recordings in an ambulatory setting. A future study will evaluate the diagnostic capabilities of the Sunrise system in home-based settings by comparing ambulatory PSG results with consecutive nocturnal recordings using the Sunrise system, with the aim of reducing the consequences of internight variability.

Fourth, relative to the anticipated lower cost associated with the use of the Sunrise system, medicoeconomic data are needed to document inequalities in access to sleep apnea diagnosis and treatment. This information will address possible increases in disease detection and intervention, with potential attendant benefits on outcomes.

Conclusions

The Sunrise system automated analysis of MM provided an accurate estimation of RDI obtained during traditional PSG in a large cohort of patients with and without OSA. This approach may provide a suitable and convenient home-based alternative to the sleep center setting and serve as a stand-alone tool for automated assessment of OSA.

ARTICLE INFORMATION

Accepted for Publication: November 24, 2019.

Published: January 22, 2020. doi:[10.1001/jamanetworkopen.2019.19657](https://doi.org/10.1001/jamanetworkopen.2019.19657)

Open Access: This is an open access article distributed under the terms of the [CC-BY-NC-ND License](https://creativecommons.org/licenses/by-nc-nd/4.0/). © 2020 Pépin J-L et al. *JAMA Network Open*.

Corresponding Author: Jean-Louis Pépin, MD, PhD, Pôle Thorax et Vaisseaux, Centre Hospitalier Universitaire (CHU) de Grenoble-Alpes (CHUGA), Université Grenoble Alpes, Institut National de la Santé et de la Recherche Médicale, 38043 Grenoble, France (jpepin@chu-grenoble.fr).

Author Affiliations: Pôle Thorax et Vaisseaux, Centre Hospitalier Universitaire (CHU) de Grenoble-Alpes (CHUGA), Université Grenoble Alpes, Institut National de la Santé et de la Recherche Médicale, Grenoble, France (Pépin); Sunrise, Namur, Belgium (Letesson, Le-Dong, Dedave, Denison); Sleep Laboratory, CHU Université Catholique de Louvain (UCL) Namur Site Sainte-Elisabeth, Namur, Belgium (Cuthbert, Martinot); Institute of Experimental and Clinical Research, UCL Bruxelles Woluwe, Brussels, Belgium (Martinot); Department of Child Health, University of Missouri, Columbia (Gozal); Child Health Research Institute, University of Missouri, Columbia (Gozal).

Author Contributions: Drs Pépin and Martinot had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Pépin, Letesson, Le-Dong, Martinot, Gozal.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: All authors.

Critical revision of the manuscript for important intellectual content: Pépin, Letesson, Le-Dong, Martinot, Gozal.

Statistical analysis: Pépin, Letesson, Le-Dong, Dedave, Martinot, Gozal.

Obtained funding: Martinot.

Administrative, technical, or material support: Dedave, Denison, Cuthbert, Martinot.

Supervision: Pépin, Letesson, Martinot, Gozal.

Conflict of Interest Disclosures: Dr Pépin reported being a scientific advisor to Sunrise; receiving grants and/or personal fees from ResMed, Philips, Fisher & Paykel, Sefam, AstraZeneca, AGIR à dom, Elevie, VitalAire, Boehringer Ingelheim, Jazz Pharmaceuticals, and Itamar Medical Ltd; and receiving research support for clinical studies from Mutualia and Air Liquide Foundation. Dr Letesson and Messrs Dedave and Denison reported receiving personal fees from Sunrise. Dr Martinot reported being a nonremunerated scientific advisor to Sunrise and being a remunerated investigator in pharmacy trials for Jazz Pharmaceuticals and Theranexus. No other disclosures were reported.

Funding/Support: This study was performed at the Sleep Laboratory, Centre Hospitalier Universitaire Université Catholique de Louvain Namur Site Sainte-Elisabeth, Namur, Belgium. The devices used in the study were provided by Sunrise, Namur, Belgium. Dr Pépin is supported by the French National Research Agency in the framework of the "Investissements d'avenir" program (ANR-15-IDEX-02) and the "e-Health and integrated care" and "Trajectories medicine multidisciplinary institute in artificial intelligence" Chairs of Excellence from the Grenoble Alpes University Foundation. Dr Gozal is supported by National Institutes of Health grants HL130984 and HL140548.

Role of the Funder/Sponsor: The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Additional Contributions: Ravzat Ashurlaeva and Liesbeth Orij (RespiSom, Namur, Belgium) spent innumerable hours providing secretarial assistance. They were not compensated for their contributions.

REFERENCES

1. Benjafield AV, Ayas NT, Eastwood PR, et al. Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. *Lancet Respir Med*. 2019;7(8):687-698. doi:10.1016/S2213-2600(19)30198-5
2. Rosenzweig I, Glasser M, Polsek D, Leschziner GD, Williams SCR, Morrell MJ. Sleep apnoea and the brain: a complex relationship. *Lancet Respir Med*. 2015;3(5):404-414. doi:10.1016/S2213-2600(15)00090-9
3. Mulgrew AT, Ryan CF, Fleetham JA, et al. The impact of obstructive sleep apnea and daytime sleepiness on work limitation. *Sleep Med*. 2007;9(1):42-53. doi:10.1016/j.sleep.2007.01.009
4. Bjornsdottir E, Keenan BT, Eysteinsdottir B, et al. Quality of life among untreated sleep apnea patients compared with the general population and changes after treatment with positive airway pressure. *J Sleep Res*. 2015;24(3):328-338. doi:10.1111/jsr.12262
5. Gottlieb DJ, Ellenbogen JM, Bianchi MT, Czeisler CA. Sleep deficiency and motor vehicle crash risk in the general population: a prospective cohort study. *BMC Med*. 2018;16(1):44. doi:10.1186/s12916-018-1025-7
6. Drager LF, McEvoy RD, Barbe F, Lorenzi-Filho G, Redline S; INCOSACT Initiative (International Collaboration of Sleep Apnea Cardiovascular Trialists). Sleep apnea and cardiovascular disease: lessons from recent trials and need for team science. *Circulation*. 2017;136(19):1840-1850. doi:10.1161/CIRCULATIONAHA.117.029400
7. Hoyos CM, Drager LF, Patel SR. OSA and cardiometabolic risk: what's the bottom line? *Respirology*. 2017;22(3):420-429. doi:10.1111/resp.12984
8. Javaheri S, Barbe F, Campos-Rodriguez F, et al. Sleep apnea: types, mechanisms, and clinical cardiovascular consequences. *J Am Coll Cardiol*. 2017;69(7):841-858. doi:10.1016/j.jacc.2016.11.069
9. Marin JM, Carrizo SJ, Vicente E, Agusti AGN. Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. *Lancet*. 2005;365(9464):1046-1053. doi:10.1016/S0140-6736(05)71141-7
10. Craig SE, Kohler M, Nicoll D, et al. Continuous positive airway pressure improves sleepiness but not calculated vascular risk in patients with minimally symptomatic obstructive sleep apnoea: the MOSAIC randomised controlled trial. *Thorax*. 2012;67(12):1090-1096. doi:10.1136/thoraxjnl-2012-202178
11. Yu J, Zhou Z, McEvoy RD, et al. Association of positive airway pressure with cardiovascular events and death in adults with sleep apnea: a systematic review and meta-analysis. *JAMA*. 2017;318(2):156-166. doi:10.1001/jama.2017.7967

12. Sia CH, Hong Y, Tan LWL, van Dam RM, Lee CH, Tan A. Awareness and knowledge of obstructive sleep apnea among the general population. *Sleep Med*. 2017;36:10-17. doi:10.1016/j.sleep.2017.03.030
13. Peppard PE, Hagen EW. The last 25 years of obstructive sleep apnea epidemiology—and the next 25? *Am J Respir Crit Care Med*. 2018;197(3):310-312. doi:10.1164/rccm.201708-1614PP
14. Tarraubella N, Sánchez-de-la-Torre M, Nadal N, et al. Management of obstructive sleep apnoea in a primary care vs sleep unit setting: a randomised controlled trial. *Thorax*. 2018;73(12):1152-1160. doi:10.1136/thoraxjnl-2017-211237
15. Abrahamyan L, Sahakyan Y, Chung S, et al. Diagnostic accuracy of level IV portable sleep monitors versus polysomnography for obstructive sleep apnea: a systematic review and meta-analysis. *Sleep Breath*. 2018;22(3):593-611. doi:10.1007/s11325-017-1615-1
16. Uddin MB, Chow CM, Su SW. Classification methods to detect sleep apnea in adults based on respiratory and oximetry signals: a systematic review. *Physiol Meas*. 2018;39(3):03TR01. doi:10.1088/1361-6579/aaafb8
17. Muto V, Berthomier C, Schmidt C, et al. Inter- and intra-expert variability in sleep scoring: comparison between visual and automatic analysis [abstract 0315]. *Sleep*. 2018;41(suppl 1):A121. doi:10.1093/sleep/zyo061.314
18. Stöberl AS, Schwarz EI, Haile SR, et al. Night-to-night variability of obstructive sleep apnea. *J Sleep Res*. 2017;26(6):782-788. doi:10.1111/jsr.12558
19. Randerath W, Bassetti CL, Bonsignore MR, et al. Challenges and perspectives in obstructive sleep apnoea: report by an ad hoc working group of the Sleep Disordered Breathing Group of the European Respiratory Society and the European Sleep Research Society. *Eur Respir J*. 2018;52(3):1702616. doi:10.1183/13993003.02616-2017
20. Gray EL, Barnes DJ. Beyond the thermistor: novel technology for the ambulatory diagnosis of obstructive sleep apnoea. *Respirology*. 2017;22(3):418-419. doi:10.1111/resp.13004
21. Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. *N Engl J Med*. 2016;375(13):1216-1219. doi:10.1056/NEJMp1606181
22. Bragazzi NL, Guglielmi O, Garbarino S. SleepOMICS: how big data can revolutionize sleep science. *Int J Environ Res Public Health*. 2019;16(2):E291. doi:10.3390/ijerph16020291
23. Alvarez-Estevéz D, Moret-Bonillo V. Computer-assisted diagnosis of the sleep apnea-hypopnea syndrome: a review. *Sleep Disord*. 2015;2015:237878. doi:10.1155/2015/237878
24. Martinot JB, Le-Dong NN, Cuthbert V, et al. Mandibular movements as accurate reporters of respiratory effort during sleep: validation against diaphragmatic electromyography. *Front Neurol*. 2017;8:353. doi:10.3389/fneur.2017.00353
25. Senny F, Destiné J, Poirrier R. Midsagittal jaw movements as a sleep/wake marker. *IEEE Trans Biomed Eng*. 2009;56(2):303-309. doi:10.1109/TBME.2008.2003264
26. Senny F, Maury G, Cambron L, Leroux A, Destiné J, Poirrier R. The sleep/wake state scoring from mandible movement signal. *Sleep Breath*. 2012;16(2):535-542. doi:10.1007/s11325-011-0539-4
27. Martinot JB, Borel JC, Cuthbert V, et al. Mandibular position and movements: suitability for diagnosis of sleep apnoea. *Respirology*. 2017;22(3):567-574. doi:10.1111/resp.12929
28. Berry RB, Budhiraja R, Gottlieb DJ, et al; American Academy of Sleep Medicine; Deliberations of the Sleep Apnea Definitions Task Force of the 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. *J Clin Sleep Med*. 2012;8(5):597-619. doi:10.5664/jcsm.2172
29. Berry RB, Brooks R, Gamaldo C, et al. AASM Scoring Manual updates for 2017 (version 2.4). *J Clin Sleep Med*. 2017;13(5):665-666. doi:10.5664/jcsm.6576
30. American Academy of Sleep Medicine. *International Classification of Sleep Disorders*. 3rd ed. Darien, IL: American Academy of Sleep Medicine; 2014.
31. Sateia MJ. *International Classification of Sleep Disorders*—third edition: highlights and modifications. *Chest*. 2014;146(5):1387-1394. doi:10.1378/chest.14-0970
32. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Project for Statistical Computing; 2017.
33. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-310. doi:10.1016/S0140-6736(86)90837-8
34. Carstensen B, Simpson J, Gurrin LC. Statistical models for assessing agreement in method comparison studies with replicate measurements. *Int J Biostat*. 2008;4(1):16. doi:10.2202/1557-4679.1107

35. Collop NA, Anderson WM, Boehlecke B, et al; Portable Monitoring Task Force of the American Academy of Sleep Medicine. Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients. *J Clin Sleep Med*. 2007;3(7):737-747.
36. Bibbins-Domingo K, Grossman DC, Curry SJ, et al; US Preventive Services Task Force. Screening for obstructive sleep apnea in adults: US Preventive Services Task Force recommendation statement. *JAMA*. 2017;317(4):407-414. doi:10.1001/jama.2016.20325
37. Collop N. Breathing related arousals: call them what you want, but please count them. *J Clin Sleep Med*. 2014;10(2):125-126. doi:10.5664/jcsm.3434
38. Tsai WH, Flemons WW, Whitelaw WA, Remmers JE. A comparison of apnea-hypopnea indices derived from different definitions of hypopnea. *Am J Respir Crit Care Med*. 1999;159(1):43-48. doi:10.1164/ajrccm.159.1.9709017
39. Bianchi MT, Goparaju B. Potential underestimation of sleep apnea severity by at-home kits: rescoring in-laboratory polysomnography without sleep staging. *J Clin Sleep Med*. 2017;13(4):551-555. doi:10.5664/jcsm.6540
40. Setty AR. Underestimation of sleep apnea with home sleep apnea testing compared to in-laboratory sleep testing. *J Clin Sleep Med*. 2017;13(4):531-532. doi:10.5664/jcsm.6534
41. 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(5):531-548. doi:10.5664/jcsm.1328
42. Mendonça F, Mostafa SS, Ravelo-García AG, Morgado-Dias F, Penzel T. Devices for home detection of obstructive sleep apnea: a review. *Sleep Med Rev*. 2018;41:149-160. doi:10.1016/j.smrv.2018.02.004
43. Rétoéry Y, David P, de Picciotto C, Niedzialkowski P, Bonay M, Petitjean M. Validity of thoracic respiratory inductive plethysmography in high body mass index subjects. *Respir Physiol Neurobiol*. 2017;242:52-58. doi:10.1016/j.resp.2017.03.009
44. Ghamari M, Castaneda D, Esparza A, Soltanpur C, Nazeran H. A review on wearable photoplethysmography sensors and their potential future applications in health care. *Biosens Bioelectron*. 2018;4(4). doi:10.15406/ijbsbe.2018.04.00125
45. Senaratna CV, Perret JL, Matheson MC, et al. Validity of the Berlin questionnaire in detecting obstructive sleep apnea: a systematic review and meta-analysis. *Sleep Med Rev*. 2017;36:116-124. doi:10.1016/j.smrv.2017.04.001
46. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44-56. doi:10.1038/s41591-018-0300-7
47. Xu Z, Gutiérrez-Tobal GC, Wu Y, et al. Cloud algorithm-driven oximetry-based diagnosis of obstructive sleep apnoea in symptomatic habitually snoring children. *Eur Respir J*. 2019;53(2):1801788. doi:10.1183/13993003.01788-2018
48. Hornero R, Kheirandish-Gozal L, Gutiérrez-Tobal GC, et al. Nocturnal oximetry-based evaluation of habitually snoring children. *Am J Respir Crit Care Med*. 2017;196(12):1591-1598. doi:10.1164/rccm.201705-0930OC
49. Veauthier C, Piper SK, Gaede G, Penzel T, Paul F. The first night effect in multiple sclerosis patients undergoing home-based polysomnography. *Nat Sci Sleep*. 2018;10:337-344. doi:10.2147/NSS.S176201

SUPPLEMENT.

eFigure 1. Behavior of the Sunrise MM Signal (Sr-Derived MM) After Incorporation Into PSG Fragments of a Transition From Wake to Sleep (upper panel), a Sleep Period Marked With Episodes of Obstructive Apnoea (middle panel) or a Respiratory-Effort Related Arousal (RERA) (lower panel)

eFigure 2. Technical Details on the Sunrise Algorithm Development Phase

eTable 1. Characteristics of the Training and Validation Samples

eTable 2. Characteristics of Interest for Sleep Mandibular Movements (MM) Analysis as Shown in eFigure 1

eFigure 3. Bland-Altman Evaluation of the Agreement Between the 2 Methods for TST Measurement

eFigure 4. Bland-Altman Evaluation of the Agreement Between the 2 Methods for the Arousal Index

eFigure 5. Differences in Sr_RDI Values Among the 3 Clinical Subgroups Diagnosed in PSG

eFigure 6. Variability in Diagnostic Performance of Sr_RDI Across All Possible Thresholds

eAppendix. Sample Size Calculation

eTable 3. Comparison With Alternate Techniques for Automated Diagnosis of OSA

eReferences.