Received: 8 August 2018 | Revised: 31 October 2018 | Accepted: 6 November 2018

DOI: 10.1111/jsr.12804

REGULAR RESEARCH PAPER

Journal of Sleep Research

# Comparison of apnoea–hypopnoea index and oxygen desaturation index when identifying obstructive sleep apnoea using type-4 sleep studies

Chamara V. Senaratna[1,2] | Adrian Lowe[1] | Jennifer L. Perret[1,3] | Caroline Lodge[1] | Gayan Bowatte[1] | Michael J. Abramson[4] | Bruce R. Thompson[5] | Garun Hamilton[6,7,*] | Shyamali C. Dharmage[1,*]

[1]Allergy & Lung Health, Melbourne School of Population & Global Health, The University of Melbourne, Melbourne, Australia

[2]University of Sri Jayewardenepura, Nugegoda, Sri Lanka

[3]Institute for Breathing and Sleep (IBAS), Heidelberg/Melbourne, Australia

[4]School of Public Health & Preventive Medicine, Monash University, Melbourne, Australia

[5]Department of Respiratory Medicine, Alfred Health, Central Clinical School, Monash University, Melbourne, Australia

[6]School of Clinical Sciences, Monash University, Clayton, Australia

[7]Department of Lung and Sleep, Monash Health, Clayton, Australia

**Correspondence**
Prof. Shyamali Dharmage, Allergy and Lung Health Unit, Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Level 3, 207, Bouverie Street, Carlton, Vic 3052, Australia.
Email: s.dharmage@unimelb.edu.au

**Funding information**
National Health and Medical Research Council, Grant/Award Number: 299901

## Summary

The concordance of different indices from type-4 sleep studies in diagnosing and categorising the severity of obstructive sleep apnoea is not known. This is a critical gap as type-4 sleep studies are used to diagnose obstructive sleep apnoea in some settings. Therefore, we aimed to determine the concordance between flow-based apnoea–hypopnoea index ($AHI_{flow50\%}$) and oxygen desaturation index ($ODI_{3\%}$) by measuring them concurrently. Using a random sub-sample of 296 from a population-based cohort who underwent two-channel type-4 sleep studies, we assessed the concordance between $AHI_{flow50\%}$ and $ODI_{3\%}$. We compared the prevalence of obstructive sleep apnoea of various severities as identified by the two methods, and determined their concordance using coefficient Kappa(κ). Participants were aged (mean ± SD) 53 ± 0.9 years (48% male). The body mass index was 28.8 ± 5.2 kg $m^{-2}$ and neck circumference was 37.4 ± 3.9 cm. The median $AHI_{flow50\%}$ was 5 (inter-quartile range 2, 10) and median $ODI_{3\%}$ was 9 (inter-quartile range 4, 15). The obstructive sleep apnoea prevalence reported using $AHI_{flow50\%}$ was significantly lower than that reported using $ODI_{3\%}$ at all severity thresholds. Although 90% of those with moderate–severe obstructive sleep apnoea classified using $AHI_{flow50\%}$ were identified by using $ODI_{3\%}$, only 46% of those with moderate–severe obstructive sleep apnoea classified using $ODI_{3\%}$ were identified by $AHI_{flow50\%}$. The overall concordance between $AHI_{flow50\%}$ and $ODI_{3\%}$ in diagnosing and classifying the severity of obstructive sleep apnoea was only fair (κ = 0.32), better for males (κ = 0.42 [95% confidence interval 0.32–0.57] versus 0.22 [95% confidence interval 0.09–0.31]), and lowest for those with a body mass index ≥ 35 (κ = 0.11). In conclusion, $ODI_{3\%}$ and $AHI_{flow50\%}$ from type-4 sleep studies are at least moderately discordant. Until further evidence is available, the use of $ODI_{3\%}$ as the measure of choice for type-4 sleep studies is recommended cautiously.

KEYWORDS
agreement, home sleep studies, home sleep-testing, oxygen desaturation index, portable

*G. H. and S. C. D. contributed equally to this study.

  

## 1 | INTRODUCTION

Given the increasing prevalence of obstructive sleep apnoea (OSA) over the last decade (Senaratna et al., 2017), there has been increasing attention on diagnostic methods. The gold-standard diagnostic test for OSA is attended, in-laboratory polysomnography, which is also known as type-1 sleep study (Qaseem et al., 2014). Due to logistical and financial constraints, other types of studies, which use portable sleep study devices and can be performed at home, are also used to diagnose OSA. These are called type-2, -3 or -4 sleep studies based on the number and complexity of data channels they use (Collop et al., 2007; Qaseem et al., 2014), and play an important role in the diagnosis and management of OSA.

How the indices that are used to define OSA, namely, apnoea–hypopnoea index (AHI), respiratory disturbance index (RDI) and/or oxygen desaturation index (ODI; Chai-Coetzer et al., 2014; Dawson et al., 2015), are generated depends on the type of sleep study. Study types 1–3 generate AHI (or RDI) utilising airflow and oxygen saturation, and are usually scored according to certain "rules" – most commonly those published by the American Academy of Sleep Medicine (AASM) (1999, Iber, Israel, Chesson, & Quan, 2007; Berry et al., 2012). However, AHI generated from a type-4 sleep study is generally based solely on nasal airflow and does not take oxygen desaturation into account ($AHI_{flow}$). Some type-4 sleep studies also commonly generate ODI, either solely or in addition to airflow-based AHI. In type-4 studies these indices are typically based on auto-analysis using testing equipment's software.

As type-4 sleep studies have shown good diagnostic utility (Qaseem et al., 2014), they are being increasingly used to diagnose OSA, especially in resource-poor settings (Gantner et al., 2010). Both $AHI_{flow}$ and ODI measured in type-4 sleep studies have been shown to correlate well with AHI measured in type-1 sleep studies (Erman, Stewart, Einhorn, Gordon, & Casal, 2007; Netzer, Eliasson, Netzer, & Kristo, 2001). There is some evidence that ODI better estimates respiratory events (Escourrou et al., 2015) and, furthermore, provides a more robust signal (Gantner et al., 2010). However, whether the diagnosis and severity classification of OSA varies based on the chosen index when type-4 portable sleep study devices are used has not been previously investigated. This knowledge is of importance, as accurate severity classification of OSA has prognostic and management implications. Given that some type-4 portable sleep study devices offer the opportunity to choose between independently-measured $AHI_{flow}$ and ODI when making a diagnosis of OSA, we aimed to determine the correlation between $AHI_{flow}$ and ODI when concurrently measured using a type-4 portable sleep study device and their concordance when used to describe the prevalence and classification of the severity of OSA.

## 2 | MATERIALS AND METHODS

We used data from the sixth decade follow-up of the Tasmanian Longitudinal Health Study (TAHS). The TAHS cohort was originally recruited in 1968 to study chronic respiratory diseases and allergies in 8,583 Tasmanian school children born in 1961 (probands; Matheson et al., 2017). They were followed up in 1974, 1979, 1991, 2002, 2010 and 2012. The last follow-up was completed in 2016 and data were collected from 3,609 probands who could be traced; 74% of them attended a respiratory (not sleep) laboratory study.

### 2.1 | Sleep studies

A random sample of 772 from among respiratory laboratory attendees were invited to undergo type-4 sleep studies using ApneaLink™ device (ResMed, Bella Vista, Australia). Those who agreed to participate were given instructions on setting up ApneaLink™ at home, switching it on before going to bed, and switching it off after getting up in the morning. ApneaLink™ recorded the following signals: nasal air-flow, snoring, oxygen desaturation, respiratory effort and pulse rate. ApneaLink™ devices were returned to the laboratory after each use, and data were downloaded using ApneaLink™ version 9.2.0 proprietary software. These were then auto-analysed using this software and user-defined criteria. A random sample of 10% of the records was manually examined to check the accuracy of the auto-analysis. The criterion for apnoea was a reduction in airflow by $\geq 80\%$ for at least 10 s, for hypopnoea there was a reduction in airflow by $\geq 50\%$ for at least 10 s, and for oxygen desaturation event a reduction in oxygen saturation by at least 3% from the baseline.

### 2.2 | Definitions

Those who had both oxygen desaturation information ($ODI_{3\%}$) and flow-based apnoea–hypopnoea ($AHI_{flow50\%}$) information for at least 4 hr of sleep were included in the analysis. $ODI_{3\%}$ and $AHI_{flow50\%}$ thresholds of $\geq 5$, $\geq 15$ and $\geq 30$ events per hr were used to categorise participants as having any, moderate–severe and severe OSA, respectively.

Based on the body mass index (BMI), participants were categorised as being normal weight ($< 25\ kg\ m^{-2}$), overweight ($\geq 25$ and $< 30\ kg\ m^{-2}$), obese class-I ($\geq 30$ and $< 35\ kg\ m^{-2}$), obese class-II ($\geq 35$ and $< 40\ kg\ m^{-2}$) and obese class-III ($\geq 40\ kg\ m\text{-}^{2}$; World Health Organization, 2000). Obese class-I was defined as obese, and classes-II and -III were collectively defined as morbidly obese. The prevalence of OSA of various severities was reported based on the above thresholds using $AHI_{flow50\%}$ and $ODI_{3\%}$ criteria separately.

### 2.3 | Analysis

Data were analysed using Stata/SE 14.1 software (StataCorp LP, College Station, TX, USA). The correlation between $AHI_{flow50\%}$ and $ODI_{3\%}$ was determined using Pearson's correlation coefficient. Bland–Altman plots were used to examine the agreement between $AHI_{flow50\%}$ and $ODI_{3\%}$. The prevalence of OSA at different OSA severity thresholds as determined using $AHI_{flow50\%}$ and $ODI_{3\%}$ was compared using Pearson's $\chi^2$-test. The severity of OSA classified using $ODI_{3\%}$ and $AHI_{flow50\%}$ was cross-tabulated, and differences in distribution were also examined using the $\chi^2$-test. Cohen's Kappa ($\kappa$)

coefficient was used to check the concordance between $AHI_{flow50\%}$ and $ODI_{3\%}$ in classifying OSA severity. The concordance was considered poor if the Kappa coefficient was < 0.00, slight if 0.00–0.20, fair if 0.21–0.40, moderate if 0.41–0.60, substantial if 0.61–0.80, and almost perfect if 0.81–1.00 (Landis & Koch, 1977).

This study was approved by the Human Research Ethics Committee of the University of Melbourne (approval number 040375). Participants provided written informed consent.

## 3 | RESULTS

Out of the 772 who were invited to undergo sleep studies, 137 declined and another 211 who agreed could not complete the sleep studies due to various reasons (Supporting Information Figure S1). Out of the remaining 424 (54.9%), 296 had both airflow and oxygen saturation recordings for at least 4 hr. Their basic characteristics are shown in Table 1. They were aged 53 years, overweight, and had a high prevalence of OSA defined using $ODI_{3\%}$ and $AHI_{flow50\%}$.

There was a strong correlation between $AHI_{flow50\%}$ and $ODI_{3\%}$ in the overall sample (Pearson's $r = .85$; 95% confidence interval [CI] 0.82, 0.88; $p < .001$). A moderation effect by gender is seen ($p$ for moderation effect < .001), where males had a significantly stronger correlation ($r = .90$; 95% CI 0.86, 0.93; $p < .001$) than females ($r = .65$; 95% CI 0.55, 0.73; $p < .001$). Similar gender differences were also seen across all BMI categories, although statistical significance was seen only in overweight and obese class-I categories (Table 2).

In the total sample, the correlation was weakest in those who had normal BMI ($r = .65$; Table 2), and this significantly increased gradually through overweight to obese ($r = .91$), then decreased in morbidly obese ($r = .78$). These differences were statistically significant between BMI categories from normal weight to obese. However, the correlation in morbidly obese was not statistically different from those who were normal weight or overweight (Table 2). This trend was also seen in males. In females, however, such a trend was not present, and the correlation was not statistically different between those in normal weight, overweight, obese and morbidly obese categories. Intra-class correlation coefficients (ICCs) were almost identical to these (Table S1).

**TABLE 1** Basic characteristics of the sample ($n = 296$)

| Characteristic | Mean ± SD or *N* (%) or median (range; inter-quartile range) |
|---|---|
| Age (years) | 52.9 ± 0.9 |
| Gender (male) | 143 (48.3) |
| BMI (kg $m^{-2}$) | 28.8 ± 5.2 |
| Neck circumference (cm) | 37.4 ± 3.9 |
| Apnoeas and hypopnoeas per hr | 5 (0, 97; 2, 10) |
| Oxygen desaturation events per hr | 9 (1, 92; 4, 15) |
| $AHI_{flow50\%} \geq 5$ | 161 (54.4) |
| $AHI_{flow50\%} \geq 15$ | 41 (13.8) |
| $ODI_{3\%} \geq 5$ | 221 (71.7) |
| $ODI_{3\%} \geq 15$ | 81 (21.4) |

$AHI_{flow50\%}$, flow-based apnoea–hypopnea index (using 50% drop in nasal pressure); BMI, body mass index; $ODI_{3\%}$, oxygen desaturation index (using 3% drop in oxygen saturation).

Bland–Altman plots for the total sample and for each gender showed a higher number of respiratory events in general when $ODI_{3\%}$ was used compared with when $AHI_{flow50\%}$ was used (Figures 1 and 2). The bias (mean difference between $ODI_{3\%}$ and $AHI_{flow50\%}$) for the total sample was 3.5. The limits of agreement were wide (lower and upper limits of agreement −9.9, 16.8). There was no difference between males (3.2; limits −10.2, 17.0) and females (3.8; limits −9.2, 16.7). The bias increased with BMI, being 0.8 (−6.9, 8.5) for those with normal weight, 2.5 (−7.8, 12.8) for overweight, 4.4 (−9.3, 18.1) for obese class-I, 5.8 (−4.8, 16.5) for obese class-II, and 18.2 (−6.5, 42.8) for obese class-III (Figures S2 and S3). At $AHI_{flow50\%}$ and $ODI_{3\%}$ thresholds of ≥ 5, ≥ 15 and ≥ 30, the use of $AHI_{flow50\%}$ underestimated the OSA prevalence, respectively, by 27%, 50% and 48% compared with the use of $ODI_{3\%}$ (Table 3). Differences between classification by $AHI_{flow50\%}$ and $ODI_{3\%}$ at all of these thresholds were statistically significant ($p < .001$). Similarly, the use of $ODI_{3\%}$ identified significantly more participants as having mild OSA, moderate OSA and severe OSA ($p < .001$ for all) than when $AHI_{flow50\%}$ was used. This effect was proportionately more pronounced in the classification of moderate and severe OSA, where the use of $ODI_{3\%}$ identified twice as many participants to have moderate OSA and severe OSA than use of $AHI_{flow50\%}$.

The concordance between $ODI_{3\%}$ and $AHI_{flow50\%}$ in classifying OSA severity was only fair (Landis & Koch, 1977; 54.4% similarly classified; $\kappa = 0.32$, 95% CI 0.24–0.40). The highest discordance in severity classification was seen when 57% of those who were classified as moderate OSA by $ODI_{3\%}$ were identified as mild OSA by $AHI_{flow50\%}$ (Table 4). Furthermore, 28% of those who were classified as severe OSA by $ODI_{3\%}$ were also identified as mild OSA by AHIflow50%. In contrast, all those who were classified as having severe OSA by AHIflow50% were similarly classified by ODI3%, and nearly 86% of those who were classified as moderate OSA by $AHI_{flow50\%}$ were classified as moderate or severe OSA by $ODI_{3\%}$. These differences were statistically significant ($p$ for Fisher's exact test < .001).

The concordance between $ODI_{3\%}$ and $AHI_{flow50\%}$ in classifying OSA severity was affected by gender and BMI. The concordance in males was moderate (59.4% similarly classified; $\kappa = 0.42$, 95% CI 0.32–0.57), but only fair in females (49.7% similarly classified; $\kappa = 0.22$, 95% CI 0.09–0.31). Similarly, the concordance was fair in those who had normal weight (63.6% similarly classified; $\kappa = 0.31$, 95% CI 0.13–0.53), who were overweight (59.2% similarly classified; $\kappa = 0.38$, 95% CI 0.23–0.52) and who were in obese class-I (48.8% similarly classified; $\kappa = 0.24$, 95% CI 0.09–0.38), but was only slight for those who were morbidly obese (33.3% similarly classified; $\kappa = 0.11$, 95% CI −0.06 to 0.30).

The ICC between auto-analyses and manual analyses of the random sub-sample of 10% ($n = 30$) was 0.9 (95% CI 0.9–1.0) for both AHI and ODI. The kappa coefficients for agreement for classification of OSA severity when autoscoring and manual scoring were used

**TABLE 2** Correlation between $ODI_{3\%}$ and $AHI_{flow50\%}$ at different BMI thresholds and categories

| | Pearson's *r* (95% CI); *p* (*n*) | | |
|---|---|---|---|
| **BMI category/threshold (kg m$^{-2}$)** | **Total** | **Male** | **Female** |
| Normal weight (< 25) | .65 (0.50, 0.76); < .001 (77) | .69 (0.43, 0.84); < .001 (29) | .60 (0.38, 0.75); < .001 (48) |
| Overweight (≥ 25 and < 30) | .85 (0.79, 0.90); < .001 (103) | .87 (0.79, 0.92); < .001 (66) | .67 (0.44, 0.82); < .001 (37) |
| Obese class-I (≥ 30 and < 35; obese) | .91 (0.87, 0.94); < .001(80) | .95 (0.91, 0.98); < .001 (37) | .62 (0.40, 0.78); < .001 (43) |
| Obese classes-II & -III[a] (≥ 35; morbidly obese) | .78 (0.60, 0.88); < .001 (36) | .80 (0.39, 0.95); .003 (11) | .62 (0.30, 0.82); < .001 (25) |
| ≥ 25 kg m$^{-2}$ | .86 (0.82, 0.89); < .001 (219) | .90 (0.86, 0.93); < .001 (114) | .63 (0.50, 0.73); < .001 (105) |
| ≥ 30 kg m$^{-2}$ | .86 (0.81, 0.90); < .001(116) | .92 (0.86, 0.95); < .001 (48) | .61 (0.44, 0.74); < .001 (68) |

BMI, body mass index; CI, confidence interval.
[a]Obese classes II and II were combined due to small numbers in these categories.

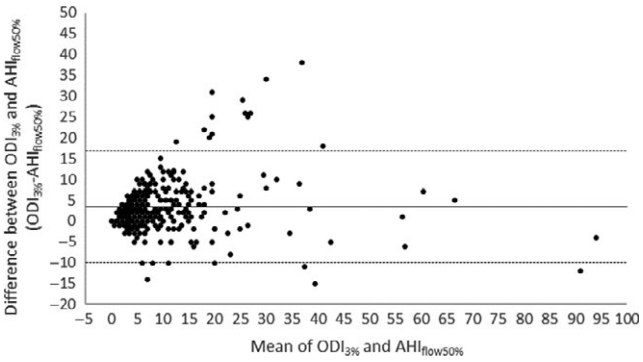


**FIGURE 1** Bland–Altman plot for the distribution of apnoea–hypopnoea index ($AHI_{flow50\%}$) and oxygen desaturation index ($ODI_{3\%}$) in the total sample

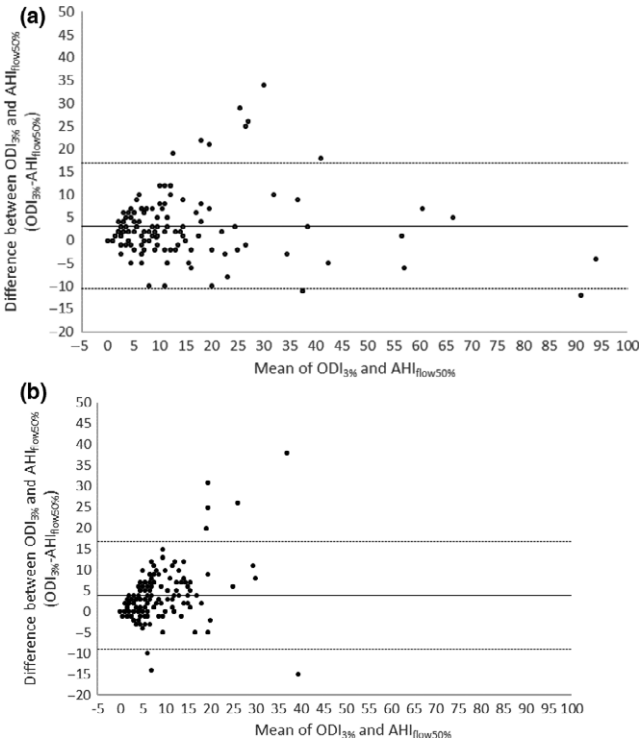


**FIGURE 2** Bland–Altman plot for the distribution of apnoea–hypopnoea index ($AHI_{flow50\%}$) and oxygen desaturation index ($ODI_{3\%}$) in (a) males and (b) females

were found to be 0.7 ± 0.1 (absolute agreement 80%) for AHI and 0.8 ± 0.1 (agreement 87%) for ODI.

## 4 | DISCUSSION

We found that $AHI_{flow50\%}$ (based on airflow only) significantly underestimated respiratory events and OSA prevalence at all thresholds compared with the use of concurrently measured $ODI_{3\%}$ from the same type-4 sleep study device. The concordance of $AHI_{flow50\%}$ and $ODI_{3\%}$ in classifying the severity of OSA was only fair and worsened with increasing BMI, but was better in males than in females. Although 90% of those with moderate or severe OSA classified using $AHI_{flow50\%}$ were also identified by using $ODI_{3\%}$, only 46% of those who had moderate or severe OSA classified using $ODI_{3\%}$ were identified by using $AHI_{flow50\%}$. These levels of disagreement were not clearly reflected by the corresponding correlations, which were high.

Both $AHI_{flow}$ and ODI from type-4 portable sleep study devices have been shown to have acceptable diagnostic utility when compared with type-1 sleep studies (Erman et al., 2007; Netzer et al., 2001). When portable devices have been used in sleep clinic populations and analysed manually, the AHIs from portable devices have shown to have good correlation and concordance with those from polysomnography (Ayappa, Norman, Suryadevara, & Rapoport, 2004). The important differences in our study were that it was conducted in the general population and results from the testing were auto-analysed. In a previous validation using auto-analysis, the $AHI_{flow}$ from the same type-4 portable sleep study device as used in our study has been shown to under-report the respiratory events, but only in those with severe OSA, where the AHI was ≥ 30 events per hr (Erman et al., 2007). Type-4 portable sleep study devices providing a flow-based AHI ($AHI_{flow}$) do not utilise ancillary measures to score hypopnoeas (arousals or oxygen desaturation), and therefore may either under- or over-score hypopnoeas, depending on the threshold of flow-reduction that is used. When $AHI_{flow}$ is used with ApneaLink™, a conservative flow-reduction of 50% (Erman et al., 2007) is often required to prevent overscoring that is likely if smaller reductions are allowed without either arousal or oxygen desaturation. Although there is some potential for under-reporting respiratory events by ODI from oximetry

**TABLE 3** Prevalence of OSA at different $AHI_{flow50\%}/ODI_{3\%}$ thresholds

| $AHI_{flow50\%}/ODI_{3\%}$ threshold | Prevalence % (95% CI); *N* | | % Difference in prevalence when $AHI_{flow50\%}$ was used compared with when $ODI_{3\%}$ used |
|---|---|---|---|
| | Using $AHI_{flow50\%}$ | Using $ODI_{3\%}$ | |
| No OSA (≥ 5) | 45.6 (40.0, 51.4); 135 | 25.3 (20.7, 30.6); 75 | ↑80.2 |
| Any OSA (≥ 5) | 54.4 (48.6, 60.0); 161 | 74.7 (69.4, 79.3)[a]; 221 | ↓27.2 |
| Mild OSA (≥ 5 to < 15) | 40.5 (35.1, 46.3); 120 | 47.3 (41.6, 53.0)[a]; 140 | ↓14.4 |
| Moderate OSA (≥ 15 to < 30) | 9.5 (6.6, 13.4); 28 | 18.9 (14.8, 23.8)[a]; 56 | ↓49.7 |
| Moderate–severe OSA (≥ 15) | 13.8 (10.3, 18.3); 41 | 27.4 (22.6, 32.8)[a]; 81 | ↓49.6 |
| Severe OSA (≥ 30) | 4.4 (2.6, 7.4); 13 | 8.4 (5.6, 12.2)[a]; 25 | ↓47.6 |

$AHI_{flow50\%}$, flow-based apnoea–hypopnea index (using 50% drop in nasal pressure); CI, confidence interval; $ODI_{3\%}$, oxygen desaturation index (using 3% drop in oxygen saturation); OSA, obstructive sleep apnoea.
[a]$p < .001$ for chi-squared test for comparison between prevalence determined using $AHI_{flow50\%}$ and $ODI_{3\%}$ for the given $AHI_{flow50\%}/ODI_{3\%}$ threshold (within the same row).

**TABLE 4** Agreement between OSA severities determined by $AHI_{flow50\%}$ and $ODI_{3\%}$

| Using $AHI_{flow50\%}$ | Using $ODI_{3\%}$ | | | |
|---|---|---|---|---|
| | No OSA (*n* = 75) *N* (%)[a] | Mild OSA (*n* = 140) *N* (%)[a] | Moderate OSA (*n* = 56) *N* (%)[a] | Severe OSA (*n* = 25) *N* (%)[a] |
| No OSA (*n* = 135) | **61 (20.6)** | 69 (23.3) | 4 (1.4) | 1 (0.3) |
| Mild OSA (*n* = 120) | 14 (4.7) | **67 (22.6)** | 32 (10.8) | 7 (2.4) |
| Moderate OSA (*n* = 28) | 0 (0.0) | 4 (1.4) | **20 (6.8)** | 4 (1.4) |
| Severe OSA (*n* = 13) | 0 (0.0) | 0 (0.0) | 0 (0.0) | **13 (4.4)** |

Bold values indicate OSA severity that have been similarly classified using $AHI_{flow50\%}$ and $ODI_{3\%}$.
OSA, obstructive sleep apnoea; $AHI_{flow50\%}$, flow-based apnoea–hypopnea index (using 50% drop in nasal pressure); $ODI_{3\%}$, oxygen desaturation index (using 3% drop in oxygen saturation); No OSA ($AHI/ODI_{3\%} < 5$); Mild OSA ($AHI/ODI_{3\%} \geq 5$ to $< 15$); Moderate OSA ($AHI/ODI_{3\%} \geq 15$ to $< 30$); Severe OSA ($AHI/ODI_{3\%} \geq 30$).
[a]Percentages are out of a total of 296; Fisher's exact test < 0.001.

(missing respiratory events that have an associated arousal rather than oxygen desaturation; Netzer et al., 2001), there is some evidence that ODI performs better and is more similar to AHI from type-1 sleep studies compared with $AHI_{flow}$ from portable sleep study devices (Escourrou et al., 2015). This suggests that ODI from type-4 portable sleep study devices, rather than $AHI_{flow}$, is a better approximation of the actual respiratory events. Furthermore, it has been shown that use of airflow channels in addition to oximetry does not significantly improve standalone oximetry agreement with type-1 sleep studies (Dawson et al., 2015).

The high correlation between $AHI_{flow50\%}$ and $ODI_{3\%}$ that we observed and its variation by gender and BMI are, to our knowledge, the first such evidence using type-4 portable sleep study devices. The overall correlation we saw is similar to what has been reported for type-3 portable sleep study devices ($r = .9$) and type-1 sleep studies ($r = .80$; Ernst et al., 2016). However, despite the increase in correlation with increasing BMI, the average *difference* between $ODI_{3\%}$ and $AHI_{flow50\%}$ ($ODI_{3\%} - AHI_{flow50\%}$) increased with increasing BMI, from 0.8 in those with normal weight to 18.2 in those in obese class-III. In addition, the proportion of people who were similarly classified by $ODI_{3\%}$ and $AHI_{flow50\%}$ decreased from 64% in those with normal weight to 33% in those who were morbidly obese.

An increase in BMI has been shown to independently and significantly predict both more severe OSA and a greater oxygen desaturation in blood during sleep-disordered breathing, especially in the supine position (Peppard, Ward, & Morrell, 2009). It has been suggested that the likely primary mechanism by which obesity accentuates oxygen desaturation is by the effect of excess weight on reducing static lung volumes, especially functional residual capacity (Ling, James, & Hillman, 2012). This previous evidence therefore suggests that relatively small reductions in ventilation that fail to record a flow-based apnoea or hypopnoea event may lead to oxygen desaturation in obese individuals that is sufficient to record an oxygen desaturation event, compared with no oxygen desaturation in those with normal weight.

Clinically, the importance of establishing the most valid classification for OSA status and severity as derived by type-4 portable sleep study devices outweighs the information gained from correlating indices on a continuous scale (Berry, 2011). Our finding that $ODI_{3\%}$ identified over 90% of those classified as having moderate–severe OSA by $AHI_{flow50\%}$ and that $AHI_{flow50\%}$ failed to identify more than half of those classified as moderate–severe OSA by $ODI_{3\%}$ has significant clinical implications. Given the direct and indirect health and economic costs of OSA (Leger, Bayon, Laaban, & Philip, 2012;

Wittmann & Rodenstein, 2004), including nearly a twofold increase in medical expenditure associated with undiagnosed OSA (Wittmann & Rodenstein, 2004) and billions of dollars of additional healthcare costs resulting from the consequences of undiagnosed OSA (Knauert, Naik, Gillespie, & Kryger, 2015), missing those who may benefit from treatment for OSA (Adult Obstructive Sleep Apnea Task Force of the American Academy of Sleep Medicine, 2009) also has important public health implications.

Furthermore, as oxygen desaturation has been linked with adverse consequences associated with OSA (Bradley & Floras, 2009; Drager, Jun, & Polotsky, 2010; Ryan, Taylor, & Mcnicholas, 2005), detecting those with nocturnal hypoxaemia may be arguably more important than detecting those with apnoeas and hypopnoeas without significant oxygen desaturation. Nevertheless, physicians' confidence in managing patients with OSA based solely on $ODI_{3\%}$ has been shown to be lower than when full polysomnographic data are available, despite the high diagnostic utility of $ODI_{3\%}$ (area under the receiver operating characteristic curve being 0.94 [95% CI 0.90–0.98]) and non-inferior clinical improvement in patients treated based solely on ODI versus full polysomnography data (Chai-Coetzer et al., 2017). This lack of confidence among the physicians may arise from inadequate prior experience in management models solely using $ODI_{3\%}$ data (cf. full polysomnography) than from any inferiority of ODI-based management. Some previous evidence suggests that if the practitioners are adequately trained in managing patients based on ODI alone, clinically important outcomes of the patients managed solely on ODI would not be inferior to those of the patients managed based on full polysomnography data (Antic et al., 2009; Chai-Coetzer et al., 2013).

Our study is the first to report on how the use of $AHI_{flow50\%}$ and $ODI_{3\%}$ from the same type-4 devices affect OSA prevalence and severity classification. Our sample represents the general population, and the findings are unlikely to be materially influenced by specific health conditions that are present in high-risk populations. However, our sample size was relatively small and therefore our findings should be interpreted with caution. Only 296 out of 424 participants who underwent sleep studies (70%) had valid data for both AHI and ODI. Although this is relatively higher than the rate of data loss that has commonly been reported (Kapoor & Greenough, 2015), data losses of up to 33% have been previously reported (Lux, Boehlecke, & Lohr, 2004). Although the prevalence of OSA when $ODI_{3\%} \geq 5$ events per hr was used seems high (72%), similar estimates have been previously reported in population-based studies (Senaratna et al., 2017; Tan et al., 2016) that used AASM 2012 scoring criteria (Berry et al., 2012), which uses the same desaturation levels. However, as the severity of OSA and the symptom-profile in the general population, from where our sample was drawn, are likely to be less severe than in patients referred to sleep centres, it is possible that the indices assessed here may be different from clinical populations. In addition, ODI may not be an accurate measure of OSA in those with severe chronic obstructive pulmonary disease (Lacasse et al., 2011; Lewis, Fergusson, Eaton, Zeng, & Kolbe, 2009) or in physically trained subjects with high lung capacity who could have long apnoea/hypopnoea events without desaturation events. Our findings are also limited by the fact that the performances of both $AHI_{flow50\%}$ and $ODI_{3\%}$ were not compared with full polysomnography, the gold-standard. The analyses of sleep records in our study were done using automated analysis and not manually. Using auto-analysis is important in our study as this is how these type-4 devices are generally used in real-world practice. Our results are, therefore, more relevant to clinical practice than if manual scoring was used. Nevertheless, we performed a sensitivity analysis by selecting a random sub-sample of 10% of studies for manual scoring. There was good reliability and agreement between auto- and manual analysis, similar to what has been reported in previous studies (Nigro, Dibur, Aimaretti, González, & Rhodius, 2011; Nigro, Malnis, Dibur, & Rhodius, 2012).

When comparing $AHI_{flow50\%}$ and $ODI_{3\%}$ from type-4 devices, it is important for clinicians to know whether one metric gives the same or different results compared with the other. Clinicians might assume that either metric can be used interchangeably, whereas we have shown that this is not the case. Although ODI becomes a subset of AHI when full polysomnography is used (where oxygen desaturation is considered when scoring hypopnoeas; Ayappa, Rapaport, Norman, & Rapoport, 2005; Redline et al., 2000), this is not the case when type-4 devices are used. In the latter they are scored independent of each other from different data, hence the importance of our study in guiding decision-making in clinical practice. Furthermore, we have not used the new SCOPER classification system (Collop et al., 2011) to identify the performance metrics of sleep study devices (as the type-1–4 terminology is still in common use). If the SCOPER system is used, the two indices ODI and AHI fall into two different categories of measurement indicating that they do not measure the same underlying conditions.

Although we have shown a clear difference in the performance of $AHI_{flow50\%}$ and $ODI_{3\%}$, it is unclear which of these (or other similar metrics) optimally predict adverse outcomes from OSA. It may also be that the optimal diagnostic metric varies according to the type of outcome measured, for example, cardiovascular versus neurocognitive. Future research on prospective cohorts should help to answer this question, and may also have the opportunity to compare ODI and AHI using different criteria (3% or 4% oxygen desaturation) and different thresholds when defining apnoea and hypopnoea.

In summary, for this middle-aged general population sample, we found that the concordance between concurrently measured $ODI_{3\%}$ and $AHI_{flow50\%}$ from type-4 portable sleep study devices in diagnosis and severity classification of OSA was unsatisfactory despite the high correlation between the two indices on a continuous scale. While we currently favour $ODI_{3\%}$ as the measure of choice for type-4 sleep studies, further adequately powered research to address variations within gender and BMI subgroups is required as these are likely to modify ODI during obstructive events.

## ACKNOWLEDGEMENTS

The authors acknowledge the National Health and Medical Research Council for funding, and the Australian government for graduate-training and scholarships. Contributions by Professors Haydn

Walters, Peter Frith, Alan James, Paul Thomas, Graham Giles, David Jones, Richard Wood-Baker, John Hopper, Melanie Matheson and Lyle Gurrin are appreciated. The authors thank ResMed for providing some of the ApneaLink™ -devices.

## CONFLICT OF INTEREST

G. H. has received equipment from third parties to support research unrelated to the current project. M. A. has received grants from Pfizer and Boehringer-Ingelheim, and assistance with conference attendance from Sanofi, the submitted work. A. L. has received grants from the National Health and Medical Research Council, during the conduct of the study. ResMed loaned some of the ApneaLink™ devices used for this study.

## AUTHOR CONTRIBUTIONS

S. D., G. H., M. A. and B. T. were involved in study design and data acquisition. C. S. analysed the data and drafted successive versions of the manuscript. All authors critically evaluated the successive drafts and approved the final version.

## ORCID

*Chamara V. Senaratna* https://orcid.org/0000-0002-5879-6174

## REFERENCES

Adult Obstructive Sleep Apnea Task Force of the American Academy of Sleep Medicine. (2009). 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. *Journal of Clinical Sleep Medicine*, *5*, 263–276.

American Academy of Sleep Medicine. (1999). Sleep-related breathing disorders in adults: Recommendations for syndrome definition and measurement techniques in clinical research. The Report of an American Academy of Sleep Medicine Task Force. *Sleep*, *22*, 667–689.

Antic, N. A., Buchan, C., Esterman, A., Hensley, M., Naughton, M. T., Rowland, S., … McEvoy, R. D. (2009). A randomized controlled trial of Nurse-led care for symptomatic moderate–severe obstructive sleep apnea. *American Journal of Respiratory and Critical Care Medicine*, *179*, 501–508. https://doi.org/10.1164/rccm.200810-1558oc

Ayappa, I., Norman, R. G., Suryadevara, M., & Rapoport, D. M. (2004). Comparison of limited monitoring using a Nasal-Cannula flow signal to full polysomnography in sleep-disordered breathing. *Sleep*, *27*, 1171–1179. https://doi.org/10.1093/sleep/27.6.1171

Ayappa, I., Rapaport, B. S., Norman, R. G., & Rapoport, D. M. (2005). Immediate consequences of respiratory events in sleep disordered breathing. *Sleep Medicine*, *6*, 123–130. https://doi.org/10.1016/j.sleep.2004.08.005

Berry, R. B. (2011). Uses and limitations of portable monitoring for diagnosis and management of obstructive sleep apnea. *Sleep Medicine Clinics*, *6*, 309–333. https://doi.org/10.1016/j.jsmc.2011.05.002

Berry, R. B., Brooks, R., Gamaldo, C. E., Harding, S. M., Marcus, C. L., & Voughn, B. V. (2012). *The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications*. USA: American Academy of Sleep Medicine.

Bradley, T. D., & Floras, J. S. (2009). Obstructive sleep apnoea and its cardiovascular consequences. *The Lancet*, *373*, 82–93. https://doi.org/10.1016/s0140-6736(08)61622-0

Chai-Coetzer, C. L., Antic, N. A., Hamilton, G. S., McArdle, N., Wong, K., Yee, B. J., … Woodman, R. (2017). Physician decision making and clinical outcomes with laboratory polysomnography or limited-channel sleep studies for obstructive sleep apnea: A randomized trial laboratory PSG or limited-channel sleep studies for OSA. *Annals of Internal Medicine*, *166*, 332–340. https://doi.org/10.7326/m16-1301

Chai-Coetzer, C., Antic, N. A., Rowland, L., Reed, R. L., Esterman, A., Catcheside, P. G., …McEvoy, R. D. (2013). Primary care vs specialist sleep center management of obstructive sleep apnea and daytime sleepiness and quality of life: A randomized trial. *JAMA*, *309*, 997–1004. https://doi.org/10.1001/jama.2013.1823

Chai-Coetzer, C. L., Douglas, J., McEvoy, D., Naughton M., Neill, A., Rochford, P., ... Bhalla, G. (2014). Guidelines for sleep studies in adults: Prepared for the Australasian Sleep Association. (pp. 1–49). Australia: Australasian Sleep Association.

Collop, N. A., Anderson, W. M., Boehlecke, B., Claman, D., Goldberg, R., Gottlieb, D. J., … Schwab, R. (2007). Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients. Portable Monitoring Task Force of the American Academy of Sleep Medicine. *Journal of Clinical Sleep Medicine*, *3*, 737–747.

Collop, N. A., Tracy, S. L., Kapur, V., Mehra, R., Kuhlmann, D., Fleishman, S. A., & Ojile, J. M. (2011). Obstructive sleep apnea devices for Out-Of-Center (OOC) testing: Technology evaluation. *Journal of Clinical Sleep Medicine*, *7*, 531–548.

Dawson, A., Loving, R. T., Gordon, R. M., Abel, S. L., Loewy, D., Kripke, D. F., & Kline, L. E. (2015). Type III home sleep testing versus pulse oximetry: Is the respiratory disturbance index better than the oxygen desaturation index to predict the apnoea-hypopnoea index measured during laboratory polysomnography? *British Medical Journal Open*, *5*, e007956.

Drager, L. F., Jun, J. C., & Polotsky, V. Y. (2010). Metabolic consequences of intermittent hypoxia: Relevance to obstructive sleep apnea. *Best Practice & Research Clinical Endocrinology & Metabolism*, *24*, 843–851. https://doi.org/10.1016/j.beem.2010.08.011

Erman, M. K., Stewart, D., Einhorn, D., Gordon, N., & Casal, E. (2007). Validation of the ApneaLink™ for the screening of sleep apnea: A novel and simple single-channel recording device. *Journal of Clinical Sleep Medicine*, *3*, 387–392.

Ernst, G., Bosio, M., Salvado, A., Dibur, E., Nigro, C., & Borsini, E. (2016). Difference between apnea-hypopnea index (AHI) and oxygen desaturation index (ODI): Proportional increase associated with degree of obesity. *Sleep and Breathing*, *20*, 1175–1183. https://doi.org/10.1007/s11325-016-1330-3

Escourrou, P., Grote, L., Penzel, T., Mcnicholas, W. T., Verbraecken, J., Tkacova, R., … Barbé, F. (2015). The diagnostic method has a strong influence on classification of obstructive sleep apnea. *Journal of Sleep Research*, *24*, 730–738. https://doi.org/10.1111/jsr.12318

Gantner, D., Ge, J. Y., Li, L. H., Antic, N., Windler, S., Wong, K., … Wang, J. G. (2010). Diagnostic accuracy of a questionnaire and simple home monitoring device in detecting obstructive sleep apnoea in a Chinese population at high cardiovascular risk. *Respirology*, *15*, 952–960. https://doi.org/10.1111/j.1440-1843.2010.01797.x

Iber, C., Israel, S. A., Chesson, A. L., & Quan, S. F.; For the American Academy of Sleep Medicine. (2007). The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications. (pp. 57). Westchester, IL: American Academy of Sleep Medicine.

Kapoor, M., & Greenough, G. (2015). Home sleep tests for obstructive sleep apnea (OSA). *The Journal of the American Board of Family Medicine*, *28*, 504–509. https://doi.org/10.3122/jabfm.2015.04.140266

Knauert, M., Naik, S., Gillespie, M. B., & Kryger, M. (2015). Clinical consequences and economic costs of untreated obstructive sleep apnea syndrome. *World Journal of Otorhinolaryngology-Head and Neck Surgery*, *1*, 17–27. https://doi.org/10.1016/j.wjorl.2015.08.001

Lacasse, Y., Sériès, F., Vujovic-Zotovic, N., Goldstein, R., Bourbeau, J., Lecours, R., … Maltais, F. (2011). Evaluating nocturnal oxygen desaturation in COPD – revised. *Respiratory Medicine*, *105*, 1331–1337. https://doi.org/10.1016/j.rmed.2011.04.003

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174. https://doi.org/10.2307/2529310

Leger, D., Bayon, V., Laaban, J. P., & Philip, P. (2012). Impact of sleep apnea on economics. *Sleep Medicine Reviews*, *16*, 455–462. https://doi.org/10.1016/j.smrv.2011.10.001

Lewis, C. A., Fergusson, W., Eaton, T., Zeng, I., & Kolbe, J. (2009). Isolated nocturnal desaturation in COPD: Prevalence and impact on quality of life and sleep. *Thorax*, *64*, 133–138. https://doi.org/10.1136/thx.2007.088930

Ling, I. T., James, A. L., & Hillman, D. R. (2012). Interrelationships between body mass, oxygen desaturation, and apnea-hypopnea indices in a sleep clinic population. *Sleep*, *35*, 89–96. https://doi.org/10.5665/sleep.1592

Lux, L., Boehlecke, B., & Lohr, K. N. (2004). AHRQ technology assessments. *Effectiveness of portable monitoring devices for diagnosing obstructive sleep apnea: Update of a systematic review* (pp. 1–62). Rockville, MD: Agency for Healthcare Research and Quality (US).

Matheson, M. C., Abramson, M. J., Allen, K., Benke, G., Burgess, J. A., Dowty, J. G., … Gurrin, L. C. (2017). Cohort profile: The Tasmanian Longitudinal Health STUDY (TAHS). *International Journal of Epidemiology*, *46*, 407–08i.

Netzer, N., Eliasson, A. H., Netzer, C., & Kristo, D. A. (2001). Overnight pulse oximetry for sleep-disordered breathing in adults: A review. *Chest*, *120*, 625–633. https://doi.org/10.1378/chest.120.2.625

Nigro, C. A., Dibur, E., Aimaretti, S., González, S., & Rhodius, E. (2011). Comparison of the automatic analysis versus the manual scoring from ApneaLink™ device for the diagnosis of obstructive sleep apnoea syndrome. *Sleep and Breathing*, *15*, 679–686. https://doi.org/10.1007/s11325-010-0421-9

Nigro, C. A., Malnis, S., Dibur, E., & Rhodius, E. (2012). How reliable is the manual correction of the autoscoring of a level IV sleep study (ApneaLink™) by an observer without experience in polysomnography? *Sleep and Breathing*, *16*, 275–279. https://doi.org/10.1007/s11325-011-0524-y

Peppard, P. E., Ward, N. R., & Morrell, M. J. (2009). The impact of obesity on oxygen desaturation during sleep-disordered breathing. *American Journal of Respiratory and Critical Care Medicine*, *180*, 788–793. https://doi.org/10.1164/rccm.200905-0773oc

Qaseem, A., Dallas, P., Owens, D. K., Starkey, M., Holty, J. E., & Shekelle, P. (2014). Diagnosis of obstructive sleep apnea in adults: A clinical practice guideline from the American College of Physicians. *Annals of Internal Medicine*, *161*, 210–220. https://doi.org/10.7326/m12-3187

Redline, S., Kapur, V. K., Sanders, M. H., Quan, S. F., Gottlieb, D. J., Rapoport, D. M., … Iber, C. (2000). Effects of varying approaches for identifying respiratory disturbances on sleep apnea assessment. *American Journal of Respiratory and Critical Care Medicine*, *161*, 369–374. https://doi.org/10.1164/ajrccm.161.2.9904031

Ryan, S., Taylor, C. T., & Mcnicholas, W. T. (2005). Selective activation of inflammatory pathways by intermittent hypoxia in obstructive sleep apnea syndrome. *Circulation*, *112*, 2660–2667. https://doi.org/10.1161/circulationaha.105.556746

Senaratna, C. V., Perret, J. L., Lodge, C., Lowe, A. J., Campbell, B. E., Matheson, M. C., … Dharmage, S. C. (2017). Prevalence of obstructive sleep apnea in the general population: A systematic review. *Sleep Medicine Reviews*, *34*, 17–81.

Tan, A., Cheung, Y. Y., Yin, J., Lim, W. Y., Tan, L. W., & Lee, C. H. (2016). Prevalence of sleep-disordered breathing in a multiethnic Asian population in Singapore: A community-based study. *Respirology*, *21*, 943–950. https://doi.org/10.1111/resp.12747

Wittmann, V., & Rodenstein, D. O. (2004). Health care costs and the sleep apnea syndrome. *Sleep Medicine Reviews*, *8*, 269–279. https://doi.org/10.1016/j.smrv.2004.01.002

World Health Organization. (2000). *Obesity: Preventing and managing the global epidemic*. Geneva: World Health Organization.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Senaratna CV, Lowe A, Perret JL, et al. Comparison of apnoea–hypopnoea index and oxygen desaturation index when identifying obstructive sleep apnoea using type-4 sleep studies. *J Sleep Res*. 2018;e12804. https://doi.org/10.1111/jsr.12804