



OPEN ACCESS

EDITED BY

Ding Zou,
University of Gothenburg, Sweden

REVIEWED BY

Jean-Benoit Martinot,
CHU UCL Namur Site Godinne, Belgium
Carlos Teixeira,
Cooperativa de Ensino Superior Politécnico e
Universitário, Portugal

*CORRESPONDENCE

Stephen D. Pittman
✉ spittman@apnimed.com

RECEIVED 20 December 2024

ACCEPTED 13 March 2025

PUBLISHED 15 April 2025

CITATION

Gell LK, Mehta K, Esmaeili N,
Taranto-Montemurro L, Sands SA, Pittman SD
and Azarbarzin A (2025) Performance
evaluation of a ring-worn pulse oximeter for
the identification and monitoring of
obstructive sleep apnea.
Front. Sleep 4:1549272.
doi: 10.3389/frsle.2025.1549272

COPYRIGHT

© 2025 Gell, Mehta, Esmaeili,
Taranto-Montemurro, Sands, Pittman and
Azarbarzin. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Performance evaluation of a ring-worn pulse oximeter for the identification and monitoring of obstructive sleep apnea

Laura K. Gell¹, Ketan Mehta¹, Neda Esmaeili²,
Luigi Taranto-Montemurro¹, Scott A. Sands²,
Stephen D. Pittman^{1*} and Ali Azarbarzin²

¹Apnimed Inc., Cambridge, MA, United States, ²Division of Sleep and Circadian Disorders, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States

Introduction: Obstructive sleep apnea (OSA) is a highly prevalent chronic disorder that is challenging to monitor clinically. While single-night laboratory-based polysomnography (PSG) is the current gold standard for OSA assessment, its utility is limited by cost and inaccessibility. Overnight pulse oximetry is a feasible approach for simplified at-home monitoring of OSA. In this study, we evaluate the performance of a modified finger-worn pulse oximetry device ("Ring") for OSA assessment.

Methods: In all, 25 patients with OSA [age: 55.5 ± 7.7 years (mean \pm SD), body mass index (BMI): 31.8 ± 5.1 kg/m², 14M:11F, and Fitzpatrick scale score I–II: 15, III–IV: 6, and V–VI: 4] completed up to four in-laboratory PSG studies with simultaneous Ring oximetry measurements (90 studies in total). Correlation and agreement analyses compared Ring-derived measures of the oxygen desaturation index (ODI_{4RING}, desaturations $\geq 4\%$) against PSG measures (ODI_{4PSG} and AHI_{4PSG}). Likewise, Ring-derived hypoxic burden (HB_{RING}) was compared against its PSG counterpart (HB_{PSG}). Receiver operator characteristic (ROC) curve analysis quantified the ability of ODI_{4RING} to identify moderate-to-severe OSA (AHI_{4PSG} > 15 events/h).

Results: Median [interquartile range (IQR)] of AHI_{4PSG} was 18.0 [9.6, 31.7] events/h. ODI_{4RING} was positively correlated with ODI_{4PSG} (Pearson $r = 0.87$, root mean square error [RMSE] = 6.6 events/h, intraclass correlation [ICC] = 0.85) and AHI_{4PSG} ($r = 0.85$, RMSE = 7.1 events/h, ICC = 0.84). The bias (mean difference) and limits of agreement (1.96 SD) between ODI_{4PSG} and ODI_{4RING} were 2.9 and 14.2 events/h, while for AHI_{4PSG} and ODI_{4RING}, the bias and limits of agreement were 1.4 and 16.3 events/h, respectively. HB_{RING} was positively correlated with HB_{PSG} ($r = 0.75$, RMSE = 24.6% min/h, ICC = 0.73), with a mean difference of 3.7% min/h and limits of agreement of 60.6% min/h. The receiver operator characteristic curve analysis of ODI_{4RING} to identify moderate-to-severe OSA produced an area under the curve of 0.92 (ODI_{4PSG} > 15 events/h, "excellent") and 0.84 (AHI_{4PSG} > 15 events/h, "excellent").

Conclusion: Our results show that a low-cost, convenient, and simple-to-use finger-worn pulse oximeter is a reliable tool for continuous monitoring of OSA severity and therapy responses. It also offers excellent discriminative value for screening moderate-to-severe OSA in this population.

KEYWORDS

oximetry, validation, obstructive sleep apnea, wearable, monitoring

1 Introduction

Obstructive sleep apnea (OSA) is the most common sleep-related breathing disease, estimated to affect approximately one in four adults (Benjafield et al., 2019; Heinzer et al., 2015), in which decreased upper airway muscle tone during sleep leads to a repeating cycle of partial airway collapse (hypopnea) and total obstruction (apnea) accompanied by oxygen desaturation and sleep fragmentation (White, 2005). Untreated OSA is associated with increased daytime sleepiness and fatigue, and increased risk of cardiovascular and cerebrovascular disease, metabolic dysfunction, and early mortality (Antic et al., 2011; Azarbarzin et al., 2018; Javaheri et al., 2017; Marin et al., 2005; Bonsignore et al., 2013; Marshall et al., 2008). Although awareness about the high prevalence of OSA in the community is increasing, it frequently remains undiagnosed, even in patients with moderate-to-severe disease. Currently, an in-laboratory polysomnography (PSG) sleep study is considered the gold standard tool for diagnosing OSA; however, it is costly and resource-intensive, and thus access is limited. Home sleep testing (HST) has become more common as an alternative to in-lab PSG, using devices that record at a minimum airflow, pulse rate, and pulse oximetry channels. Devices and sensors must still be applied by an experienced professional or under their supervision, and studies are scored manually by a registered technologist. Thus, despite alleviating some of the resource and patient burden, delays to diagnosis remain high, and it is estimated that the majority of those with OSA are undiagnosed and thus untreated (Heinzer et al., 2015). Therefore, the need for a simple and easy-to-use screening tool for OSA that could help identify those with undiagnosed or suspected OSA is significant.

Moreover, despite the chronic nature of OSA, clinicians rely on PSG or HST measurement of OSA severity derived from a single night to make treatment decisions. However, the known within-subject night-to-night variability in respiratory events makes this problematic. It has been estimated that single-night PSG studies lead to misdiagnosis between 20 and 60% of the time (Roeder et al., 2020; Punjabi et al., 2020; Skiba et al., 2015; Tschopp et al., 2021; Lechat et al., 2022), whereas multi-night monitoring reduces the likelihood of misdiagnosis (Lechat et al., 2022). More importantly, PSG and HST with multiple channels (type 3) are not well suited for serial assessments during therapy titration. Simple, reliable tools are needed to objectively quantify changes in OSA severity over time and evaluate responses to therapy. As more treatment options emerge for patients, including the recent approval of pharmacotherapy for OSA (Malhotra et al., 2024), the value of such insights will increase. An incomplete treatment response could indicate the need to modify therapy or dose/level or consider a combination of therapies or interventions; however, clinical tools to facilitate ongoing monitoring are currently lacking.

Single-channel pulse oximetry to determine the oxygen desaturation index (ODI) has been shown to be a useful screening tool for OSA severity (Chiner et al., 1999; Dumitrache-Rujinski et al., 2013) that could be used by patients at home and over multiple nights for continued monitoring of OSA; however, accuracy varies across devices (Rashid et al., 2021). In this study, we evaluate using a new modified, wearable finger-worn Ring pulse oximeter device combined with custom algorithms for use in the ongoing monitoring of OSA severity during a randomized

crossover trial across different therapeutic conditions. In the primary analysis, the bias and limits of agreement between Ring and gold-standard PSG measurements of OSA severity (ODI4 and AHI4) were assessed. The secondary analysis assessed the intraclass and Pearson correlations between metrics and the ability of Ring ODI4 to identify moderate-to-severe ODI4. An exploratory analysis assessed the agreement between novel hypoxic burden measurements derived from the Ring oximeter alone and standard PSG measures.

2 Materials and methods

2.1 Study participants and design

This Ring oximetry ancillary study was conducted as part of a larger study (NCT05793684), approved by the WIRB-Copernicus Group Institutional Review Board (Aishah et al., 2024). Twenty-five participants with mild-to-severe OSA met eligibility criteria and were randomly assigned to the parent study across three sites: Brigham and Women's Hospital, Boston, Massachusetts ($N = 1$); Clayton Sleep Institute, St. Louis, Missouri ($N = 14$), and Santa Monica Clinical Trials, Los Angeles, California ($N = 10$). The exclusion criteria included clinically significant cardiac disease, neurological disorders, non-OSA sleep disorders, or uncontrolled hypertension. All participants provided informed written consent prior to study participation. During the study, participants completed up to four in-laboratory PSG studies under different treatment conditions (baseline, placebo, and treated with 500 mg Viloxazine, 500/75 mg Viloxazine-Trazadone). In total, 90 in-laboratory PSG studies were performed, and simultaneous Ring oximetry measurements were recorded successfully on all PSG nights.

2.2 Sleep study measurements

In-laboratory PSGs included measurements of electroencephalogram (EEG), electrooculogram (EOG), nasal pressure, thermistor, body position, and pulse oximetry using a finger probe. Oxygen saturation from PSG oximetry was sampled at a minimum rate of 1 Hz (10 Hz preferred), the signal averaging window was required to be between 1 and 3 s. In-lab PSGs were scored by a centralized PSG center, following the American Academy of Sleep Medicine (ASSM; Berry et al., 2017), with access to all PSG channels but blinded to the Ring oximetry data. The apnea-hypopnea index (AHI) was based on the AASM rule 1B for identification of hypopneas (which specifies a $\geq 30\%$ reduction in airflow for ≥ 10 s and oxygen desaturation of $\geq 4\%$, AHI_{4PSG}), as has been used recently to determine eligibility in several prominent OSA clinical trials (Malhotra et al., 2024; Schweitzer et al., 2023). We also measured the ODI based on $\geq 4\%$ desaturation (ODI_{4PSG}) and hypoxic burden (HB, area under the desaturation curve), based on manually scored respiratory events (Azarbarzin et al., 2018). Participants also wore a finger-based Ring pulse oximeter (Product Name: Pulse Oximeter, Model: S9; Shenzhen Viatom Technology Co. Ltd., Shenzhen, China) with modified firmware to meet our requirements. Oxygen saturation from the Ring was recorded at

1 Hz; oxygen desaturation from the Ring oximeter was determined as the number of desaturations $\geq 4\%$ from baseline per hour over total recording time using custom algorithms (ODI4_{RING}). HB from the Ring oximeter was also calculated (HB_{RING}), as the area under the desaturation curve based on automatically detected oxygen desaturations $\geq 2\%$ as described and validated by [Esmaeili et al. \(2023\)](#), divided by the total recording time.

2.3 Statistical analysis

The primary outcome was the bias and limits of agreement between ODI4_{RING} and ODI4_{PSG}. Repeated linked measurements under different conditions within subjects may lead to an underestimation of the limits of agreement if standard Bland–Altman analyses (in which it is assumed measurements are independent) are used. To account for this, we instead implemented a mixed-model approach, as described by [Carstensen et al. \(2008\)](#), including the subject-by-oximetry method and subject-by-treatment interactions to ensure the variance is correctly attributed between terms. This mixed-model approach provides a modestly more conservative limit of agreement estimate than simple difference variance reporting because we are modeling how much the difference variance appears to be reduced by the repeated non-independent estimates within subjects. The mean difference was then determined from the oximetry method coefficient and limits of agreement by the formula $1.96 \times \sqrt{(2 \times \tau^2 + \sigma_{\text{PSG}}^2 + \sigma_{\text{RING}}^2)}$, where τ^2 is the oximetry method by subject interaction variance and σ_{RING}^2 and σ_{PSG}^2 are the respective within-oximetry method residual variances. Agreement and limits were displayed on Bland–Altman plots for interpretation. The analysis was repeated to compare ODI4_{RING} and AHI4_{PSG}. For secondary outcomes, the intraclass correlation coefficients (ICCs) were also determined between measurements (0.75–0.9: good reliability, >0.9: excellent reliability). Pearson correlation analysis was used to assess the strength of the linear relationship between the Ring and gold-standard PSG measurements and the spread of observed data around this model using the root mean square error (RMSE). A correlation coefficient, r , >0.7 is considered a strong correlation, with a minimum RMSE requirement of fewer than 10 events/h (i.e., the difference between “mild” and “moderate” OSA classification). Receiver operator characteristic (ROC) curve analyses and area under the curve (AUC) analyses were used to evaluate the identification of moderate-to-severe OSA using the Ring with a cutoff of AHI4_{PSG} > 15 and ODI4_{PSG} > 15. An AUC >0.8 (“excellent”) was considered sufficient sensitivity and specificity for screening purposes. Optimal thresholds for identifying moderate-to-severe OSA were determined from the optimal operating point of the ROC curve, minimizing the misclassification cost where, given the intended use as a prescreening tool, false negatives were given twice the cost weighting of false positives. Pointwise confidence intervals for AUC, sensitivity, and specificity were calculated using bootstrapping with 1,000 iterations. In an exploratory analysis, we repeated the previously described analyses to assess the agreement between Ring- and PSG-derived hypoxic burden measurements; in ROC analyses, the ability of the Ring oximeter to screen for high HB (HB_{PSG} > 60) was evaluated. In

the *post-hoc* subgroup analysis, descriptive statistics are reported to describe the bias in different subgroups using the Fitzpatrick scale for skin tone. Statistical analyses were performed using MATLAB (Natick, MA) and R version 4.3.2.

3 Results

Patient characteristics are given in [Table 1](#). The median [interquartile range (IQR)] AHI4_{PSG} was 18.0 [9.6, 31.7] events/h and ODI4_{PSG} 19.5 [12.7, 32.5] events/h. Of the 90 sleep studies, 55 exhibited moderate-to-severe OSA (AHI4 > 15) per PSG scoring, 25 exhibited mild OSA ($5 \leq \text{AHI4} < 15$), and 10 exhibited no OSA (AHI4 < 5). Ring analysis showed a median ODI4_{RING} of 16.8 (10.7, 26.9) events/h. The maximum data rejection rate for the Ring SpO₂ data (null values as a percentage of total data points) across all recordings was 0.13%. Example signal traces from PSG and Ring recordings are shown in [Figure 1](#).

From mixed-model analyses, the bias (mean difference) and limits of agreement between ODI4_{PSG} and ODI4_{RING} were 2.9 ± 14.2 events/h ([Figure 2A](#)), and between AHI4_{PSG} and ODI4_{RING},

TABLE 1 Patient characteristics.

Demographics and obesity	
Age (years)	56 [50–62]
Sex, N (M:F)	14:11
BMI (Kg/m ²)	32 [28–38]
Race, N (Black:White:Asian:Other)	3:18:1:3
Ethnicity, N (Hispanic:Non-Hispanic)	3:22
Fitzpatrick scale (I:II:III:IV:V:VI)	3:12:5:1:3:1
Polysomnography	
Total sleep time (min)	387.3 [353.0, 424.0]
AHI4 (events/h)	18.0 [9.6, 31.7]
ODI4 (events/h)	19.5 [12.7, 32.5]
Hypoxic burden (%.min/h)	62.9 [41.2, 91.2]
Arousal index (events/h)	28.8 [19.6, 37.6]
Stage 1 (%TST)	12.2 [7.4, 17.0]
Stage 2 (%TST)	65.9 [60.8, 72.4]
Stage 3 (%TST)	8.7 [2.9, 14.0]
REM (%TST)	9.4 [4.6, 15.5]
T90 (%TST)	3.4 [0.8, 9.3]
Epworth sleepiness scale	9.5 [6.0, 12.0]
Ring metrics	
Total recording time (min)	491 [487.3, 496.7]
ODI4 _{RING} (events/h)	16.8 [10.7, 26.9]
HB _{RING} (%.min/h)	59.7 [38.6, 83]

The Fitzpatrick scale is a descriptive classification of skin tone, with I being the lightest and VI being the darkest tone. Data are presented as median [interquartile range]. AHI4, apnea-hypopnea index with 4% oxygen desaturation; HB, hypoxic burden; ODI4, oxygen desaturation index with 4% oxygen desaturation; T90, sleep time under 90% oxygen desaturation; TST, total sleep time; REM, rapid eye movement; BMI, Body Mass Index.

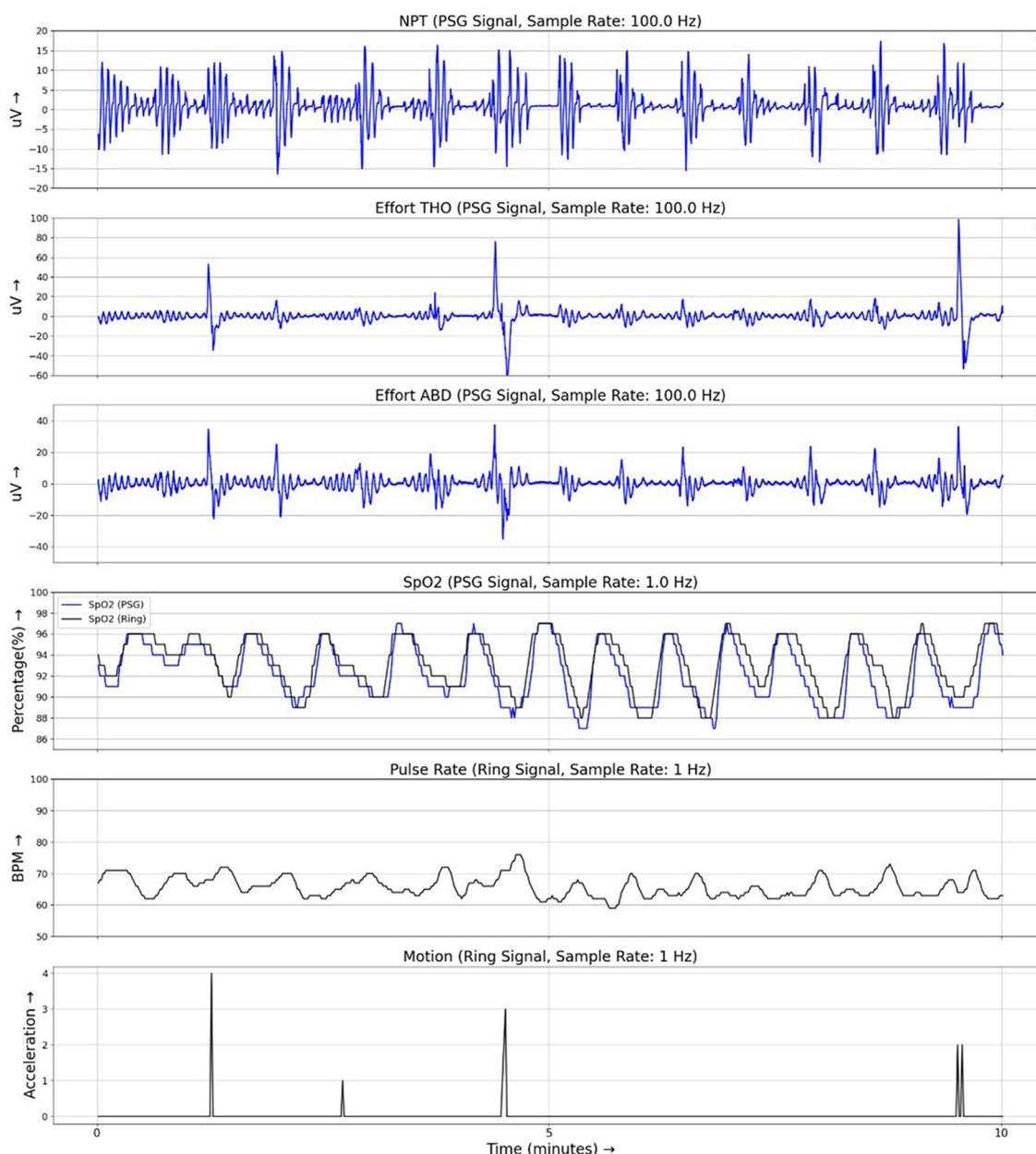
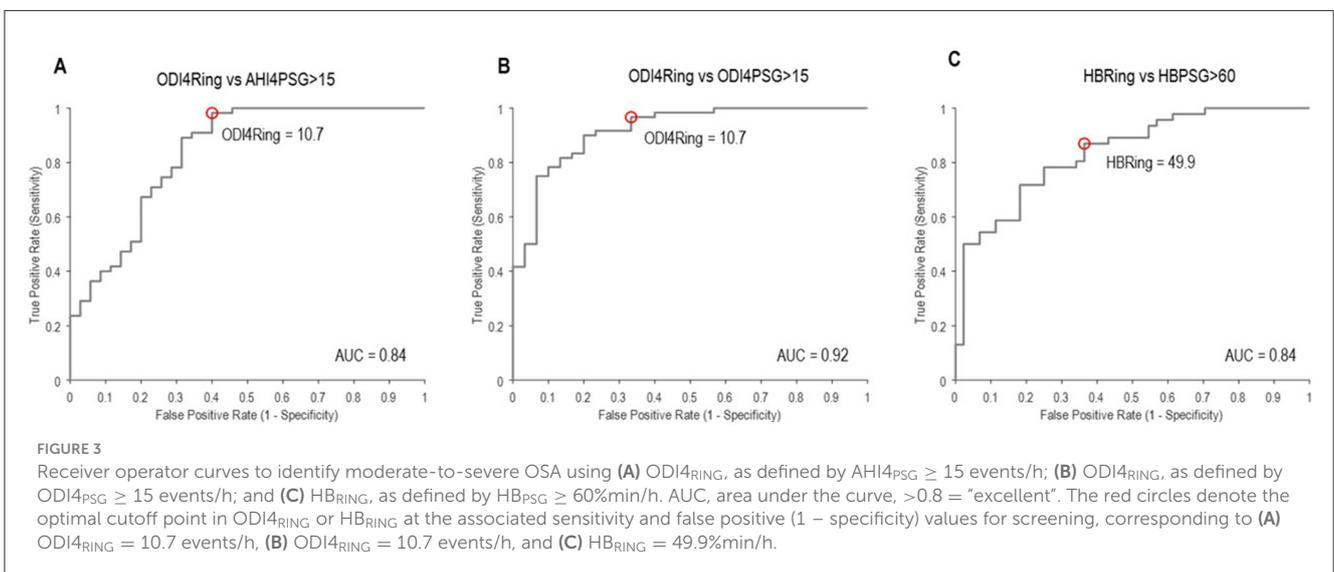
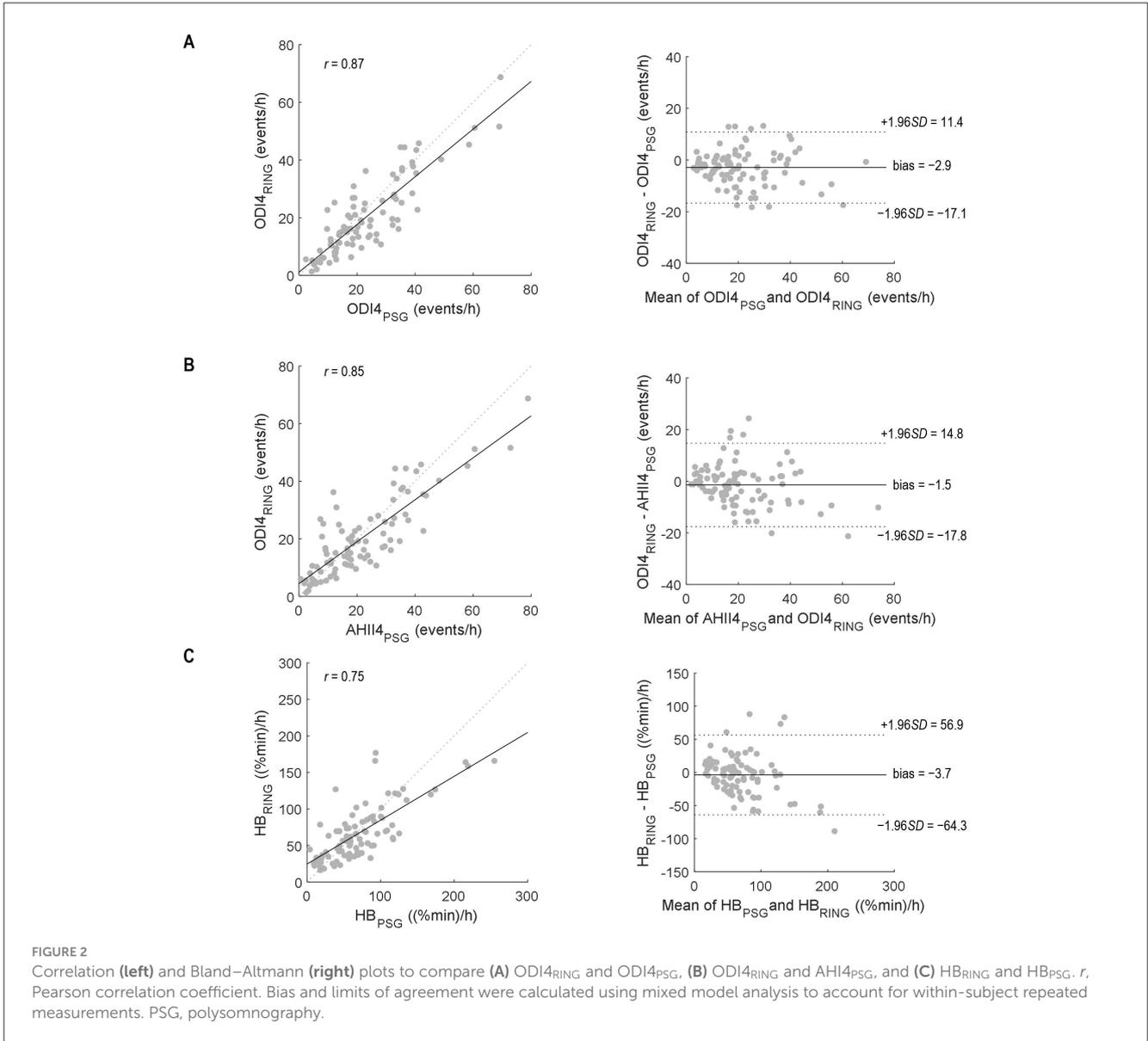


FIGURE 1

Example traces from simultaneous recordings of PSG and Ring data during a 10-min period of respiratory events. PSG signals are shown in blue and Ring signals are shown in black. NPT, nasal pressure; Effort THO, respiratory effort signal from the thoracic belt; Effort ABD, respiratory effort signal from the abdominal belt; SpO₂, oximetry signal.

they were 1.5 ± 16.3 events/h (Figure 2B), respectively. We note that limits of agreement were only slightly underestimated when ignoring the effects of repeated measurements using standard Bland–Altman analyses (± 13.8 and ± 16.2 events/h, respectively), suggesting that correlations between within-subject repeated measures in different treatment conditions are low. ODI_{4RING} was positively correlated with ODI_{4PSG} (Pearson $r = 0.87$, 95% CI [0.81, 0.91], RMSE = 6.6 events/h; Figure 2A), and AHI_{4PSG} ($r = 0.85$ [0.78–0.90], RMSE = 7.1 events/h; Figure 2B). The ICC for ODI_{4RING} vs. ODI_{4PSG} and vs. AHI_{4PSG} was 0.85 (95% CI [0.78, 0.90]) and 0.84 [0.76–0.89], respectively, and for the

ROC curve analysis of ODI_{4RING} to predict moderate-to-severe OSA produced an area-under-curve (AUC) of 0.84 (95% CI [0.73, 0.92]; AHI_{4PSG} >15 events/h, “excellent”, Figure 3A) and 0.92 [0.84–0.97] (ODI_{4PSG} >15 events/h, “excellent”; Figure 3B). The optimal cutoff for the screening of moderate-to-severe OSA by AHI_{4PSG} > 15 events/h (i.e., the gold standard) was determined to be ODI_{4RING} = 10.7 events/h, which was associated with a sensitivity of 0.98 [0.91–1.00] and a specificity of 0.60 [0.42–0.74] (Figure 3A). The optimal cutoff for the screening of moderate-to-severe OSA by ODI_{4PSG} > 15 events/h was determined to be ODI_{4RING} = 10.7 events/h, which was associated with a sensitivity



of 0.97 [0.88–1.00] and a specificity of 0.67 [0.47–0.81] (Figure 3B). In the exploratory analysis, HB from the Ring oximetry, HB_{RING}, was positively correlated with HB_{PSG} ($r = 0.75$ [0.65–0.83], RMSE = 24.6%min/h; Figure 2C), with a mean difference of 3.7%min/h and limits of agreement of $\pm 60.6\%$ min/h (Figure 2C). The ICC for HB_{RING} vs. HB_{PSG} was 0.73 [0.62–0.81], ROC analysis using HB_{RING} to predict moderate-to-severe HB_{PSG} > 60%min/h, AUC = 0.84 [0.74–0.91], and optimal threshold was determined to be 49.9%min/h, associated with a sensitivity of 0.87 [0.75–0.94] and specificity of 0.64 [0.47–0.76] (Figure 3C).

4 Discussion

This study has shown that the Ring oximeter device combined with custom algorithms can reliably detect OSA events and monitor OSA severity. The mean differences between ODI_{4RING} and PSG measurements were small, the ring slightly underestimates the severity of OSA on average (−2.9 events/h vs. AHI_{4PSG}, −1.4 events/h vs. ODI_{4PSG}). A visual analysis of the Bland–Altman plot suggests that bias tends to be greater at high values of ODI₄. The limits of agreement of ODI_{4RING} with AHI_{4PSG} and with ODI_{4PSG} were ± 14.2 events/h and ± 16.3 events/h, respectively: 95% of the differences lie within this range. Good reliability was observed between ODI_{4RING} and both AHI_{4PSG} and ODI_{4PSG} measurements (ICC = 0.84–0.85). Correlations between ODI_{4RING} and both AHI_{4PSG} and ODI_{4PSG} were strong ($r = 0.85$ – 0.87), with an RMSE of 6.6 and 7.1 events/h, respectively, meaning that the standard deviation of the difference between the observed and predicted values was less than the difference between mild and moderate OSA classification (10 events/h). The ROC curves showed that the Ring can identify moderate-to-severe sleep apnea per AHI_{4PSG} criteria with excellent discriminative value (AUC = 0.84), with an optimal threshold cutoff selected of ODI_{4RING} = 10.7 events/h for high sensitivity (0.98) but lower specificity (0.60). The optimization model was preferentially weighted for high sensitivity to maximize the number of true positives included, considering the intended application of the Ring as a prescreening tool; however, in different situations, alternative optimal operating points on the ROC curve could be chosen to increase specificity at the cost of lower sensitivity, for example, to make treatment decisions.

We think that the observed underestimate in ODI₄ with the Ring vs. PSG is predominately attributable to the use of the total recording time as a denominator in the Ring ODI₄ calculations vs. total sleep time for PSG metrics: Indeed, in exploratory analysis, using PSG total sleep time instead as the denominator for Ring ODI₄, we instead see a small positive mean bias (+2.8 events/h with ODI_{4RING} vs. ODI_{4PSG}; see Supplementary Figure S2), suggesting that the Ring could, in fact, be more sensitive to oxygen desaturations than the PSG finger probe if sleep time is better accounted for. A greater bias at higher ODI₄ values is not seen in a sleep-time-adjusted analysis. Sensor placement at the fingertip with the PSG probe vs. the base of the finger using the Ring may also contribute to differences. Nevertheless, systematically underestimating OSA severity may have important clinical implications on diagnosis or classification

if not accounted for in the interpretation of metrics, for example, by using adjusted thresholds. We note that the optimal screening cutoff for moderate to severe OSA identified using ODI_{4RING} is lower than the standard AHI criteria (10.8 vs. 15) and still produced excellent discrimination.

We also have demonstrated for the first time that it is feasible to calculate HB from a stand-alone wearable device using the oximetry signal alone. HB_{RING} and HB_{PSG} showed moderate-to-good agreement (ICC = 0.73). Hypoxic burden has been shown to be sensitively associated with a greater risk of cardiovascular disease and mortality (Azarbarzin et al., 2018, 2020), adding to the increasing evidence that intermittent hypoxemia plays a key role in the systemic long-term physiological consequences of OSA. Thus, in the future, monitoring oximetry-based metrics may be very useful for predicting disease risk and stratifying those who may most benefit from treatment.

The findings presented here are comparable with other wearable oximetry devices. A recent systematic review showed considerable variability in oximeters' performance across studies, with mean differences between oximetry to AHI_{PSG} that ranged from −13.7 to 4.8 events/h (Khor et al., 2023) and sensitivity and specificity of ODI values that ranged from 49 to 97% and 64 to 100%, respectively, for classifying AHI_{PSG} > 15 events/h (Khor et al., 2023). A wireless finger-worn oximeter and cloud-based analysis system (Oxistart, Biologix Sistemas Ltd., Brazil) were shown to accurately detect OSA, with a mean difference of 2.9 events/h and limits of agreement of ± 16.5 events/h compared to AHI_{4PSG} and an AUC 0.96 classifying moderate-to-severe OSA (Pinheiro et al., 2020). However, direct comparisons are challenging given differences in populations, OSA diagnostic criteria, ODI thresholds, and the selected optimal operating threshold between studies. The WatchPAT (Itamar Medical Inc., Caesarea, Israel) measures peripheral arterial tonometry in addition to oximetry, heart rate, and actigraphy and was found to detect AHI > 15, with an average sensitivity and specificity of 92.21 and 72.39%, respectively, in a recent meta-analysis of previous evaluation studies, with considerable variability between studies (Ifikhar et al., 2022). A commercial wrist-worn oximetry device (Galaxy Watch 4, Samsung, South Korea) can distinguish moderate-to-severe OSA (AUC: 0.80–0.91); however, device data rejection rates were high (26.5–52.3%; Jung et al., 2022; Browne et al., 2024).

Oximetry monitoring using a wearable device also provides specific advantages over the current gold standard, PSG. Unlike PSG, with many channels to precisely monitor multiple parameters across one night, the Ring could easily be used over multiple nights at home to provide accurate, multi-night, and averaged metrics to follow changes over time, for example, to assess treatment efficacy. The substantially lower cost of an oximetry device compared to a full PSG system allows for its use in a broader range of settings, including underserved communities and lower-income countries, where a full PSG is not feasible. Currently, night-to-night variability is recognized as a significant source of inconsistent measurement of OSA within a patient and can often lead to misdiagnosis (Roeder et al., 2020; Punjabi et al., 2020). Repeated measurements of OSA severity to reduce the effects of night-to-night variability correlate better with adverse cardiovascular outcomes than with single-night measures and reduce the misdiagnosis of OSA (Lechat et al.,

TABLE 2 Measurement bias by Fitzpatrick scale subgroup.

Fitzpatrick scale	AHI4 _{PSG} – ODI4 _{RING} (events/h)	ODI4 _{PSG} – ODI4 _{RING} (events/h)	HB _{PSG} – HB _{RING} (%min/h)
I, II (N = 15)	0.2 [–3, 7.6]	1.7 [–1.6, 7.5]	1.8 [–12.4, 17]
III, IV (N = 6)	1.7 [–1.8, 5.5]	2.8 [0.0, 5.5]	3 [–8, 27]
V, VI (N = 4)	1.6 [–11.1, 4.6]	2.3 [–2.2, 5.5]	–1.8 [–20.5, 20]

Data are presented as median [interquartile range]. The Fitzpatrick scale is a numerical classification schema for human skin color, where I is the palest and VI is the darkest tone.

AHI4_{PSG}, apnea–hypopnea index from PSG (hypopneas scored by >4% desaturation criteria); ODI4_{PSG}, oxygen desaturation index from PSG (hypopneas scored by >4% desaturation criteria); ODI4_{RING}, oxygen desaturation index from Ring oximeter (hypopneas scored by >4% desaturation criteria); HB_{PSG}, hypoxic burden from PSG; HB_{RING}, hypoxic burden from ring oximeter.

2023). Algorithmically identifying events also avoids uncertainty introduced by the inter-scorer variability in human scoring of AHI measures (Thomas et al., 2020; Collop, 2002). These sources of variability are also inherent in current single-night HST approaches to quantifying residual OSA severity on therapy and could potentially be mitigated through unobtrusive multi-night oximetry to get a more accurate representation of the true ongoing therapeutic response. Providing clinicians and/or patients with simple objective measurements of treatment efficacy or disease progression over time could help inform decisions about the best management strategies on an individualized basis.

It is known that oximetry-based metrics, including ODI and hypoxic burden, may be systematically underestimated in people with darker skin due to an underestimation of oxygen desaturation using pulse oximetry (Sjoding et al., 2020). The current study was not sufficiently powered to perform a formal analysis of the effects of skin pigmentation on the reliability of oximetry metrics. However, we note that the average bias was similar across the groups in our study (Table 2). Future studies are required to better explore this in a large diverse cohort.

Including other clinical features or spectral or non-linear characteristics of the oximetry signal into a multivariate classifier of OSA could further improve performance as a screening tool (Terrill, 2020; Behar et al., 2019), but the current study sought to limit the complexity of this initial validation study. In general, multivariate models have shown marginal improvements in accuracy compared to classifying with ODI alone (Uddin et al., 2018).

4.1 Strengths and limitations

This study was performed in a sleep laboratory under controlled conditions; metrics determined by the Ring device at home in an uncontrolled environment may exhibit greater variability compared with more controlled conditions. However, it could be argued that measurements made using minimally disruptive equipment in a usual sleep environment may be a better representation of true disease severity. Oximetry-based monitoring uses the total recording time as the denominator to assess metrics compared to total sleep time used in a PSG study

with simultaneous EEG recording, which likely contributes to the average underestimate of oximetry metrics from the Ring. Measurement bias may be greater in those with reduced sleep efficiency, for example, patients with comorbid insomnia or other complex sleep disorders. Ongoing work to determine sleep–wake staging using Ring accelerometry, oximetry, and pulse data could help address these concerns. Our study is a proof-of-concept trial conducted as part of a separate clinical trial; therefore, we did not perform *a priori* sample size calculations. *Post-hoc* power simulations indicated the current sample size provided a power of 0.991 to detect a clinically meaningful effect of 5 events/h (based on the median absolute difference between two repeated gold-standard PSG measurements of OSA severity; see the [Supplementary material](#) for details). Nevertheless, further studies with larger sample sizes across diverse populations are warranted to validate these findings. This study involved a prescreened population of OSA patients under different therapeutic conditions, only 10/90 PSGs exhibited no OSA per the gold-standard assessment (AHI4 < 5). We think that this is representative of a target population for the intended use of continued monitoring of OSA and in individual responses to therapy; however, it is likely not reflective of a screening population. Including more non-OSA control subjects, that is, better reflecting a community cohort, would better describe the performance for a target use for screening, and may result in different, potentially improved, sensitivity and specificity of the device. The ROC optimal operating thresholds may be different in this population, where more subjects are expected to have low AHI. Finally, the oximetry metrics explored in this study do not distinguish between central and obstructive respiratory events, whose gold-standard classification requires measuring respiratory effort. The diagnostic accuracy from the Ring alone may be lower in those with complex or central sleep apnea. The current findings may also not be representative of OSA patients with excluded comorbidities (e.g., significant cardiac disease or uncontrolled hypertension). However, pulse oximetry could still be used as a useful screening tool that could help clinicians identify those that may require a full PSG, and future algorithm development could enhance the ability to predict central events from obstructive events using event-based features.

4.2 Conclusion

Ring oximetry measures of OSA severity showed strong correlations with current gold-standard PSG criteria and were able to identify moderate to severe OSA with excellent discriminative value. This study shows the Ring oximeter has substantial promise as an accessible tool for the identification and multi-night monitoring of OSA severity.

Data availability statement

The datasets presented in this article are not readily available because data sharing requests will be considered from research groups that submit a research proposal and an appropriate statistical analysis plan and dissemination plan. Data will be shared

via a secure data access system. Requests to access the datasets should be directed to ltaranto@apnimed.com.

Ethics statement

The studies involving humans were approved by WIRB-Copernicus Group Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

LG: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Validation, Writing – original draft, Writing – review & editing. KM: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. NE: Formal analysis, Writing – review & editing, Investigation. LT-M: Writing – review & editing, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft. SS: Software, Writing – review & editing. SP: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. AA: Writing – review & editing, Formal analysis, Investigation.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was funded by Apnimed, Inc.

Conflict of interest

LG is a consultant for Apnimed Inc, KM, LT-M, and SP are employees of Apnimed Inc. AA reports grant support from Somnifix and serves as a consultant for Respicardia, Eli Lilly, Inspire, Cerebra, and Apnimed. Apnimed is developing pharmacological treatments for obstructive sleep apnea. AA is also the co-inventor of intellectual property pertaining to wearablesleep apnea phenotyping, relevant to the current manuscript, via

References

- Aishah, A., Kim, M., Norman, D., Ojile, J., Gell, L., Pho, H., et al. (2024). 0547 Vil-tra. *Sleep* 47, A234–A234. doi: 10.1093/sleep/zsae067.0547
- Antic, N. A., Catcheside, P., Buchan, C., Hensley, M., Naughton, M. T., Rowland, S., et al. (2011). The effect of CPAP in normalizing daytime sleepiness, quality of life, and neurocognitive function in patients with moderate to severe OSA. *Sleep* 34, 111–119. doi: 10.1093/sleep/34.1.111
- Azarbarzin, A., Sands, S. A., Stone, K. L., Taranto-Montemurro, L., Messineo, L., Terrill, P. I., et al. (2018). The hypoxic burden of sleep apnoea predicts cardiovascular disease-related mortality: the Osteoporotic Fractures in Men Study and the Sleep Heart Health Study. *Eur. Heart J.* 40, 1149–1157. doi: 10.1093/eurheartj/ehy624

his Institution. AA's interests were reviewed by Brigham and Women's Hospital and Mass General Brigham in accordance with their institutional policies. SS reports grant support from Apnimed, ProSomnus, and Dynaflex and has served as a consultant for Apnimed, Nox Medical, Inspire Medical Systems, Eli Lilly, Respicardia, LinguaFlex, and Achaemenid. He receives royalties for intellectual property pertaining to combination pharmacotherapy for sleep apnea via his institution. He is also the co-inventor of intellectual property pertaining to wearable sleep apnea phenotyping, relevant to the current manuscript, also via his institution. His industry interactions are actively managed by his institution. The authors declare that this study received funding from Apnimed, Inc. The funder had the following involvement in the study: study design, data collection and analysis, decision to publish, preparation of the manuscript.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frsle.2025.1549272/full#supplementary-material>

- Berry, R. B., Brooks, R., Gamaldo, C., Harding, S. M., Lloyd, R. M., Quan, S. F., et al. (2017). Scoring manual updates for 2017 (version 2.4). *J. Clin. Sleep Med.* 13, 665–666. doi: 10.5664/jcsm.6576
- Bonsignore, M. R., Borel, A. L., Machan, E., and Grunstein, R. (2013). Sleep apnoea and metabolic dysfunction. *Eur. Respir. Rev.* 22, 353–364. doi: 10.1183/09059180.00003413
- Browne, S. H., Vaida, F., Umlauf, A., Kim, J., DeYoung, P., Owens, R. L., et al. (2024). Performance of a commercial smart watch compared to polysomnography reference for overnight continuous oximetry measurement and sleep apnea evaluation. *J. Clin. Sleep Med.* 20, 1479–1488. doi: 10.5664/jcsm.11178
- Carstensen, B., Simpson, J., and Gurrin, L. C. (2008). Statistical models for assessing agreement in method comparison studies with replicate measurements. *Int. J. Biostat.* 4:16. doi: 10.2202/1557-4679.1107
- Chiner, E., Signes-Costa, J., Arriero, J. M., Marco, J., Fuentes, I., Sergado, A., et al. (1999). Nocturnal oximetry for the diagnosis of the sleep apnoea hypopnoea syndrome: a method to reduce the number of polysomnographies? *Thorax* 54, 968–971. doi: 10.1136/thx.54.11.968
- Collop, N. A. (2002). Scoring variability between polysomnography technologists in different sleep laboratories. *Sleep Med.* 3, 43–47. doi: 10.1016/S1389-9457(01)00115-0
- Dumitrache-Rujinski, S., Calcaianu, G., Zaharia, D., Toma, C. L., and Bogdan, M. (2013). The role of overnight pulse-oximetry in recognition of obstructive sleep apnea syndrome in morbidly obese and non obese patients. *Maedica* 8, 237–242.
- Esmaili, N., Labarca, G., Hu, W. H., Vena, D., Messineo, L., Gell, L., et al. (2023). Hypoxic burden based on automatically identified desaturations is associated with adverse health outcomes. *Ann. Am. Thor. Soc.* 20, 1633–1641. doi: 10.1513/AnnalsATS.202303-248OC
- Heinzer, R., Vat, S., Marques-Vidal, P., Marti-Soler, H., Andries, D., Tobback, N., et al. (2015). Prevalence of sleep-disordered breathing in the general population: the HypnoLaus study. *Lancet Respir. Med.* 3, 310–318. doi: 10.1016/S2213-2600(15)00043-0
- Iftikhar, I. H., Finch, C. E., Shah, A. S., Augunstein, C. A., and Ioachimescu, O. C. A. (2022). meta-analysis of diagnostic test performance of peripheral arterial tonometry studies. *J. Clin. Sleep Med.* 18, 1093–1102. doi: 10.5664/jcsm.9808
- Javaheri, S., Barbe, F., Campos-Rodriguez, F., Dempsey, J. A., Khayat, R., Javaheri, S., et al. (2017). Sleep apnea: types, mechanisms, and clinical cardiovascular consequences. *J. Am. Coll. Cardiol.* 69, 841–858. doi: 10.1016/j.jacc.2016.11.069
- Jung, H., Kim, D., Lee, W., Seo, H., Seo, J., Choi, J., et al. (2022). Performance evaluation of a wrist-worn reflectance pulse oximeter during sleep. *Sleep Health* 8, 420–428. doi: 10.1016/j.sleh.2022.04.003
- Khor, Y. H., Khung, S. W., Ruehland, W. R., Jiao, Y., Lew, J., Munsif, M., et al. (2023). Portable evaluation of obstructive sleep apnea in adults: a systematic review. *Sleep Med. Rev.* 68:101743. doi: 10.1016/j.smrv.2022.101743
- Lechat, B., Naik, G., Reynolds, A., Aishah, A., Scott, H., Loffler, K. A., et al. (2022). Multinight prevalence, variability, and diagnostic misclassification of obstructive sleep apnea. *Am. J. Respir. Crit. Care Med.* 205, 563–569. doi: 10.1164/rccm.202107-1761OC
- Lechat, B., Nguyen, D. P., Reynolds, A., Loffler, K., Escourrou, P., McEvoy, R. D., et al. (2023). Single-night diagnosis of sleep apnea contributes to inconsistent cardiovascular outcome findings. *Chest* 164, 231–240. doi: 10.1016/j.chest.2023.01.027
- Malhotra, A., Grunstein, R. R., Fietze, I., Weaver, T. E., Redline, S., Azarbarzin, A., et al. (2024). Tirzepatide for the treatment of obstructive sleep apnea and obesity. *N. Engl. J. Med.* 391, 1193–1205. doi: 10.1056/NEJMoa2404881
- Marin, J. M., Carrizo, S. J., Vicente, E., and Agusti, A. G. (2005). Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. *Lancet* 365, 1046–1053. doi: 10.1016/S0140-6736(05)71141-7
- Marshall, N. S., Wong, K. K., Liu, P. Y., Cullen, S. R., Knudman, M. W., Grunstein, R. R., et al. (2008). Sleep apnea as an independent risk factor for all-cause mortality: the Busselton Health Study. *Sleep* 31, 1079–1085. doi: 10.5665/sleep/31.8.1079
- Pinheiro, G. D. L., Cruz, A. F., Domingues, D. M., Genta, P. R., Drager, L. F., Strollo, P. J., et al. (2020). Validation of an overnight wireless high-resolution oximeter plus cloud-based algorithm for the diagnosis of obstructive sleep apnea. *Clinics* 75:e2414. doi: 10.6061/clinics/2020/e2414
- Punjabi, N. M., Patil, S., Crainiceanu, C., and Aurora, R. N. (2020). Variability and misclassification of sleep apnea severity based on multi-night testing. *Chest* 158, 365–373. doi: 10.1016/j.chest.2020.01.039
- Rashid, N. H., Zaghi, S., Scapuccin, M., Camacho, M., Certal, V., Capasso, R., et al. (2021). The value of oxygen desaturation index for diagnosing obstructive sleep apnea: a systematic review. *Laryngoscope* 131, 440–447. doi: 10.1002/lary.28663
- Roeder, M., Bradicich, M., Schwarz, E. I., Thiel, S., Gaisl, T., Held, U., et al. (2020). Night-to-night variability of respiratory events in obstructive sleep apnoea: a systematic review and meta-analysis. *Thorax* 75, 1095–1102. doi: 10.1136/thoraxjnl-2020-214544
- Schweitzer, P. K., Taranto-Montemurro, L., Ojile, J. M., Thein, S. G., Drake, C. L., Rosenberg, R., et al. (2023). The combination of aroxybutynin and atomoxetine in the treatment of obstructive sleep apnea (MARIPOSA): a randomized controlled trial. *Am. J. Respir. Crit. Care Med.* 208, 1316–1327. doi: 10.1164/rccm.202306-1036OC
- Sjoding, M. W., Dickson, R. P., Iwashyna, T. J., Gay, S. E., and Valley, T. S. (2020). Racial bias in pulse oximetry measurement. *N. Engl. J. Med.* 383, 2477–2478. doi: 10.1056/NEJMc2029240
- Skiba, V., Goldstein, C., and Schotland, H. (2015). Night-to-night variability in sleep disordered breathing and the utility of esophageal pressure monitoring in suspected obstructive sleep apnea. *J. Clin. Sleep Med.* 11, 597–602. doi: 10.5664/jcsm.4764
- Terrill, P. I. (2020). A review of approaches for analysing obstructive sleep apnoea-related patterns in pulse oximetry data. *Respirology* 25, 475–485. doi: 10.1111/resp.13635
- Thomas, R. J., Chen, S., Eden, U. T., and Prerau, M. J. (2020). Quantifying statistical uncertainty in metrics of sleep disordered breathing. *Sleep Med.* 65, 161–169. doi: 10.1016/j.sleep.2019.06.003
- Tschopp, S., Wimmer, W., Caversaccio, M., Borner, U., and Tschopp, K. (2021). Night-to-night variability in obstructive sleep apnea using peripheral arterial tonometry: a case for multiple night testing. *J. Clin. Sleep Med.* 17, 1751–1758. doi: 10.5664/jcsm.9300
- Uddin, M. B., Chow, C. M., and Su, S. W. (2018). Classification methods to detect sleep apnea in adults based on respiratory and oximetry signals: a systematic review. *Physiol. Meas.* 39:03TR01. doi: 10.1088/1361-6579/aaafb8
- White, D. P. (2005). Pathogenesis of obstructive and central sleep apnea. *Am. J. Respir. Crit. Care Med.* 172, 1363–1370. doi: 10.1164/rccm.200412-1631SO