DR. CHAMARA VISANKA SENARATNA (Orcid ID : 0000-0002-5879-6174)

Article type : Regular Research Paper

Title: Comparison of apnoea-hypopnoea index and oxygen-desaturation index when Identifying obstructive sleep apnoea using Type-4 sleep studies

Short title: AHI or ODI in 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], [m*]Shyamali C. Dharmage[1]

[*] Equally contributed

[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, Australia

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

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

[m] Corresponding author

Corresponding author: Prof. Shyamali Dharmage

Postal address: 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

Phone: +61 3 8344 0737

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/jsr.12804

Fax: + 61 3 9349 5815

All authors have seen and approved the final version of the manuscript.

Conflicts of interests: GH has received equipment from third parties to support research unrelated to the current project. MA has received grants from Pfizer and Boehringer-Ingelheim, and assistance with conference attendance from Sanofi, outside the submitted work. AL has received grants from National Health and Medical Research Council, during the conduct of the study. ResMed loaned some of the ApneaLinik™ devices used for this study.

Author contributions: SD, GH, MA and BT were involved in study design and data acquisition. CS analysed the data and drafted successive versions of the manuscript. All authors critically evaluated the successive drafts and approved the final version.

Word count: 4992

Number of references: 40

ABBREVIATION LIST

AASM - American Academy of Sleep Medicine

AHI – Apnoea-hypopnoea index

AHI_{flow} - Flow-based apnoea-hypopnoea index

$AHI_{flow50\%}$ - Flow-based apnoea-hypopnoea index (using 50% reduction of airflow from baseline)

BMI – body-mass index

CI – confidence interval

ODI – Oxygen desaturation index

$ODI_{3\%}$ - Oxygen desaturation index (using 3% reduction of oxygen saturation from baseline)

OR – odds ratio

OSA - Obstructive sleep apnoea

RDI – Respiratory disturbance index

SD – Standard deviation

TAHS - Tasmanian Longitudinal Cohort Study

Summary

Concordance of different indices from type-4 sleep-studies in diagnosing and categorising severity of Obstructive sleep apnoea (OSA) is not known. This is a critical gap as Type-4 sleep studies are used to

diagnose OSA 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 prevalence of OSA of various severity as identified by the two methods and determined their concordance using co-efficient Kappa(κ). Participants were aged (mean±SD) 53±0.9 years (48% male). Body-mass index (BMI) was 28.8±5.2kg/m^2 and neck circumference 37.4±3.9cm. Median $AHI_{flow50\%}$ was 5 [IQR 2, 10] and median $ODI_{3\%}$ 9 [IQR 4, 15]. OSA 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 OSA classified using $AHI_{flow50\%}$ were identified by using $ODI_{3\%}$, only 46% of those with moderate-severe OSA classified using $ODI_{3\%}$ were identified by $AHI_{flow50\%}$. The overall concordance between $AHI_{flow50\%}$ and $ODI_{3\%}$ in diagnosing and classifying severity of OSA was only fair (κ=0.32), better for males (κ=0.42 [95%CI 0.32-0.57] vs 0.22 [95%CI 0.09-0.31]), and lowest for those with BMI≥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, use of $ODI_{3\%}$ as the measure of choice for type-4 sleep-studies is recommended cautiously.

Key-words: portable; home sleep-studies; home sleep-testing; agreement; ODI; AHI

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 a 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 types 2, 3 or 4 sleep-studies based on the number and complexity of data channels they use (Qaseem et al., 2014, Collop et al., 2007) and play an important role in 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) (Dawson et al., 2015, Chai-Coetzer et al., 2014) are generated depends on the type of sleep-study. Types 1-3 studies 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 et al., 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 by testing equipment's software.

As Type-4 sleep-studies have shown good diagnostic utility (Qaseem et al., 2014), they are been 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 et al., 2007, Netzer et al., 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 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 is important knowledge, as correct 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 severity of OSA.

Methods

We used data from the 6th 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 collected data from 3609 probands who could be traced; 74% of them attended a respiratory (not sleep) laboratory study.

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 were manually examined to check the accuracy of the auto-analysis. The criterion for apnoea was a reduction in airflow by ≥80% for at

least 10s, for hypopnoea was a reduction in airflow by ≥50% for at least 10s, and for oxygen desaturation event a reduction in oxygen saturation by at least 3% from the baseline.

Definitions

Those who had both oxygen desaturation information ($ODI_{3\%}$) and flow-based apnoea-hypopnoea ($AHI_{flow50\%}$) information for at least 4 hours of sleep were included in the analysis. $ODI_{3\%}$ and $AHI_{flow50\%}$ thresholds of ≥5, ≥15, and ≥30 events/hour 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 having normal weight (<25kg/m^2), overweight (≥25 and <30 kg/m^2), obese class-I (≥30 and <35 kg/m^2), obese class-II (≥35 and <40 kg/m^2), and obese class-III (≥40 kg/m^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 were reported based on the above thresholds using separately $AHI_{flow50\%}$ and $ODI_{3\%}$ criteria.

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 $ODI_{3\%}$. Prevalence of OSA at different OSA severity thresholds as determined using $AHI_{flow50\%}$ and $ODI_{3\%}$ were compared using Pearson's χ^2 test. Severity of OSA classified using $ODI_{3\%}$ and $AHI_{flow50\%}$ were cross-tabulated and differences in distribution were also examined using the χ^2 test. Cohen's Kappa (κ) coefficient was used to check the concordance between $AHI_{flow50\%}$ and $ODI_{3\%}$ in classifying OSA severity. The concordance was considered poor if Kappa co-efficient 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 and Koch, 1977).

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

Results

Out of the 772 who were invited to undergo sleep-studies, 137 declined and another 211 who agreed could not complete sleep-studies due to various reasons (Figure S1). Out of the remaining

424 (54.9%), 296 had both airflow and oxygen saturation recordings for at least four hours. Their basic characteristics are shown in Table 1. They were aged 53 years, overweight, and had high prevalence of OSA defined using $ODI_{3\%}$ and $AHI_{flow50\%}$.

[Insert Table 1 here]

There was a strong correlation between $AHI_{flow50\%}$ and $ODI_{3\%}$ in the overall sample (Pearson's r=0.85; 95%CI 0.82, 0.88; p<0.001). A moderation effect by gender is seen (p for moderation effect<0.001), where males had a significantly stronger correlation (r=0.90; 95%CI 0.86, 0.93; p<0.001) than females (r=0.65; 95%CI 0.55, 0.73; p<0.001). 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).

[Insert Table 2 here]

In the total sample, the correlation was weakest in those who had normal BMI (r=0.65; Table 2), and this significantly increased gradually through overweight to obese (r=0.91), then decreased in morbidly obese (r=0.78). 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 had 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. Intra-class correlation coefficient (ICC) were almost identical to these (Table S1).

Bland-Altman plots for the total sample and for each gender showed higher number of respiratory events in general when $ODI_{3\%}$ was used compared to when $AHI_{flow50\%}$ was used (Figures 1-2). 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 to 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<0.001). Similarly, the use of $ODI_{3\%}$ identified significantly more participants as having mild OSA, moderate OSA, and severe OSA (p<0.001 for all) than when $AHI_{flow50\%}$ was used. This effect was proportionately more pronounced in classification of moderate and severe OSA, where use of $ODI_{3\%}$ identified twice as many participants to have moderate OSA and severe OSA than use of $AHI_{flow50\%}$.

[Insert Figure 1 here]

[Insert Figure 2 here]

[Insert Table 3 here]

The concordance between $ODI_{3\%}$ and $AHI_{flow50\%}$ in classifying OSA severity was only fair (Landis and Koch, 1977) (54.4% similarly classified; κ=0.32, 95%CI 0.24-0.40). 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 $AHI_{flow50\%}$. In contrast, all those who were classified as having severe OSA by $AHI_{flow50\%}$ were similarly classified by $ODI_{3\%}$ 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<0.001).

[Insert Table 4 here]

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; κ=0.42, 95%CI 0.32-0.57), but only fair in females (49.7% similarly classified; κ=0.22, 95%CI 0.09-0.31). Similarly, the concordance was fair in those who had normal-weight (63.6% similarly classified; κ=0.31, 95%CI 0.13-0.53), who were overweight (59.2% similarly classified; κ=0.38, 95%CI 0.23-0.52), and who were in obese class-I (48.8% similarly classified; κ=0.24, 95%CI 0.09-0.38), but was only slight for those who were morbidly obese (33.3% similarly classified; κ=0.11, 95% CI -0.06-0.30) .
The intraclass correlation coefficient (ICC) between auto- 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 auto- and manual scoring was used were 0.7±0.1 (absolute agreement 80%) for AHI and 0.8±0.1 (agreement 87%) for ODI.

Discussion

We found $AHI_{flow50\%}$ (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 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 et al., 2004). 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 AHI was ≥30 events per hour (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 of the respiratory events by ODI from oximetry (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 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=0.9) and type-1 sleep-studies (r=0.80) (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.

Increase in the 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 et al., 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 et al., 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 individual 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 et al., 2012, Wittmann and Rodenstein, 2004), including nearly a two-fold increase in medical expenditure associated with undiagnosed OSA (Wittmann and Rodenstein, 2004) and billions of dollars of additional healthcare-cost resulting from consequences of undiagnosed OSA (Knauert et al., 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 to the adverse consequences associated with OSA (Bradley and Floras, 2009, Drager et al., 2010, Ryan et al., 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 OSA patients 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 ROC curve being 0.94 (95%CI 0.90-0.98)] and non-inferior clinical improvement in patients treated based solely on ODI vs 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 (Chai-Coetzer et al., 2013, Antic et al., 2009).

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 and Greenough, 2015), data losses of up to 33% has been previously reported (Lux et al., 2004). Although the prevalence of OSA when $ODI_{3\%} \geq 5$ events/hour 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 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 COPD (Lacasse et al., 2011, Lewis et al., 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 et al., 2011, Nigro et al., 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 to 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) (Redline et al., 2000, Ayappa et al., 2005) 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 types 1-4 terminology are 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 e.g. cardiovascular vs 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.

Acknowledgments

We 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. We thank ResMed for providing some of the ApneaLink™-devices.

References

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*, 1999, 22: 667-89.

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: 263-76.

Antic, N. A., Buchan, C., Esterman, A. *et al.* A Randomized Controlled Trial of Nurse-led Care for Symptomatic Moderate–Severe Obstructive Sleep Apnea. *American Journal of Respiratory and Critical Care Medicine*, 2009, 179: 501-08.

Ayappa, I., Norman, R. G., Suryadevara, M. and Rapoport, D. M. Comparison of Limited Monitoring Using a Nasal-Cannula Flow Signal to Full Polysomnography in Sleep-Disordered Breathing. *Sleep*, 2004, 27: 1171-79.

Ayappa, I., Rapaport, B. S., Norman, R. G. and Rapoport, D. M. Immediate consequences of respiratory events in sleep disordered breathing. *Sleep Med.*, 2005, 6: 123-30.

Berry, R. B. Uses and limitations of portable monitoring for diagnosis and management of obstructive sleep apnea. *Sleep Med. Clin.*, 2011, 6: 309-33.

Berry, R. B., Brooks, R., Gamaldo, C. E., Harding, S. M., Marcus, C. L. and Voughn, B. V. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, terminology and technical specifications*. American Academy of Sleep Medicine, USA, 2012 (

Bradley, T. D. and Floras, J. S. Obstructive sleep apnoea and its cardiovascular consequences. *The Lancet*, 2009, 373: 82-93.

Chai-Coetzer, C., Antic, N. A., Rowland, L. and Et Al. Primary care vs specialist sleep center management of obstructive sleep apnea and daytime sleepiness and quality of life: A randomized trial. *JAMA*, 2013, 309: 997-1004.

Chai-Coetzer, C. L., Antic, N. A., Hamilton, G. S. *et al.* Physician Decision Making and Clinical Outcomes With Laboratory Polysomnography or Limited-Channel Sleep Studies for Obstructive Sleep ApneaA Randomized TrialLaboratory PSG or Limited-Channel Sleep Studies for OSA. *Ann. Intern. Med.*, 2017, 166: 332-40.

Chai-Coetzer, C. L., Douglas, J., Mcevoy, D. *et al.* Guidelines for sleep studies in adults: Prepared for the Australasian Sleep Association. In. Australasian Sleep Association, Australia, 2014: 1-69.

Collop, N. A., Anderson, W. M., Boehlecke, B. *et al.* Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients. Portable Monitoring Task Force of the American Academy of Sleep Medicine. *J. Clin. Sleep Med.*, 2007, 3: 737-47.

Collop, N. A., Tracy, S. L., Kapur, V. *et al.* Obstructive Sleep Apnea Devices for Out-Of-Center (OOC) Testing: Technology Evaluation. *Journal of Clinical Sleep Medicine : JCSM : Official Publication of the American Academy of Sleep Medicine*, 2011, 7: 531-48.

Dawson, A., Loving, R. T., Gordon, R. M. *et al.* 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? *BMJ Open*, 2015, 5: e007956.

Drager, L. F., Jun, J. C. and Polotsky, V. Y. Metabolic consequences of intermittent hypoxia: Relevance to obstructive sleep apnea. *Best Pract. Res. Clin. Endocrinol. Metab.*, 2010, 24: 843-51.

Erman, M. K., Stewart, D., Einhorn, D., Gordon, N. and Casal, E. Validation of the ApneaLink™ for the screening of sleep apnea: A novel and simple single-channel recording device. *J. Clin. Sleep Med.*, 2007, 3: 387-92.

Ernst, G., Bosio, M., Salvado, A., Dibur, E., Nigro, C. and Borsini, E. Difference between apnea-hypopnea index (AHI) and oxygen desaturation index (ODI): proportional increase associated with degree of obesity. *Sleep and Breathing*, 2016, 20: 1175–83.

Escourrou, P., Grote, L., Penzel, T. *et al.* The diagnostic method has a strong influence on classification of obstructive sleep apnea. *J. Sleep Res.*, 2015, 24: 730-38.

Gantner, D., Ge, J. Y., Li, L. H. *et al.* Diagnostic accuracy of a questionnaire and simple home monitoring device in detecting obstructive sleep apnoea in a Chinese population at high cardiovascular risk. *Respirology*, 2010, 15: 952-60.

Iber, C., Israel, S. A., Chesson, A. L., Quan, S. F. and For the American Academy of Sleep Medicine The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications. In. American Academy of Sleep Medicine, Westchester, Illinois, 2007: 57.

Kapoor, M. and Greenough, G. Home sleep tests for obstructive sleep apnea (OSA). *The Journal of the American Board of Family Medicine*, 2015, 28: 504-09.

Knauert, M., Naik, S., Gillespie, M. B. and Kryger, M. Clinical consequences and economic costs of untreated obstructive sleep apnea syndrome. *World Journal of Otorhinolaryngology-Head and Neck Surgery*, 2015, 1: 17-27.

Lacasse, Y., Sériès, F., Vujovic-Zotovic, N. *et al.* Evaluating nocturnal oxygen desaturation in COPD – revised. *Respir. Med.*, 2011, 105: 1331-37.

Landis, J. R. and Koch, G. G. The measurement of observer agreement for categorical data. *Biometrics*, 1977, 33: 159-74.

Leger, D., Bayon, V., Laaban, J. P. and Philip, P. Impact of sleep apnea on economics. *Sleep Med. Rev.*, 2012, 16: 455-62.

Lewis, C. A., Fergusson, W., Eaton, T., Zeng, I. and Kolbe, J. Isolated nocturnal desaturation in COPD: prevalence and impact on quality of life and sleep. *Thorax*, 2009, 64: 133-38.

Ling, I. T., James, A. L. and Hillman, D. R. Interrelationships between body mass, oxygen desaturation, and apnea-hypopnea indices in a sleep clinic population. *Sleep*, 2012, 35: 89-96.

Lux, L., Boehlecke, B. and Lohr, K. N. AHRQ Technology Assessments. In), *Effectiveness of Portable Monitoring Devices for Diagnosing Obstructive Sleep Apnea: Update of a Systematic Review*. Agency for Healthcare Research and Quality (US), Rockville (MD), 2004.

Matheson, M. C., Abramson, M. J., Allen, K. *et al.* Cohort Profile: The Tasmanian Longitudinal Health STUDY (TAHS). *Internatinoal Journal of Epidemiology*, 2017, 46: 407–08i.

Netzer, N., Eliasson, A. H., Netzer, C. and Kristo, D. A. Overnight pulse oximetry for sleep-disordered breathing in adults: a review. *Chest*, 2001, 120: 625-33.

Nigro, C. A., Dibur, E., Aimaretti, S., González, S. and Rhodius, E. Comparison of the automatic analysis versus the manual scoring from ApneaLink™ device for the diagnosis of obstructive sleep apnoea syndrome. *Sleep. Breath.*, 2011, 15: 679-86.

Nigro, C. A., Malnis, S., Dibur, E. and Rhodius, E. How reliable is the manual correction of the autoscoring of a level IV sleep study (ApneaLink™) by an observer without experience in polysomnography? *Sleep. Breath.*, 2012, 16: 275-79.

Peppard, P. E., Ward, N. R. and Morrell, M. J. The impact of obesity on oxygen desaturation during sleep-disordered breathing. *Am. J. Respir. Crit. Care Med.*, 2009, 180: 788-93.

Qaseem, A., Dallas, P., Owens, D. K., Starkey, M., Holty, J. E. and Shekelle, P. Diagnosis of obstructive sleep apnea in adults: a clinical practice guideline from the American College of Physicians. *Ann. Intern. Med.*, 2014, 161: 210-20.

Redline, S., Kapur, V. K., Sanders, M. H. *et al.* Effects of Varying Approaches for Identifying Respiratory Disturbances on Sleep Apnea Assessment. *Am. J. Respir. Crit. Care Med.*, 2000, 161: 369-74.

Ryan, S., Taylor, C. T. and Mcnicholas, W. T. Selective activation of inflammatory pathways by intermittent hypoxia in obstructive sleep apnea syndrome. *Circulation*, 2005, 112: 2660-67.

Senaratna, C. V., Perret, J. L., Lodge, C. *et al.* Prevalence of Obstructive Sleep Apnea in the general population: A systematic review. *Sleep Med. Rev.*, 2017, 34: 17-81.

Tan, A., Cheung, Y. Y., Yin, J., Lim, W. Y., Tan, L. W. and Lee, C. H. Prevalence of sleep-disordered breathing in a multiethnic Asian population in Singapore: A community-based study. *Respirology*, 2016, 21: 943-50.

Wittmann, V. and Rodenstein, D. O. Health care costs and the sleep apnea syndrome. *Sleep Med. Rev.*, 2004, 8: 269-79.

World Health Organization *Obesity: preventing and managing the global epidemic*. World Health Organization, Geneva, 2000

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)
Body-mass index (kg/m^2)	28.8±5.2
Neck circumference (cm)	37.4±3.9
Apnoeas and hypopnoeas/Hour	5 (0, 97; 2, 10)
Oxygen desaturation events/Hour	9 (1, 92; 4, 15)
$AHI_{flow50\%}$≥5	161 (54.4)
$AHI_{flow50\%}$≥15	41 (13.8)
$ODI_{3\%}$≥5	221 (71.7)
$ODI_{3\%}$≥15	81 (21.4)

$AHI_{flow50\%}$=Flow-based apnoea-hypopnea index (using 50% drop in nasal pressure); $ODI_{3\%}$=Oxygen desaturation index (using 3% drop in oxygen saturation)

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

BMI category / threshold (kg/m^2)	Pearson's r (95% CI); p (n)		
	Total	**Male**	**Female**
Normal weight (<25)	0.65 (0.50, 0.76); <0.001 (77)	0.69 (0.43, 0.84); <0.001 (29)	0.60 (0.38, 0.75); <0.001 (48)
Overweight (≥25 and <30)	0.85 (0.79, 0.90); <0.001 (103)	0.87 (0.79, 0.92); <0.001 (66)	0.67 (0.44, 0.82); <0.001 (37)
Obese class I (≥30 and <35; obese)	0.91 (0.87, 0.94); <0.001(80)	0.95 (0.91, 0.98); <0.001 (37)	0.62 (0.40, 0.78); <0.001 (43)
Obese classes II & III [a] (≥35; morbidly obese)	0.78 (0.60, 0.88); <0.001 (36)	0.80 (0.39, 0.95); 0.003 (11)	0.62 (0.30, 0.82); <0.001 (25)
≥25 kg/m^2	0.86 (0.82, 0.89); <0.001 (219)	0.90 (0.86, 0.93); <0.001 (114)	0.63 (0.50, 0.73); <0.001 (105)
≥30 kg/m^2	0.86 (0.81, 0.90); <0.001(116)	0.92 (0.86, 0.95); <0.001 (48)	0.61 (0.44, 0.74); <0.001 (68)

$AHI_{flow50\%}$=Flow-based apnoea-hypopnea index (using 50% drop in nasal pressure); $ODI_{3\%}$=Oxygen desaturation index (using 3% drop in oxygen saturation); BMI=body-mass index; [a] Obese classes II and II were combined due to small numbers in these categories

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) *; 221	↓27.2
Mild OSA (≥5 to <15)	40.5 (35.1, 46.3); 120	47.3 (41.6, 53.0) *; 140	↓14.4
Moderate OSA (≥15 to <30)	9.5 (6.6, 13.4); 28	18.9 (14.8, 23.8) *; 56	↓49.7
Moderate-severe OSA (≥15)	13.8 (10.3, 18.3); 41	27.4 (22.6, 32.8) *; 81	↓49.6
Severe OSA (≥30)	4.4 (2.6, 7.4); 13	8.4 (5.6, 12.2) *; 25	↓47.6

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); *p<0.001 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 severity determined by $AHI_{flow50\%}$ and $ODI_{3\%}$

	Using $ODI_{3\%}$			
Using $AHI_{flow50\%}$	**No OSA (n=75)**	**Mild OSA (n=140)**	**Moderate OSA (n=56)**	**Severe OSA (n=25)**
	N (%)[a]	**N (%)[a]**	**N (%)[a]**	**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)**

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 ($AHI/ODI_{3\%}$≥5 to <15); Moderate OSA ($AHI/ODI_{3\%}$ ≥15 to <30); Severe OSA ($AHI/ODI_{3\%}$ ≥30); [a] Percentages are out of the total of 296; Fisher's exact test<0.001

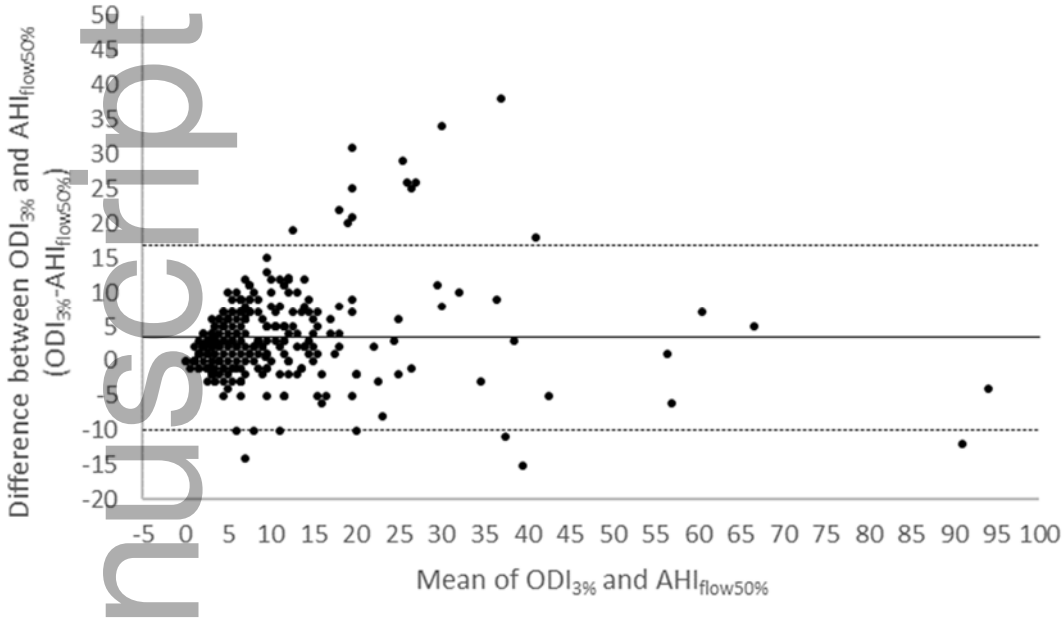
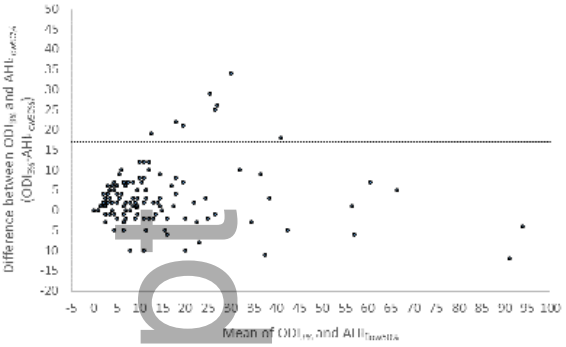
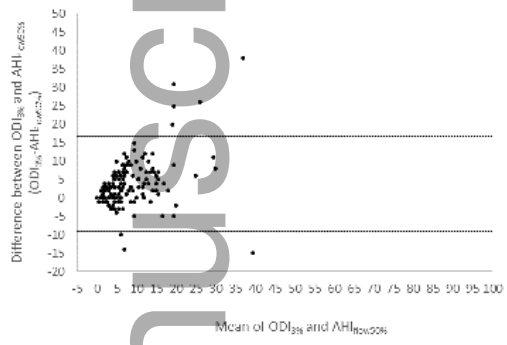


Figure 1. Bland-Altman plot for the distribution of apnoea-hypopnoea index ($AHI_{50\%}$) and oxygen desaturation index ($ODI_{3\%}$) in the total sample.



(a)



(b)

Figure 2. Bland-Altman plot for the distribution of apnoea-hypopnoea index (AHI) and oxygen desaturation index ($ODI_{3\%}$) in (a) males and (b) females.