ORIGINAL ARTICLE

Multinight Prevalence, Variability, and Diagnostic Misclassification of Obstructive Sleep Apnea

Bastien Lechat[1], Ganesh Naik[1], Amy Reynolds[1], Atqiya Aishah[1,2], Hannah Scott[1], Kelly A. Loffler[1], Andrew Vakulin[1], Pierre Escourrou[3], R. Doug McEvoy[1], Robert J. Adams[1], Peter G. Catcheside[1], and Danny J. Eckert[1]

[1]Adelaide Institute for Sleep Health and Flinders Health and Medical Research Institute Sleep Health, College of Medicine and Public Health, Flinders University, Adelaide, Australia; [2]School of Medical Sciences, University of New South Wales, Sydney New South Wales, Australia; and [3]Centre Interdisciplinaire du Sommeil, Paris, France

ORCID IDs: 0000-0003-0760-0714 (B.L.); 0000-0003-1790-9838 (G.N.); 0000-0001-9534-8699 (A.R.); 0000-0001-9383-1724 (A.A.); 0000-0002-0707-8068 (H.S.); 0000-0003-3302-5995 (K.A.L.); 0000-0002-3919-1313 (A.V.); 0000-0003-0068-0563 (P.E.); 0000-0002-5759-0094 (R.D.M.); 0000-0001-6254-0385 (R.J.A.); 0000-0002-9372-6788 (P.G.C.); 0000-0003-3503-2363 (D.J.E.).

Abstract

Rationale: Recent studies suggest that obstructive sleep apnea (OSA) severity can vary markedly from night to night, which may have important implications for diagnosis and management.

Objectives: This study aimed to assess OSA prevalence from multinight in-home recordings and the impact of night-to-night variability in OSA severity on diagnostic classification in a large, global, nonrandomly selected community sample from a consumer database of people that purchased a novel, validated, under-mattress sleep analyzer.

Methods: A total of 67,278 individuals aged between 18 and 90 years underwent in-home nightly monitoring over an average of approximately 170 nights per participant between July 2020 and March 2021. OSA was defined as a nightly mean apnea–hypopnea index (AHI) of more than 15 events/h. Outcomes were multinight global prevalence

and likelihood of OSA misclassification from a single night's AHI value.

Measurements and Main Results: More than 11.6 million nights of data were collected and analyzed. OSA global prevalence was 22.6% (95% confidence interval, 20.9–24.3%). The likelihood of misdiagnosis in people with OSA based on a single night ranged between approximately 20% and 50%. Misdiagnosis error rates decreased with increased monitoring nights (e.g., 1-night F1-score = 0.77 vs. 0.94 for 14 nights) and remained stable after 14 nights of monitoring.

Conclusions: Multinight in-home monitoring using novel, noninvasive under-mattress sensor technology indicates a global prevalence of moderate to severe OSA of approximately 20%, and that approximately 20% of people diagnosed with a single-night study may be misclassified. These findings highlight the need to consider night-to-night variation in OSA diagnosis and management.

Keywords: sleep-disordered breathing; misdiagnosis; wearables; polysomnography

At a Glance Commentary

Scientific Knowledge on the Subject: Recent studies that have assessed obstructive sleep apnea (OSA) over several consecutive nights indicate that OSA severity can vary markedly from night to night, which may have important implications for diagnosis, management, and prevalence estimates. Before recent advances in noninvasive home sleep monitoring technology, it was not feasible to examine night-to-night variation in OSA severity and its potential impact on diagnostic classification and prevalence estimates over extended periods in the home setting at scale.

What This Study Adds to the Field: Multinight in-home assessment of OSA in a self-selected population from a consumer database of more than 65,000 people with approximately 6 months of nightly data per participant indicates an estimated global prevalence of moderate to severe OSA (>15 respiratory events/h sleep) of approximately 20%. The impact of night-to-night variation in OSA severity means that 20% or more of patients diagnosed based on single-night testing (current practice) may be misclassified. OSA diagnostic confidence is high and stable after 14 nights of monitoring. Simple, noninvasive multinight assessment for OSA may be a feasible, cost-effective approach to minimize diagnostic misclassification to help complement routine clinical practice and better define OSA for research trials.

Obstructive sleep apnea (OSA) is a common sleep-related breathing disorder, estimated to affect nearly 1 billion people worldwide (1). Untreated OSA is associated with adverse health and safety consequences, including increased risk of cardiovascular disease (2), depression (3), traffic accidents (4), reduced quality of life

(5), and all-cause mortality (2, 6). The current gold standard measurement of OSA severity is the apnea–hypopnea index (AHI), a frequency count of the number of 10-second or longer apneas (complete cessation of airflow) and hypopneas (reductions in airflow) associated with sleep disruption (cortical arousal) and/or hypoxemia per hour of sleep. Current standard clinical practice to diagnose OSA requires a single overnight in-laboratory or in-home polysomnography or polygraphy study in which an AHI between 5 and 15 is classified as mild, 15–30 moderate, and more than 30 events/h as severe OSA (7).

However, diagnosis of OSA based on the AHI from a single night's sleep study may be problematic owing to potential night-to-night physiological and measurement variability. Furthermore, substantial variation in interscorer agreement in manual scoring of the AHI leads to uncertainty of diagnosis in approximately 25–45% of patients (8). The latest evidence on night-to-night variability in AHI from a sample of 10,000 adults who completed three nights of in-home testing indicates that up to 20% of people with OSA may be misdiagnosed based on a single night study (9). This is concordant with earlier reports that indicate between 20% and 60% of people with OSA may be misdiagnosed based on a single-night study depending on the sampled population (e.g., sleep clinic vs. general population) (10–15). However, previous studies were limited in sample size, and misdiagnosis rates were only calculated on a small (<14) number of nights, which limits generalizability. Furthermore, the number of nights required to achieve a confident estimate of the AHI and reliable diagnostic classification remains to be determined.

Accordingly, this study sought to use validated (16), novel noninvasive under-mattress sensor technology in a large global population sample to estimate OSA prevalence from multinight recordings and to assess the impact of night-to-night variability in OSA severity on diagnostic classification.

Deidentified data that support the findings of this study, including individual data, are available from the corresponding author upon request subject to ethical and data custodian (Withings) approval.

Methods

Participants

Data were acquired from 87,610 participants from a consumer-user database of people who registered to use under-mattress sleep sensor technology (Withings Sleep Analyzer [WSA]) to track their nightly sleep in their homes between July 1, 2020, and March 30, 2021. This large nonrandomly selected sample was composed of participants from 151 countries/regions. All participants provided informed consent for their deidentified data to be used for research purposes. This study was approved by the Flinders University Human Research Ethics Committee (project number: 4,291).

Monitoring Equipment

The WSA is a contactless, nonintrusive sleep monitoring device that is positioned under the user's mattress. The device contains pneumatic and sound sensors to detect body movements, respiratory rate, heart rate, snoring, and episodes of breathing cessation. These signals are used to estimate sleep macrostructure (total sleep time), sleep timing (bedtime and wake-time), and the AHI using automated algorithms (*see* the online supplement for a more detailed technology description). Clinical validation shows good agreement with polysomnography-derived AHI (16, 17) with high predictive performance (88% sensitivity and 88% specificity) to classify moderate to severe OSA (≥15 events/h sleep). A further internal validation study in 32 participants (26 men and 6 women) independently studied at the Adelaide Institute for Sleep Health laboratory (*see* the online supplement) showed similar diagnostic performance characteristics.

OSA Classification

Data from the Withings database were extracted for all participants aged at least 18 years with at least 28 nights of recordings who used their device for an average of at least four times per week during the 9-month assessment period. Mean and 95% confidence intervals (CIs) for AHI were calculated from all available nights separately for each participant. The probability of OSA above/below standard clinical cutoffs (<5 = no OSA, ≥5 and <15 = mild, ≥15 and <30 = moderate, and ≥30 events/h = severe OSA) was estimated for each participant (7). Absence versus presence of

OSA was defined as an AHI of 15 events/h or more, indicating OSA of at least moderate severity (7).

Global Prevalence Estimates

Assuming an OSA prevalence in the population of approximately 25% (95% CI, 20–30%) (18), we calculated that at least 289 individuals would be required to estimate OSA prevalence in each country/region. Therefore, OSA prevalence by country/region was only calculated in countries/regions where data from at least 300 individuals were available. Each prevalence estimate (p) has variance equal to $p(1-p)n$, where n is the total number of participants in that country. The 95% CI of the prevalence per country is therefore

$$\pm 1.96 \sqrt{\left(p(1-p)n \right)}.$$

Global prevalence is expressed as the mean and 95% CI across countries. Age- and sex-stratified global OSA prevalence was also calculated (*see* analysis in the online supplement).

OSA Misclassification

The reliability to classify OSA based on 1 night, 2 nights, 7 nights, and 14 nights was determined using receiver operator characteristic curves, precision-recall curves, and detection error tradeoff curves. Briefly, true labels were determined based on the mean AHI of all available nights for each individual and according to defined clinical cutoffs. X-night diagnosis labels were determined using the mean AHI of X randomly selected nights for each participant. True positive, false negative, false positive, and true negative values were then calculated for an X-night diagnosis, from which true positive rates, false positive rates, positive predictive values, false negative rates, false positive rates, F1-score, and Matthew's correlation coefficients (*see* the online supplement) were derived. This process of randomly selecting nights and calculating rates was repeated 100 times for each X-night diagnosis, allowing the calculation of mean and 95% CI of predictive performance metrics.

A sensitivity analysis was performed to further validate these findings in which only data recorded in March 2021 were included. In this sensitivity analysis, only participants with at least 28 days of recording during March 2021 were included. This sensitivity analysis aimed to reduce the potential impact of uncontrolled confounders (e.g., weight loss/gain) that may have resulted in increased AHI variability measured across the full recording period.

Statistical Analysis

Effects of available sociodemographic variables (age, body mass index, and sex) and AHI variability on the likelihood of misclassification and inconclusive diagnostic classification were studied using logistic regression. Given that participants with a mean AHI of less than 5 events/h were not at risk of being misclassified, these participants were excluded. Inconclusive diagnosis was defined as a 1-night probability of being misclassified higher than 10%. AHI variability was defined using tertiles of the average standard deviation of AHI across all nights. Statistical analysis was performed in the R programming language (19), using the rms modeling package (20).

Results

Participant Characteristics

Of the 87,610 registered users in the database, 10,954 (12.1%) and 8,501 (9.7%) were excluded as they were new users with 28 or fewer days of recording or were infrequent device users (average of less than 4 times a week), respectively. A further 745 (0.8%) participants were excluded as they were less than 18 years old. The characteristics for the remaining 67,278 participants are summarized in Table 1. Most participants resided in Europe (62%) and North America (30%) (*see* Figure E2 in the online supplement for the remaining distributions globally).

Global Prevalence of OSA

Of the 20 countries with at least 300 users, the global prevalence of OSA was 22.6% (95% CI, 20.9–24.3%). Prevalence estimates varied between 15% (Japan) and 29%

Table 1. Baseline Participant Characteristics

		Overall	OSA Severity			
			None	**Mild**	**Moderate**	**Severe**
n		67,278	30,051	21,573	9,982	5,672
Age, yr		47 (13)	42 (11)	49 (12)	55 (12)	57 (12)
BMI, kg/m^2		27 (5)	26 (5)	28 (5)	29 (5)	31 (6)
Sex	Male	52,533 (78)	21,775 (72)	17,239 (80)	8,450 (85)	5,069 (89)
	Female	14,745 (22)	8,276 (28)	4,334 (20)	1,532 (15)	603 (11)
Continent	Europe	41,627 (62)	17,313 (58)	14,000 (65)	6,642 (67)	3,672 (65)
	North America	19,907 (30)	9,459 (32)	6,099 (28)	2,721 (27)	1,628 (29)
	Asia	4,242 (6)	2,598 (9)	972 (5)	424 (4)	248 (4)
	Oceania	920 (2)	432 (1)	304 (2)	113 (2)	71 (2)
	South America	192 (0)	71 (0)	75 (0)	32 (0)	14 (0)
	Africa	152 (0)	66 (0)	58 (0)	18 (0)	10 (0)
Number of nights, *n*		174 (72)	171 (72)	176 (71)	176 (71)	173 (73)
Use, nights/week		6 (1)	6 (1)	6 (1)	6 (1)	6 (1)
Mean AHI, events/h		11 (13)	2 (1)	9 (3)	21 (4)	46 (15)
AHI variability*, events/h		6 (4)	3 (1)	6 (2)	10 (3)	14 (6)
Mean total sleep time, min		448 (50)	451 (43)	451 (48)	447 (54)	425 (73)

Definition of abbreviations: AHI = apnea–hypopnea index; BMI = body mass index; OSA = obstructive sleep apnea.
*Defined as SD AHI across all available nights.

Table 2. Prevalence of Obstructive Sleep Apnea (AHI ⩾ 15 events/h) for Countries with More than 300 Participants

	Number of Participants	Age, Median (IQR)	Sex, Female, %	OSA Prevalence, Mean (95% CI), %	OSA Prevalence in Benjafield et al. (1), %
Japan	2,885	43 (34–51)	10.6	15.3 (14.0–16.6)	14.0
China	388	41 (32–49)	17.0	17.5 (13.7–21.0)	8.8
Netherlands	1,533	46 (37–53)	18.9	19.0 (17.0–21.0)	28.5
Australia	1,530	43 (36–52)	18.8	19.3 (16.5–22.2)	4.8
Canada	794	48 (36–56)	25.3	19.3 (17.2–22.2)	4.8
Norway	503	45 (36–52)	17.4	19.8 (16.4–23.4)	13.1
Finland	1,183	46 (37–54)	21.7	20.0 (17.7–22.3)	29.5
United States	18,061	46 (36–55)	25.9	22.0 (21.4–22.6)	14.5
Poland	392	44 (37–52)	16.1	22.4 (18.3–26.6)	17.8
Sweden	4,864	47 (37–57)	20.0	23.0 (21.8–24.1)	12.7
UK	1,239	46 (37–54)	21.0	23.0 (21.8–24.2)	4.8
France	10,300	48 (37–57)	23.4	23.3 (22.5–24.1)	36.3
Portugal	302	49 (39–57)	25.6	23.5 (18.7–28.3)	12.5
Belgium	1,067	47 (38–55)	23.1	23.5 (20.9–26.0)	15.7
Switzerland	2,725	48 (39–56)	21.0	23.6 (21.9–25.1)	36.6
Denmark	575	47 (39–56)	18.3	24.7 (21.1–28.2)	28.5
Spain	782	48 (40–57)	20.1	26.4 (23.4–29.6)	16.2
Austria	791	48 (38–57)	18.1	28.3 (25.1–31.4)	28.4
Italy	1,147	50 (38–57)	16.8	28.4 (25.8–31.0)	12.0
Germany	12,252	50 (41–58)	20.4	29.0 (28.1–29.8)	32.9

Definition of abbreviations: AHI = apnea–hypopnea index; CI = confidence interval; IQR = interquartile range; OSA = obstructive sleep apnea.

(Germany), as shown in Table 2 and Figure E2. The United States, France, Germany, and the UK had the highest proportion of participants and had OSA prevalence rates of 21.6, 23.1, 29.0, and 22.9%, respectively. OSA prevalence in men was almost twice as high as in women (25% vs. 13%), as shown in Table E1. OSA prevalence for participants aged 50 years or less was 12% (95% CI, 11–14), which was much lower than for participants aged more than 50 years (mean [95% CI], 39% [36–42%]). Demographics were similar to data previously published by the World Health Organization and the United Nations, except for the lowest proportions of women per country (Table E2).

Likelihood of OSA Misclassification
The likelihood of detecting OSA in participants with a confident classification of OSA (according to their reference mean value) for a single night diagnosis was 79% (Figure E4A). Conversely, an average of 21% of the diagnoses would have resulted in false negatives based on a single night (detection of no OSA rather than OSA). Misclassification rates were dependent on the mean AHI value and the AHI cutoffs values selected, as shown in Figure E4B.

Multiple-night diagnosis (using the mean across X nights) showed substantially better predictive performance than a one-night diagnosis (as shown in Figures 1A, 1B, and 1C). Performance increased from a 1-night diagnosis to a 14-night diagnosis, with increased area under the receiver operator characteristic curve (Figure 1A), better precision-recall score (Figure 1B) and a decrease in both false negative and false positive rates (Figure 1C). Using 10% as the maximally acceptable error rate, false positive rates of 16.8% (0.2), 8.5% (0.2), 2.0% (0.1), and 1.0% (0) occurred for a 1-, 2-, 7- and 14-night diagnostic assessment period, respectively. Furthermore, the F1-score (a measure of overall predictive performance [*see* the online supplement]) was 0.77, 0.83, 0.91, and 0.94 for a 1-, 2-, 7-, and 14-night diagnosis, respectively. Predictive performance tended to saturate after 14 days (Figure E5). A similar predictive performance was observed in sensitivity analyses (Figures E6 and E7).

Misclassification of OSA severity (mild, moderate, and severe) based on a one-night diagnosis was high for participants who had mild and/or moderate OSA according to the reference values (Figure 2). Mild and moderate OSA were only characterized correctly on 54% and 52% of the nights respectively, whereas no-OSA (AHI < 5 events/h) and severe OSA (AHI ⩾ 30 events/h) were correctly classified in 85% and 77% of cases, respectively.

Associations between inconclusive diagnostic status and AHI variability, independently of demographic factors, are presented in Table E3. Age, sex, and body mass index effects were observed but were considered likely to be negligible compared with AHI variability (Table E3).

Discussion

Data from novel, noninvasive under-mattress sensor technology acquired across an average of 8 months of in-home nightly monitoring in a large, nonrandomly selected population sample from 20 countries/regions indicates a global prevalence of moderate to severe OSA of approximately 20%. Misdiagnosis probability based on a single night of testing, as is current standard clinical practice, was high at approximately 20% in the current selected population sample and increased to approximately 50% in people with mild to moderate OSA.

The main strengths of this study approach include the simplicity of objective monitoring in the real-world home sleeping environment rather than an in-laboratory setting, and many months rather than one night of recording. Thus, estimates of OSA prevalence and impacts of night-to-night variability were based on a very large volume of physiological and clinical information
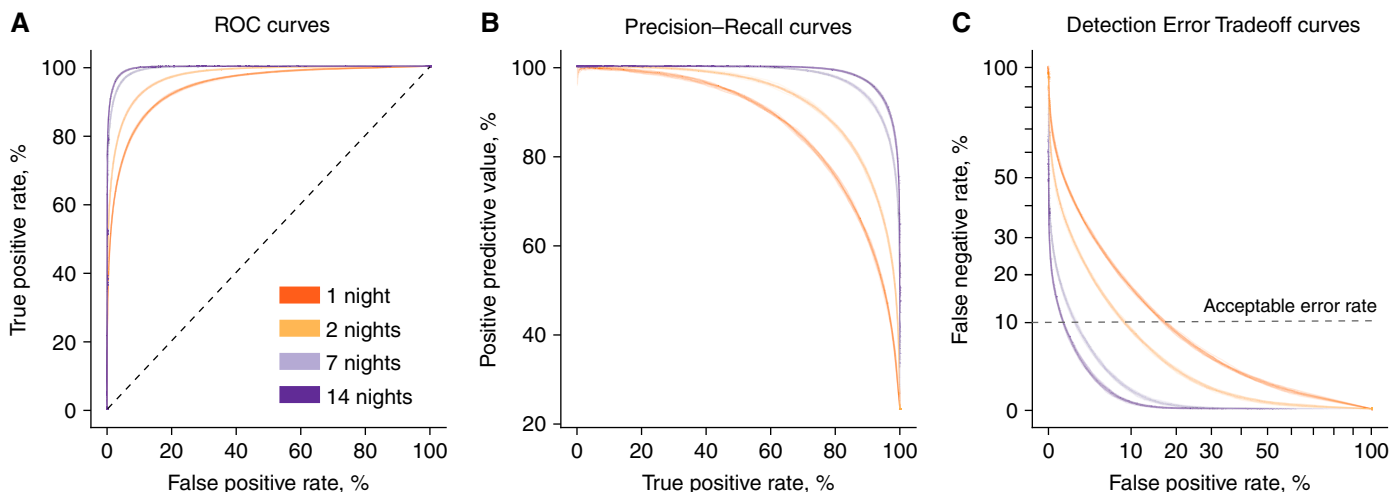
Figure 1. Diagnostic performance of an X-night diagnosis assessed using (*A*) receiver operator characteristic curves, (*B*) precision–recall curves, and (*C*) detection error tradeoff curves for a 1-night (dark orange), 2-night (orange), 7-night (light purple), and 14-night (dark purple) assessment period. ROC = receiver operating characteristic.

collected per night, including nearly 12 million nights of data collected over an average of more than 8 months. As such, this dataset is the largest standardized, objective assessment of OSA collected to date. Data were also acquired across diverse geographical locations, with large datasets available in 20 countries to provide the first multinight, objective global estimates of OSA prevalence. Finally, the lengthy recording duration (between 1 and 8 mo) allowed for much more precise quantification of misdiagnosis probability owing to AHI variability than has previously been possible. This provides a clear advantage over other

recent studies that have derived misdiagnosis probability from a smaller number of nights (<14 nights) and smaller number of participants.

Both biological and measurement/ technical factors may importantly contribute to night-to-night variation in AHI. Sources of potential biological variation in AHI include body position, environmental effects, alcohol, and medication use. A key contributor to measurement/technical AHI variability with conventional in-laboratory polysomnography is high manual interscorer variability (8), particularly

problematic for arousal and hypopnea quantification. Polysomnography versus under-mattress sensor data have different strengths/weaknesses with respect to AHI detection. For example, AHI derived from the under-mattress sensor includes fewer input variables in which to detect respiratory events than the more extensive and direct but also more invasive monitoring approaches of conventional polysomnography. Thus, the ability to accurately identify respiratory events and differentiate between different types of respiratory events may be more challenging with under-mattress sensor
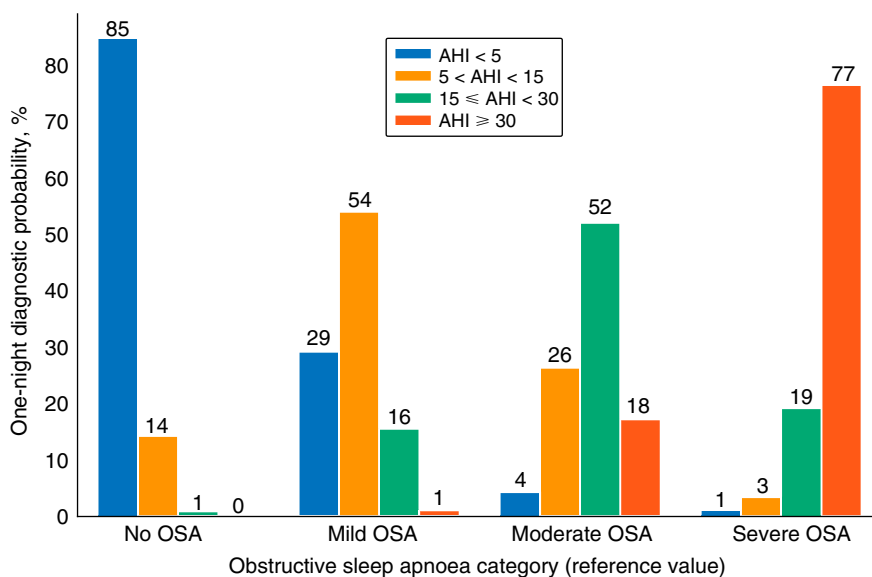

Figure 2. One-night diagnostic probability of obstructive sleep apnea (OSA) severity. AHI = apnea–hypopnea index.

technology. There may also be currently unknown technical reasons that cause higher night-to-night variation with the new under mattress device compared with polysomnography and other sleep study systems. However, an advantage of under-mattress sensor technology is that scoring decisions are made automatically according to objective preset criteria based on the acquired physiological signals and thus are not prone to human scoring variability. These differences may explain, at least in part, individual differences in AHI values on any given night between measurement approaches. This requires further investigation. Nonetheless, recognizing these potential caveats, published validation data (16) and our own validation assessment (*see* the online supplement) show reasonably good agreement in AHI quantified via the under-mattress device compared with in-laboratory polysomnography.

Most of the participants in the current study were male, and likely from moderate to high socioeconomic backgrounds, self-selected by the voluntary purchase of this relatively recently released consumer device. Therefore, this study sample may not be representative of the wider general population and does not resolve the long-standing issue in sleep research of underrepresentation of women in such analyses, and the underlying assumption that thresholds derived from studies primarily of men will be valid when generalized to women. Nonetheless, overall demographics data for the cohort were comparable to World Health Organization and United Nations data. In addition, OSA prevalence in this study was calculated from a nonrandomly selected population that voluntarily purchased the device, possibly owing to concerns of snoring or sleep apnea. If so, reported OSA prevalence may have been overestimated.

However, despite these potential limitations, our multinight moderate to severe OSA prevalence estimates are consistent with recent single-night population study estimates of between 20% and 40% (1, 18). The current multicountry estimates are, however, more consistent than the previously reported findings by Benjafield and colleagues, which ranged from 3% to 40%

(1). This is likely owing to the standardized approach to data collection in the current study compared with the global OSA prevalence estimates derived by Benjafield and colleagues (1), which relied on a literature search of 16 population-based studies and extrapolation to the remaining countries based on matched demographics and the associated limitations of single-night testing.

Misclassification probabilities derived in this study are also in accordance with published data, where misclassification probabilities of a one-night diagnosis range between 10% and 60% (9–14). The wide range of misclassification probability in the literature likely relates, at least in part, to differences in the selected populations which vary depending on the mean AHI, as shown in the current study. For example, participants with an AHI close to 15 events per hour showed a misdiagnosis probability of approximately 50–60%. Overall, most participants with an AHI between 5 and 35 events/h demonstrated a misdiagnosis probability of more than 10%. However, confidence of AHI probability, even with single-night testing, increases in people with an AHI of more than 35 events/h. This knowledge has immediate implications for current clinical decision-making.

The current findings highlight the important new information acquired from multinight monitoring of sleep and OSA. This includes increased confidence of OSA diagnosis, particularly in those with mild or moderate to severe OSA. These findings also pave the way for future prospective large-scale investigations to gain knowledge of the effects of different sleep patterns and night-to-night variability in OSA severity on key health consequences such as altered mood, mental health, sleepiness, workplace and traffic accident, cognitive impairment, and cardiometabolic risks. Indeed, recent preliminary findings indicate high internight variability in OSA severity in people with atrial fibrillation whereby nights of increased OSA severity are associated with increased risk of next-day atrial fibrillation events (21).

Multinight testing approaches could also be invaluable for clinical trials both in terms of assessing eligibility criteria and for long-term monitoring of therapeutic interventions. Determining the potential

differential effects of responses to therapies in people with versus without major internight variability in OSA severity may also be informative. Indeed, standard one-size-fits-all treatment approaches may not be appropriate and may explain, at least in part, the heterogeneity of results from prior treatment trials (22, 23). Furthermore, the study findings directly challenge the use of AHI cutoffs from single-night assessments to diagnose and manage OSA given the large misdiagnosis probability around clinical thresholds. Several studies have clearly shown that OSA is multifaceted and that specific traits and clinical subtypes (e.g., co-occurrence of insomnia, sleepiness, sleep fragmentation, and heart rate surges to apneic events) can help to identify people at greater risk of adverse health outcomes (24–29). Similar to another recent study (9), this study assumed that the average AHI across multiple nights is a reliable marker of OSA severity. However, other metrics may be more useful to better characterize OSA severity and its consequences. Thus, well-designed studies that compare the predictive performance of multinight versus single-night metrics of OSA severity to detect and track key health outcomes and disease consequences are required.

The current findings also highlight the potential utility for noninvasive multinight assessments of sleep and OSA in the home to support current clinical diagnosis and management practices, which may provide benefits in terms of cost-effectiveness and greater access to care over current routine practice. Nonetheless, our findings do not preclude the need for gold-standard polysomnography, particularly in more complex cases or in people with major or multiple comorbidities for whom additional information on hypoxemia and OSA endotypes may be clinically indicated and informative (30–32). Further prospective work in randomly selected populations is required to investigate the generalizability of the current findings. ∎

References

1. Benjafield AV, Ayas NT, Eastwood PR, Heinzer R, Ip MSM, Morrell MJ, *et al.* Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. *Lancet Respir Med* 2019;7: 687–698.
2. Kendzerska T, Gershon AS, Hawker G, Leung RS, Tomlinson G. Obstructive sleep apnea and risk of cardiovascular events and all-cause mortality: a decade-long historical cohort study. *PLoS Med* 2014;11: e1001599.
3. Lang CJ, Appleton SL, Vakulin A, McEvoy RD, Vincent AD, Wittert GA, *et al.* Associations of undiagnosed obstructive sleep apnea and excessive daytime sleepiness with depression: an Australian population study. *J Clin Sleep Med* 2017;13:575–582.
4. Terán-Santos J, Jiménez-Gómez A, Cordero-Guevara J; Cooperative Group Burgos-Santander. The association between sleep apnea and the risk of traffic accidents. *N Engl J Med* 1999;340:847–851.
5. Appleton SL, Vakulin A, McEvoy RD, Vincent A, Martin SA, Grant JF, *et al.* Undiagnosed obstructive sleep apnea is independently associated with reductions in quality of life in middle-aged, but not elderly men of a population cohort. *Sleep Breath* 2015;19:1309–1316.
6. Punjabi NM, Caffo BS, Goodwin JL, Gottlieb DJ, Newman AB, O'Connor GT, *et al.* Sleep-disordered breathing and mortality: a prospective cohort study. *PLoS Med* 2009;6:e1000132.
7. Epstein LJ, Kristo D, Strollo PJ Jr, Friedman N, Malhotra A, Patil SP, *et al.*; Adult Obstructive Sleep Apnea Task Force of the American Academy of Sleep Medicine. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. *J Clin Sleep Med* 2009;5:263–276.
8. Thomas RJ, Chen S, Eden UT, Prerau MJ. Quantifying statistical uncertainty in metrics of sleep disordered breathing. *Sleep Med* 2020; 65:161–169.
9. Punjabi NM, Patil S, Crainiceanu C, Aurora RN. Variability and misclassification of sleep apnea severity based on multi-night testing. *Chest* 2020;158:365–373.
10. Skiba V, Goldstein C, Schotland H. Night-to-night variability in sleep disordered breathing and the utility of esophageal pressure monitoring in suspected obstructive sleep apnea. *J Clin Sleep Med* 2015;11:597–602.
11. Stöberl AS, Schwarz EI, Haile SR, Turnbull CD, Rossi VA, Stradling JR, *et al.* Night-to-night variability of obstructive sleep apnea. *J Sleep Res* 2017;26:782–788.
12. Tschopp S, Wimmer W, Caversaccio M, Borner U, Tschopp K. Night-to-night variability in obstructive sleep apnea using peripheral arterial tonometry: a case for multiple night testing. *J Clin Sleep Med* 2021;17: 1751–1758.
13. Joosten SA, O'Donoghue FJ, Rochford PD, Barnes M, Hamza K, Churchward TJ, *et al.* Night-to-night repeatability of supine-related obstructive sleep apnea. *Ann Am Thorac Soc* 2014;11:761–769.
14. Anitua E, Duran-Cantolla J, Almeida GZ, Alkhraisat MH. Predicting the night-to-night variability in the severity of obstructive sleep apnea: the case of the standard error of measurement. *Sleep Sci* 2019;12:72–78.
15. Roeder M, Bradicich M, Schwarz EI, Thiel S, Gaisl T, Held U, *et al.* Night-to-night variability of respiratory events in obstructive sleep apnoea: a systematic review and meta-analysis. *Thorax* 2020;75:1095–1102.
16. Edouard P, Campo D, Bartet P, Yang RY, Bruyneel M, Roisman G, *et al.* Validation of the Withings Sleep Analyzer, an under-the-mattress device for the detection of moderate-severe sleep apnea syndrome. *J Clin Sleep Med* 2021;17:1217–1227.
17. Yang R-Y, Bendjoudi A, Buard N, Boutouyrie P. Pneumatic sensor for cardiorespiratory monitoring during sleep. *Biomed Phys Eng Express* 2019;5:055014.
18. Heinzer R, Vat S, Marques-Vidal P, Marti-Soler H, Andries D, Tobback N, *et al.* Prevalence of sleep-disordered breathing in the general population: the HypnoLaus study. *Lancet Respir Med* 2015;3:310–318.
19. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing; 2019. Available from: www.r-project.org.
20. Harrell FE Jr. Regression modeling strategies. Springer; 2015.
21. Linz D, Brooks AG, Elliott AD, Nalliah CJ, Hendriks JML, Middeldorp ME, *et al.* Variability of sleep apnea severity and risk of atrial fibrillation: the VARIOSA-AF study. *JACC Clin Electrophysiol* 2019;5:692–701.
22. McEvoy RD, Antic NA, Heeley E, Luo Y, Ou Q, Zhang X, *et al.*; SAVE Investigators and Coordinators. CPAP for prevention of cardiovascular events in obstructive sleep apnea. *N Engl J Med* 2016;375:919–931.
23. 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;136:1840–1850.
24. Hirotsu C, Betta M, Bernardi G, Marques-Vidal P, Vollenweider P, Waeber G, *et al.* Pulse wave amplitude drops during sleep: clinical significance and characteristics in a general population sample. *Sleep (Basel)* 2020;43:zsz322.
25. Azarbarzin A, Sands SA, Younes M, Taranto-Montemurro L, Sofer T, Vena D, *et al.* The sleep apnea-specific pulse-rate response predicts cardiovascular morbidity and mortality. *Am J Respir Crit Care Med* 2021;203:1546–1555.
26. Kwon Y, Wiles C, Parker BE, Clark BR, Sohn MW, Mariani S, *et al.* Pulse arrival time, a novel sleep cardiovascular marker: the multi-ethnic study of atherosclerosis. *Thorax* 2021;76:1124–1130.
27. Sweetman A, Lack L, McEvoy RD, Smith S, Eckert DJ, Osman A, *et al.* Bi-directional relationships between co-morbid insomnia and sleep apnea (COMISA). *Sleep Med Rev* 2021;60:101519.
28. Lechat B, Scott H, Naik G, Hansen K, Nguyen DP, Vakulin A, *et al.* New and emerging approaches to better define sleep disruption and its consequences. *Front Neurosci* 2021;15:751730.
29. Lechat B, Hansen KL, Melaku YA, Vakulin A, Micic G, Adams RJ, *et al.* A novel EEG derived measure of disrupted delta wave activity during sleep predicts all-cause mortality risk. *Ann Am Thorac Soc* 2021; doi: 10.1513/AnnalsATS.202103-315OC.
30. Azarbarzin A, Sands SA, Stone KL, Taranto-Montemurro L, Messineo L, Terrill PI, *et al.* 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* 2019;40: 1149–1157.
31. Mazzotti DR, Keenan BT, Lim DC, Gottlieb DJ, Kim J, Pack AI. Symptom subtypes of obstructive sleep apnea predict incidence of cardiovascular outcomes. *Am J Respir Crit Care Med* 2019;200:493–506.
32. Eckert DJ. Phenotypic approaches to obstructive sleep apnoea: new pathways for targeted therapy. *Sleep Med Rev* 2018;37:45–59.