Sleep Health 1 (2015) 133–137

Contents lists available at ScienceDirect

Sleep Health

Journal of the National Sleep Foundation

journal homepage: http://www.elsevier.com/locate/sleh

Agreement between simple questions about sleep duration and sleep diaries in a large online survey☆

Christopher B. Miller, PhD [a,b], Christopher J. Gordon, PhD [a,b], Leanne Toubia, BHthSc [a], Delwyn J. Bartlett, PhD [a,c], Ronald R. Grunstein, MD PhD [a,c,d], Angela L. D'Rozario, PhD [a,c], Nathaniel S. Marshall, PhD [a,b,*]

[a] NeuroSleep and Woolcock Institute of Medical Research, University of Sydney, Sydney, NSW, Australia
[b] Sydney Nursing School, University of Sydney, Sydney, NSW, Australia
[c] Central Clinical School, Faculty of Medicine, University of Sydney, Sydney, NSW, Australia
[d] Sleep and Respiratory Failure Service, Royal Prince Alfred Hospital, Sydney, Australia

ARTICLE INFO

Article history:
Received 17 December 2014
Received in revised form 12 February 2015
Accepted 27 February 2015

Keywords:
Agreement
Sleep Survey
Sleep Duration
Epidemiology
Validation
Self-reported Sleep

ABSTRACT

Introduction: Self-reported habitual sleep duration has been used widely in epidemiologic research, yet this measure remains to be validated. We evaluated whether simple sleep duration questions concord with sleep diaries in an online sample.
Methods: Australian adults aged 18+ years completed an internet survey examining measures of sleep, sociodemographic risk factors, and a 7-day sleep diary. We examined single-question (how many hours of sleep would you normally get?) and 2-question assessments (difference between sleep and wake times) to a 7-day sleep diary estimation of sleep duration. Using Bland-Altman plots and associated statistics, we tested systematic differences, precision, and systematic bias. We also evaluated whether the differences were consistent along the entire range of the measurement and whether they were associated with any sociodemographic risk factors (Spearman rho).
Results: Data were analyzed from 1662 participants (67.3% female). Bland-Altman plots displayed visual discrepancies between both 1-question and 2-question reports of sleep duration compared with sleep diaries. Both the single- (−17 minutes) and double-question (8 minutes) sleep duration estimates differed significantly (both $P < .001$). These simple estimates only agreed to within ±2.5-3 hours compared with diary estimates. The measure was also weakly systematically biased (rho = +0.204 and +0.309, $P < .001$) through the measurement range. There were significant differences and associations between differences in sleep duration estimation and determinants of health.
Conclusions: Simple questions estimating habitual sleep duration are imprecise and systematically biased in a large online survey. The amount of difference is correlated with well-known sociodemographic risk factors.

Introduction

Single questions of subjective total sleep time (TST) are frequently used to quantify sleep duration in large-scale epidemiologic studies. Single questions of "how long did you sleep last night?" are quick, easy to administer, and inexpensive.[1] Single questions have been found to be important for research, as a number of studies have reported links between reduced TST and ill health including risk and incidence of obesity,[2] type 2 diabetes,[3] hypertension,[4] cardiovascular disease,[5] and mortality.[6,7] Despite its widespread use, it does not appear that the single-sleep duration question has been validated in a general population sample.

Two studies are widely cited to establish the validity of the single-sleep duration question. The first compared sleep diaries

☆ The ABC National Science Week Big Sleep Survey was supported by the National Health and Medical Research Council of Australia (NHMRC) Centre for Integrated Research and Understanding of Sleep at The University of Sydney providing personnel and expertise (NHMRC grant no. 571421 to Prof Grunstein providing salary support to Drs D'Rozario and Marshall). The Australian Broadcasting Corporation hosts National Science Week and this survey and provided the information technology support and advertising to collect these data. Prof Grunstein is supported by an NHMRC Practitioner Fellowship (1022730) and Dr D'Rozario by an NHMRC Dora Lush PhD Scholarship (no. 633172). The NHMRC had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The ABC did have a role in design data collection but no input into analyses, decision to publish or preparation of the manuscript.
* Corresponding author. Nathaniel Marshall, Sydney Nursing School, The University of Sydney, Australia, NSW 2006. Tel.: +61 293510829.
E-mail address: nathaniel.marshall@sydney.edu.au (N.S. Marshall).

http://dx.doi.org/10.1016/j.sleh.2015.02.007
2352-7218/

to actigraphy in those who were legally blind.[8] The second investigated sleep using polysomnography (PSG), actigraphy, and postwaking assessments of subjective sleep (30 minutes after) in 21 flight crew involved in delivery flights across the Pacific Ocean from the United States to Southeast Asia.[9] Neither of these studies addressed the validity or reliability of the single-question habitual sleep duration in the general population.

Other studies have attempted to validate simple questions against actigraphy in more general population samples. One large cohort study (n = 600) has reported that a single question of sleep duration is only moderately correlated to actigraphy but is biased because the amount people get their sleep duration wrong by is correlated with important sociodemographic drivers of health.[10] Further studies have also shown poor agreement between actigraphy and self-reported assessments of sleep.[11,12] Increasingly, health researchers are looking to use the internet for sampling, as this is low cost per participant compared with traditional study methodologies. One of the limitations with online study is that, logistically, it is difficult to use actigraphy in large numbers of participants (10,000 +). Subjective measures of sleep duration may therefore be preferable to objective measures in epidemiologic research, as both actigraphy and PSG are difficult and expensive to implement in large-scale research studies.[1]

Sleep diaries are the criterion-standard subjective assessment of sleep, important for the assessment of sleep and for assessing sleep disorders, like insomnia.[13,14] Diaries, however, are time-consuming and burdensome to both participants (time to fill in) and researchers (data processing) and may not have any marginal utility if 1 or 2 questions of habitual sleep duration prove to be relatively valid and imprecise measurements. An understudied alternative might be to estimate sleep duration by asking only 2 questions about wake and sleep times. Therefore, we compared simple questions of sleep duration with sleep diary measurements in a relatively large online survey: the Australian Broadcasting Corporation (ABC), "Big Sleep Survey."

Methods

The "Big Sleep Survey" captured Australian sleep habits through the ABC (a government-funded media organization akin to the British Broadcasting Corporation (BBC) in the United Kingdom or the Public Broadcasting Service (PBS) in the United States) in conjunction with the NHMRC Centre for Integrated Research and Understanding of Sleep at the University of Sydney. The principal question in this study concerned the use of technology in teenagers' sleep, which has been reported (see[15]). The survey was conducted between August 2010 and was open until January 2011 as part of Australian National Science Week 2010. All participants were recruited through national media broadcasting (television, internet, and radio) and asked to complete an online questionnaire and keep a paper-based (print out) sleep diary for 7 days. Questionnaire completion coincided with early spring until midsummer across the continent of Australia (longitude, 113°E and 153°E; latitude, 11°S and 38°S). Participants consented electronically before the beginning of the study and were not paid for questionnaire completion, but those who did were enrolled into a draw to win an electronic tablet device. After the survey, participants were provided with feedback about their sleep habits and directed to seek treatment if symptoms consistent with sleep disorders were revealed. Ethical approval was provided by the University of Sydney Human Research Ethics Committee (protocol no. 12590) to analyze data obtained after 12th August 2010, and only these data are presented.

Survey

The survey was completed by Australians of any age and consisted of a broad range of questions relating to demographics, measures of sleep onset, offset, on both work and nonworkdays, and circumstances that might have influenced sleep quality and quantity (alcohol, caffeine, exercise, smoking, and pre-bed sleep habits). Participants were also asked about their medical history, medication use, shift work status, and diagnosis of sleep disorders. Embedded within the survey were the Epworth sleepiness scale (ESS)[16] and insomnia severity index (ISI).[17] All respondents completed the first 3 questions of the ISI, and those that responded with "a few nights per week" or "every or almost every night" to the first 3 questions from the index were prompted to complete the full questionnaire.

Self-reported sleep duration measures

Self-reported habitual sleep duration was assessed in this study using 3 methods: (1) a single question for sleep duration from the online survey; (2) through the use of 2 questions from the survey; and (3) from sleep diaries using 2 questions. For the single-sleep duration assessment, participants were asked to answer "on average, how many hours of sleep would you normally get (excluding naps)?" This was measured in hours (1 numeric input box) and minutes (a second input box). For the 2-question sleep duration survey assessment, participants were asked to nominate "specifically on the working day, what time would you (1) wake up and (2) go to sleep?" Participants were also asked to specify this on the nonworking day. This response was measured in hours and 5-minute epochs (2 drop down boxes). The difference between these 2 time points was calculated to estimate TST and weighted on the assumption that the most common ratio of working to nonworking days was 5:2. For the 2-question 7-day sleep diary assessment, every day for 1 week, participants were asked "what time did you attempt to sleep?" and "what was your final wake-up time?" and to specify if it was a working or nonworking day. The difference between these 2 time points was used to estimate TST, and these daily estimations of sleep duration were then averaged over all of the days reported. Participants first completed the survey and then were asked to upload their sleep diary data to the survey Web site once complete, and responses were recorded to the nearest minute. To our knowledge, this might be the largest online sample of sleep diary data from a general population sample collected.

Analysis of data

All data were analyzed in SPSS (version 21). Participants aged older than 18 years were included in these analyses. Participants with TST less than 3 hours or greater than 13 hours were excluded, as they were deemed biologically implausible sleep durations using the same rule reported previously.[18] Bland-Altman plots[19] were constructed to visually display the agreement between simple measures and sleep diary measures of sleep. The 2 measures were then assessed and compared with sleep diaries for the following measures of agreement as specified previously[19]: (1) the estimated bias, which is the systematic difference in sleep duration between survey and diary (assessed by a *t* test and visually by looking at the average error line compared with the zero line on the Bland-Altman plot); (2) the 95% limits of agreement (± 1.96 SDs) is used to determine precision of the measurement; and (3) the systematic bias is where the difference in sleep duration between the 2 methods changes depending on how far through the range of measurement you are (tested with Spearman rho correlations to see, for instance, whether long sleepers systematically overestimate their sleep and short sleepers systematically underestimate theirs.). Within each individual, we also calculated the difference between the methods and correlated this difference against widely recognized sociodemographic risk factors of