

Apnea-hypopnea index: time to wake up

Eyal Shaha

Division of Epidemiology and Biostatistics, Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, AZ, USA

Abstract: Despite the widespread use of the apnea-hypopnea index in research, its scientific and statistical properties have not been examined thoroughly. The index may be viewed either as a rate (number of events per hour of sleep) or as a ratio of two variables (number of events/number of hours of sleep). When considered as a rate, the apnea-hypopnea index may be modeled as the dependent variable, provided that researchers explicitly state which physical property they assume to be measuring. On the other hand, the index is rarely, if ever, the preferred model of exposure to sleep-disordered breathing (an independent variable), regardless of whether it is considered a rate or a ratio variable. Continued indiscriminate use of the apnea-hypopnea index in sleep research should be discouraged.

Keywords: apnea-hypopnea index, respiratory disturbance index, ratio variables, rates

Introduction

“We feel that the use of an ‘Apnea Index,’ ie, the number of apneas per sleep-hour, frequently gives a better indication of the seriousness of the disorder.”¹ The term “apnea index” was probably coined in this quote, a parenthetical remark embedded in the second page of a book that was published in 1978.¹ Apart from one misgiving,² the apnea index was endorsed by research groups in the early 1980s,^{3–8} but was later replaced by the apnea-hypopnea index (AHI), the average number of apneas and hypopneas per hour of sleep.^{9–11} Since 1983, the AHI has also been popularized under another self-explanatory name, the respiratory disturbance index.¹¹

Both indices proved helpful in clinical practice, helping to diagnose sleep apnea and make treatment decisions. Clinical measures, however, are not necessarily the preferred measures in biomedical science, where modeled variables should accord with sound research methodology. In retrospect, sleep researchers have adopted the AHI and its predecessor too quickly, skipping a formal appraisal of the two variables. Only recently do we encounter dissenting writers who question the wisdom of using the AHI in research and suggest numerous alternative measures.^{12–16} This commentary offers a methodical analysis of the scientific merit of the AHI, and questions the indiscriminate use of this variable in research.

First question: what are we counting?

Paraphrasing a familiar idiom on apples and oranges, the merit of counting often depends on the answer to the following question: Are we counting apples alone and oranges alone, or are we counting a mix? Whichever definition of a respiratory

Correspondence: Eyal Shaha
Division of Epidemiology and Biostatistics,
Mel and Enid Zuckerman College of
Public Health, University of Arizona,
1295 N Martin Ave, Tucson, AZ, USA
Tel +1 520 626 8025
Fax +1 520 626 2767
Email shaha@email.arizona.edu

event is used, it is inconceivable that all apneas and all hypopneas are the same kind of exposure, as far as effects are concerned. For instance, the clinical consequences of not breathing for 60 seconds cannot be identical to the consequences of shallow breathing for 30 seconds, and the consequences of accompanying hypoxemia surely depend on its magnitude and duration, not on an arbitrary cutoff point. Not surprisingly, the hypopnea definition affects the count of hypopneas,¹⁷ yet no agreed-upon definition will solve the underlying problem: all kinds of apples and all kinds of oranges are counted identically in one overly simplified variable. Of course, no alternative variable is free of weaknesses either, but deficient measures can still be ranked as better and worse.

Setting that issue aside, how well do we count apneas and hypopneas? Judging from the literature, the answer seems variable, although several difficulties should be mentioned. The AHI depends on physiological and technical variables including sleep position, sleep stage, (rapid eye movement [REM] versus non-REM), scoring rules, and device-related parameters, such as sensitivity, calibration, and the type and quality of the signals.^{9,12,18–22} Interestingly, in the era of digital technology, much of the scoring is still done manually, leaving the door open for human error. Yet all of these counting-related issues are the prelude for deeper methodological pitfalls of the AHI. The first insight may be gained by a closer look at the name of the variable.

The AHI: an index or a rate?

Rate is a well-defined term in mathematics, denoting the ratio of the change in one quantity to the change in another. In epidemiology and medicine, the denominator of the rate is usually the passing of time, which can be counted in years (cancer rates), days (case fatality rates), hours (albumin excretion rate), minutes (heart rate), or even seconds (forced expiratory volume in one second). Some rates quantify the change in discrete variables (eg, number of heart beats), whereas others quantify the change in continuous variables (eg, volume of exhaled air). Although there are several kinds of rates, ie, population rate, individual rate, instantaneous rate, hazard rate, and average rate, all of them include some quantity of interest in the numerator and time in the denominator.

The acronym AHI contains the word “index”, but this measure is not some human-made index like the Standard and Poor’s 500; it is actually a rate. More precisely, the AHI may be written as the average of the hourly rates of a respiratory event, as shown next.

Suppose a person sleeps k hours where k is an integer, and N denotes the number of respiratory events during that time. Let n_i denote the number of events in the i -th hour of sleep ($i=1,2, \dots k$), such that $N = \sum n_i$. But the number of events in the i -th hour (n_i) is related to the rate of events in the i -th hour (denoted R_i) as follows: n_i events/1 hour = R_i events/hour. So $n_i = R_i$, and, therefore, $AHI = N/k = (\sum n_i)/k = (\sum R_i)/k$, which is the average of k hourly rates.

There is nothing sacred about dividing time into hours. If sleep duration is divided into q intervals of 10 minutes, rather than k intervals of one hour, another AHI (per 10 minutes) may be written as the average of q event rates in 10-minute intervals, $(\sum R_i)/q$. We cannot, however, shrink the time intervals infinitely and consider an average of time point (ie, instantaneous) rates. The instantaneous rate is a derivative of some quantity (Y) as a function of time (t), but $Y=f(t)$ is not differentiable when Y is the count of events by time t (because it is not a continuous function). The instantaneous rate of a respiratory event does not exist.

Rates: neither causes nor effects

We distinguish between two kinds of variables: natural variables and derived variables. The former describe properties of objects at discrete time points (eg, body mass, height, and waist circumference), whereas the latter are made up by mathematics (eg, body mass index and waist-to-hip ratio). Natural variables make up the causal structure of the natural world; derived variables do not. Natural variables are causes and effects of other natural variables; derived variables are not.²³

Rates are not natural variables, and the average rate of a respiratory event is no exception. The AHI is not caused by anything and does not affect anything; there is no causal parameter behind the association of the AHI with any variable. To think that the AHI itself (the product of mathematics) affects the level of sleepiness is analogous to thinking that body mass index, also a product of mathematics, affects the level of blood glucose. Both are examples of thought bias,^{23–25} which is ubiquitous in science. And as harsh and scary as the last claim might sound, it is solidly true and yet to be challenged.

Are all rates, therefore, useless? Not at all. First, many rates, such as the mortality rate or the rate of lung cancer, do not claim to be more than what they are, ie, the frequency of some event. Second, some rates may also be used to impute the unknown values of natural variables of interest. For instance, airflow (liters per second) serves to measure the speed at which air moves, and the forced expiratory volume

in one second provides information about airway obstruction, which is actually the airway diameter. Likewise, the heart rate tells us not only about the frequency of a heartbeat, but also about the condition, or location, of the governing pacemaker. Derived variables, including rates, may sometimes serve as substitutes for natural variables,²⁵ provided we explicitly state what they are supposed to measure.

The AHI substitutes for ... what?

Scientific inference thrives on explicit, precise assumptions, which may be false, but should not include the apologetic phrase “do not really know”. Scientists are allowed, perhaps even encouraged, to remain uncertain about whether A is a cause of B, but they may not declare uncertainty about what A and B stand for.²⁶ A measurement in science must be coupled with an explicit theory of what is assumed to be measured, because a measured variable is merely a substitution for some variable of interest.^{23,25} Yet all too often a measurement takes on a life of its own, absurdly becoming the focus of the inquiry. In retrospect, that was the case with the AHI. For the most part, researchers did not bother to state what exactly is being measured by that index. What does the average rate of a respiratory event substitute for?

If probed about the causal variable of interest, many researchers would probably answer “severity of sleep-disordered breathing” or some related idea, such as “the level of physiological stress” or “the tendency of the upper airway to collapse” (in obstructive events). “Severity”, “stress”, and “tendency”, however, are not well specified natural variables. They are abstract nouns. If anything, the AHI must substitute for some physical property: upper airway pressure, upper airway diameter (in the obstructive type), excitation of respiratory muscles (in the central type), and the volume of inhaled air to name a few examples. That a single derived variable can substitute for more than one natural variable does not dismiss the duty of researchers to state which natural variable they claim to be measuring. Moreover, inexplicit measurements open the door to ridiculous inference, as illustrated next.

Consider a linear regression model that an analyst might fit:

$$\text{Log (AHI + 0.1)} = \beta_0 + \beta_1 \text{ NECK} + \beta \cdot \mathbf{V} + e$$

where NECK is neck circumference, \mathbf{V} denotes a vector of all so-called covariates, and β denotes a vector of their coefficients.

Without an explicit theory of what is being measured by the AHI and NECK, a reader may claim that both variables substitute for upper airway pressure at a given moment. If so, β_1 in this model is an absurdity, ie, an attempt to estimate the effect of a variable on itself. On the other hand, if the researchers state that neck circumference substitutes for the volume of fat around the upper airway and the AHI substitutes for subsequent upper airway diameter, they may claim that β_1 estimates the effect of the former on the latter. As that example shows, a regression model can correspond to two theories, one logical and another absurd, but a model may not correspond to “We are not sure what we have measured”. To model a variable without an explicit theory of what is assumed to be measured is a mathematical exercise, not science,^{26,27} because no one can corroborate or negate an unspecified causal theory.

The AHI: a good measure of exposure to sleep apnea?

If we are interested in the causes of sleep apnea, the AHI may be the dependent variable, substituting for a specified physical property as illustrated in the last example. On the other hand, the AHI is rarely, if ever, the preferred choice whenever we study the possible consequences of sleep apnea, such as sleepiness, stroke, blood pressure, and death. Several lines of reasoning argue against modeling the AHI as an independent variable.

First and foremost, the hallmark of sleep apnea, whether obstructive, central, or mixed, is reduced volume of inhaled air (which might lead to hypoxemia, sympathetic surges, and microarousals). That type of exposure calls for modeling assumptions, and it is unclear why the average frequency of a respiratory event is a better model than the number (and type) of respiratory events, the total duration of respiratory events,¹³ or other proposed measures.^{14–16} By analogy, which variable better captures exposure to cigarette smoke while awake: the rate of cigarette smoking or the total number of cigarettes smoked? Why do researchers prefer to model, for instance, pack-years of smoking rather than a smoking disturbance index, ie, the average number of cigarettes smoked per hour of wakefulness? Any logic behind a preference for the respiratory disturbance index in sleep research should have carried to the smoking disturbance index in smoking research.

Second, modeling the AHI leads to excessive information bias that can be avoided by using alternative variables. Consider two people who share the same value of the AHI, say 10, one of whom usually sleeps 7 hours each night

whereas the other usually sleeps 5 hours. The AHI classifies them as “equally exposed”, which is obviously wrong when we study the effects of repeated exposure to sleep apnea. In contrast, the total number of respiratory events classifies these people differently (70 versus 50), possibly ranking them, for example, according to the volume of inhaled air during sleep. If researchers think that the interval between successive events might play the role of an effect modifier,²⁸ they may fit an appropriate model; for example, include a product term that contains “event-free sleep duration”. As already discussed, distinguishing between apneas and hypopneas and taking event duration into account should be considered as well.

Third, sleep duration is not only the denominator for the rate of a respiratory event, but is also a variable that substitutes for whatever sleep is supposed to achieve. Since sleep and sleep apnea share some effects of interest (eg, daytime sleepiness), the logic of putting sleep duration in the denominator of the AHI is far from clear. If the reasoning appeals to confounding, that is plainly wrong. One theoretical exception aside,²⁹ confounding bias is not removed by dividing the exposure variable (the number of respiratory events) by the confounder (sleep duration). Rather, sleep duration should enter the model as a covariate. Furthermore, dividing one variable by another is often a bad idea in research.

The AHI: a ratio of two variables

As we just saw, the denominator of the AHI plays a dual role. On the one hand, it is the usual time interval for rate computation, but on the other hand it is a variable that captures sleep duration per se. Given the latter role, we may view the AHI not only as a rate that substitutes for some physical property, but also as a ratio of two variables, ie, the number of respiratory events (N) divided by sleep duration (S). Informally, N may be some measure of sleep apnea during the night and S may be some measure of “battery charge” upon awakening. Is there any justification for modeling the ratio of two variables as an independent variable? May we fit models such as the following?

$$\text{Systolic blood pressure} = \beta_0 + \beta_1 N/S + \beta \cdot V + e$$

$$\text{Log (death rate)} = \beta_0 + \beta_1 N/S + \beta \cdot V$$

To answer this question, we should first ask ourselves what we have in mind when we model N/S in these equations. If we claim that N/S substitutes for a natural variable, and name it, we are partially safe. The ratio may be considered an

imputation of the unknown values of that variable, analogous to any measurement.²³ Of course, the substitution may still be criticized for having information bias, as argued in the previous section.

In all other circumstances, modeling a ratio such as N/S is unjustified because the model entails excessive bias or misspecification of causal relations. If only N (or whatever N is measuring) affects the outcome of interest, division by S adds information bias by creating a variable (N/S), the distribution of which differs from that of N. If only S (or whatever S is measuring) affects the outcome, the error might be compounded by unexplained modeling of 1/S, rather than S, on top of detrimental multiplication by N. And if both N and S affect the outcome, neither the effect of N nor the effect of S is estimated by the coefficient of their ratio (misspecification). At best, we may view the ratio N/S as an interaction term. For instance:

$$\text{Log (death rate)} = \beta_0 + \beta_1 N(1/S) + \beta \cdot V$$

But that does not solve the problem either. Interaction models, serving to estimate effect modification between two causes of some outcome,²⁸ can be interpreted properly only when the components of the product term are included as well.³⁰ If researchers hold a theory of effect modification between sleep apnea and sleep duration, they should fit the following model:

$$\text{Log (death rate)} = \beta_0 + \beta_1 N + \beta_2 S + \beta_3 NS + \beta \cdot V$$

Lastly, even under the paradigm of predicting the outcome, ignoring cause-and-effect relationships, modeling the ratio alone might be worse than modeling one of its components, or worse than modeling the ratio together with its components.³⁰

Conclusion

Although there is some justification for modeling the AHI as the dependent variable, modeling of the AHI as an independent variable is rarely justified. The number and type of respiratory events overnight, their total duration,¹³ and other recently proposed variables¹⁴⁻¹⁶ are better models of exposure to sleep apnea. Interestingly enough, shortly before coining the term “apnea index”,¹ the same first author preferred to report the total number, or duration, of nocturnal apneas,^{31,32} rather than their number per sleep-hour. What caused him to change his mind is unclear, but the effect of his casual remark lasted for many years.

Should previous studies that used the AHI as an exposure variable be declared useless? Not at all. In retrospect, we may (awkwardly) view the AHI as a measure of the total number of respiratory events, just as we may (awkwardly) view the body mass index as a measure of body fat. In both cases, we can save the model by admitting to have mistakenly added some information bias by dividing the numerator (the preferred measure) by another variable. Moreover, we may argue that the bias is negligible or small, because sleep duration, *S*, does not vary greatly among many people, and *N/S* is highly correlated with *N*. Notwithstanding their saving of past work, post hoc arguments do not justify perpetuating a recognized methodological mistake, regardless of its practical significance.

Many of the arguments presented here extend to other ratio variables that are commonly used in sleep research: the arousal index, percentage of time in desaturation, and percentage of time in each sleep stage. In fact, the modeling of ratios has been harshly criticized in other branches of science as well.^{29,30,33–37} However, the lessons offered are far more general. First, a variable that is useful in clinical practice is not necessarily the preferred variable in biomedical science. Second, we should never accept a new variable into science on the basis of feelings or authority. Third, if the analyzed variable is a derived variable, researchers should explicitly state which natural variable they claim to be measuring. And if they are unwilling to commit to a clear theory,³⁸ they are not in the business of science.²⁶

Acknowledgment

The author thanks Doron Shahar for insightful comments on the draft manuscript.

Disclosure

The author reports no conflicts of interest in this work.

References

- Guilleminault C, van den Hoed J, Mitler MM. Clinical overview of the sleep apnea syndromes. In: Guilleminault C, Dement WC, editors. *Sleep Apnea Syndromes*. New York, NY, USA: Alan R Liss, Inc.; 1978.
- Hudgel DW. "Apnea index": need for improving the description of respiratory variability during sleep. *Am Rev Respir Dis*. 1986;133:708–709.
- Perks WH, Cooper RA, Bradbury S, et al. Sleep apnoea in Scheie's syndrome. *Thorax*. 1980;35:85–91.
- Ancoli-Israel S, Kripke DF, Mason W, Messin S. Sleep apnea and nocturnal myoclonus in a senior population. *Sleep*. 1981;4:349–358.
- Sullivan CE, Issa FG, Berthon-Jones M, Eves L. Reversal of obstructive sleep apnoea by continuous positive airway pressure applied through the nares. *Lancet*. 1981;1:862–865.
- Harman E, Wynne JW, Block AJ, Malloy-Fisher L. Sleep-disordered breathing and oxygen desaturation in obese patients. *Chest*. 1981;79:256–260.
- McEvoy RD, Thornton AT. Treatment of obstructive sleep apnea syndrome with nasal continuous positive airway pressure. *Sleep*. 1984;7:313–325.
- Wittig RM, Romaker A, Zorick FJ, Roehrs TA, Conway WA, Roth T. Night-to-night consistency of apneas during sleep. *Am Rev Respir Dis*. 1984;129:244–246.
- Cartwright RD. Effect of sleep position on sleep apnea severity. *Sleep*. 1984;7:110–114.
- Wilhoit SC, Suratt PM, Evans RJ, Brown ED, Kaiser DL. Comparison of indices used to detect hypoventilation during sleep. *Respiration*. 1985;47:237–242.
- Bliwise DL, Carey E, Dement WC. Nightly variation in sleep-related respiratory disturbance in older adults. *Exp Aging Res*. 1983;9:77–81.
- Cherniack NS. The sleep apnea number game: counting the apnea-hypopnea index. *Respiration*. 2009;77:21–22.
- Muraja-Murro A, Nurkkala J, Tiihonen P, et al. Total duration of apnea and hypopnea events and average desaturation show significant variation in patients with a similar apnea-hypopnea index. *J Med Eng Technol*. 2012;36:393–398.
- Kulkas A, Tiihonen P, Julkunen P, Mervaala E, Töyräs J. Novel parameters indicate significant differences in severity of obstructive sleep apnea with patients having similar apnea-hypopnea index. *Med Biol Eng Comput*. 2013;51:697–708.
- Tam S, Woodson BT, Rotenberg B. Outcome measurements in obstructive sleep apnea: beyond the apnea-hypopnea index. *Laryngoscope*. 2014;124:337–343.
- Balakrishnan K, James KT, Weaver EM. Composite severity indices reflect sleep apnea disease burden more comprehensively than the apnea-hypopnea index. *Otolaryngol Head Neck Surg*. 2013;148:324–330.
- Ruehland WR, Rochford PD, O'Donoghue FJ, Pierce RJ, Singh P, Thornton AT. The new AASM criteria for scoring hypopneas: impact on the apnea hypopnea index. *Sleep*. 2009;32:150–157.
- Punjabi NM, Bandeen-Roche K, Marx JJ, Neubauer DN, Smith PL, Schwartz AR. The association between daytime sleepiness and sleep-disordered breathing in NREM and REM sleep. *Sleep*. 2002;25:307–314.
- Peregrim I, Grešová S, Pallayová M, et al. Does obstructive sleep apnea worsen during REM sleep? *Physiol Res*. 2013;62:569–575.
- Aarab G, Lobbezoo F, Hamburger HL, Naeije M. Variability in the apnea-hypopnea index and its consequences for diagnosis and therapy evaluation. *Respiration*. 2009;77:32–37.
- Thornton AT, Singh P, Ruehland WR, Rochford PD. AASM criteria for scoring respiratory events: interaction between apnea sensor and hypopnea definition. *Sleep*. 2012;35:425–432.
- Crowley KE, Rajaratnam SM, Shea SA, Epstein LJ, Czeisler CA, Lockley SW; Harvard Work Hours, Health and Safety Group. Evaluation of a single-channel nasal pressure device to assess obstructive sleep apnea risk in laboratory and home environments. *J Clin Sleep Med*. 2013;9:109–116.
- Shahar E, Shahar DJ. Causal diagrams, information bias, and thought bias. *Pragmatic and Observational Research*. 2010;1:33–47.
- Shahar E, Shahar DJ. Causal diagrams and change variables. *J Eval Clin Pract*. 2012;18:143–148.
- Shahar E, Shahar DJ. Causal diagrams and three pairs of biases. In: Lunet N, editor. *Epidemiology – Current Perspectives on Research and Practice*. Available from: <http://www.intechopen.com/books/epidemiology-current-perspectives-on-research-and-practice>. Accessed March 16, 2014.
- Shahar E. Shahar responds to "causal diagrams and measurement bias". *Am J Epidemiol*. 2009;170:963–964.
- Shahar E. A method to detect an unknown confounder: something from nothing? *J Eval Clin Pract*. 2012;18:702–703.
- Shahar E, Shahar DJ. On the definition of effect modification. *Epidemiology*. 2010;21:587.
- Allison DB, Paultre F, Goran MI, Poehlman ET, Heymsfield SB. Statistical considerations regarding the use of ratios to adjust data. *Int J Obes*. 1995;19:644–652.

30. Kronmal RA. Spurious correlation and the fallacy of the ratio standard revisited. *J R Stat Soc Ser A*. 1993;156:379–392.
31. Guilleminault C, Eldridge FL, Tilkian A, Simmons FB, Dement WC. Sleep apnea syndrome due to upper airway obstruction: a review of 25 cases. *Arch Intern Med*. 1977;137:296–300.
32. Guilleminault C, Tilkian A, Lehrman K, Forno L, Dement WC. Sleep apnoea syndrome: states of sleep and autonomic dysfunction. *J Neurol Neurosurg Psychiatry*. 1977;40:718–725.
33. Schuessler K. Analysis of ratio variables: opportunities and pitfalls. *Am J Sociol*. 1974;80:379–396.
34. Packard GC, Boardman TJ. The use of percentages and size-specific indices to normalize physiological data for variation in body size: wasted time, wasted effort? *Comp Biochem Physiol A*. 1999;122:37–44.
35. Raubenheimer D. Problems with ratio analysis in nutritional studies. *Funct Ecol*. 1995;9:21–29.
36. Forshee RA, Storey ML. Controversy and statistical issues in the use of nutrient densities in assessing diet quality. *J Nutr*. 2004;134:2733–2737.
37. Tu Y-K, Clerehugh V, Gilthorpe MS. Ratio variables in regression analysis can give rise to spurious results: illustration from two studies in periodontology. *J Dent*. 2004;32:143–151.
38. Hernan MA, Cole SR. Causal diagrams and measurement bias. *Am J Epidemiol*. 2009;170:959–962.

Nature and Science of Sleep

Dovepress

Publish your work in this journal

Nature and Science of Sleep is an international, peer-reviewed, open access journal covering all aspects of sleep science and sleep medicine, including the neurophysiology and functions of sleep, the genetics of sleep, sleep and society, biological rhythms, dreaming, sleep disorders and therapy, and strategies to optimize healthy sleep. The journal welcomes

original research, clinical & epidemiological studies, reviews & evaluations, case reports and extended reports. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/nature-and-science-of-sleep-journal>