



Special Article

Parallel Assessment Challenges in Nutritional and Sleep Epidemiology

Galit Levi Dunietz*, Erica C. Jansen, Shelley Hershner, Louise M. O'Brien, Karen E. Peterson, and Ana Baylin

* Correspondence to Dr. Galit Levi Dunietz, Division of Sleep Medicine, 1500 E. Medical Center Drive, C728B Med Inn Building, University of Michigan, Ann Arbor, MI 48109-5845 (e-mail: gldt@med.umich.edu).

Initially submitted September 25, 2019; accepted for publication October 15, 2020.

Sleep has been consistently linked to health outcomes in clinical studies, but only in recent years has sleep become a focus in epidemiologic studies and public health. In particular, the sizable prevalence of insufficient sleep in the population warrants well-designed epidemiologic studies to examine its impact on public health. As a developing field, sleep epidemiology encounters methodological challenges similar to those faced by nutritional epidemiology research. In this article, we describe a few central challenges related to assessment of sleep duration in population-based studies in comparison with measurement challenges in nutritional epidemiology. In addition, we highlight 3 strategies applied in nutritional epidemiology to address measurement challenges and suggest ways these strategies could be implemented in large-scale sleep investigations.

calibration; causal diagrams; directed acyclic graph; measurement error; method of triads; nutritional epidemiology; sleep epidemiology

Abbreviations: OSA, obstructive sleep apnea; PSG, polysomnography.

Despite consistent reports about the negative impact of insufficient sleep on health (1), more than a third of American adults are sleep deprived (2), representing a major public health burden. Indeed, insufficient sleep has been linked to deleterious health outcomes such as cardiometabolic morbidity (3–7), depression (8), and mortality risk (9) in addition to motor vehicle accidents (10). To maintain optimal health, the American Academy of Sleep Medicine recommends that adults obtain at least 7 hours of sleep per night (11).

Sleep duration is not the only important construct as over the last 70 years, the field of sleep medicine has made significant advances with the nosology of more than 80 sleep disorders and the development of technology for diagnosis and treatment of sleep disorders (12). More recently, the rapid growth of sleep epidemiology has garnered attention from wide audiences, even beyond the scientific community (13). Yet population-based sleep research has methodological challenges surrounding the assessment of sleep. These measurement challenges, inherent in large studies, limit interpretation of prior findings and comparison across studies.

Interestingly, the field of nutritional epidemiology has faced similar methodological obstacles in the assessment of diet. Beyond their mutual measurement challenges, sleep

and nutrition are entwined via complex biological, behavioral, and environmental pathways (14). The association between sleep and nutrition is bidirectional; sleep duration and quality affect macronutrient and caloric consumption (15, 16), while intake of particular foods (e.g., tart cherries, kiwi fruit, and fatty fish) (16, 17) and dietary patterns (18) might affect sleep duration, quality, and timing. Sleep and nutrition are influenced by circadian rhythms, socioeconomic status, and physical activity, and both, independently and jointly, contribute to cardiometabolic morbidity. Specifically, sleep duration, variability, and timing as well as dietary composition, quantity, and mealtimes influence adiposity, type 2 diabetes, and cardiovascular disease (19–21).

Population-based nutritional studies date back to the 1940s and have had an impact on dietary recommendations worldwide (22). As nutritional epidemiology has expanded, methods and tools have been developed and refined to minimize the inherent measurement error in diet assessment. We believe that the conceptual framework that guided the evolution of these methods and tools in nutritional epidemiology could inform the growing field of sleep epidemiology, which includes both sleep clinicians and investigators developing population-based sleep research.

This article illustrates 3 parallel methodological challenges in sleep and nutritional epidemiology to guide epidemiologists and clinicians with sleep-research pursuits in large-scale samples. Sleep assessment and analytical approaches are common limitations in population-based sleep research. In particular, implementation of causal frameworks in large-scale sleep research and addressing systematic errors in sleep assessment could improve the accuracy of effect estimates. Here, we provide a practical perspective to address these challenges by implementing methods to minimize bias that are commonly used in nutritional epidemiology. Measurement errors are inevitable in all epidemiologic studies, but careful consideration of assessment instruments and causal pathways during the planning stages could alleviate and perhaps avoid these pitfalls. The aim of this paper is to increase awareness among investigators with an interest in epidemiologic sleep research about techniques widely used in the nutritional epidemiology field that could be implemented primarily in the assessment of sleep duration, either in new studies or with preexisting data.

PARALLEL METHODOLOGICAL CHALLENGES IN SLEEP AND NUTRITIONAL EPIDEMIOLOGY

Application of the causal framework in sleep and nutrition research

Directed acyclic graphs, or causal diagrams, serve as a visual presentation of a specified causal pathway. Causal diagrams aid in the formalization of research questions and identify confounding and mediating pathways between the exposure, outcome, and third variables within a causal framework.

Sleep is a multidimensional state represented by its duration, architecture (sleep stages), timing, continuity, quality, and intraperson night-to-night variability; thus, determining which aspects of sleep to measure in a particular study can be a challenge. Depending on the research question, the different dimensions of sleep might confound, interact with, or mediate one another in relation to a health outcome (7). Here, we use causal diagrams to illustrate the role of sleep duration as a confounder or a mediator in 2 separate research questions (23, 24). In older adults with poor sleep quality, long sleep duration is associated with cardiovascular mortality, suggesting that sleep quality modifies the relationship between sleep duration and cardiovascular mortality (25). Alternatively, sleep duration might be confounded by concurrent sleep disorders such as insomnia, obstructive sleep apnea (OSA), or restless leg syndrome. Indeed, a population-based cohort reported that untreated OSA was associated with high mortality risk, independent of sleep duration (Figure 1A) (24). Sleep architecture, specifically, the percentage time in slow wave sleep, has been inversely associated with body mass index in older men, independent of OSA and total sleep duration (26). Finally, sleep duration might be an intermediate variable on a pathway between an exposure and an outcome. For example, in a large study among employed workers, sleep duration mediated the asso-

ciation between working long hours per week and occupational injury risk (Figure 1B) (23).

There is a parallel in nutritional epidemiology. Diet is similarly a multifaceted behavior that represents quantity, composition, and timing of nutrient intake. In relation to health outcomes, diet can be a predictor, a source of confounding, a mediator, or an effect modifier. For example, total energy intake often confounds the association of red meat intake and cardiovascular morbidity. Thus, nutritional epidemiologists routinely measure and account for total energy intake as a potential confounder with regression techniques (i.e., the residual method or by directly including total energy intake in model adjustment) (27). However, in certain scenarios, total energy intake acts as a mediator on the causal pathway between an exposure and an outcome and thus should not be handled as a confounder due to the risk of overadjustment (28).

To illustrate, associations of sugar-sweetened beverage intake and weight gain might be mediated by total energy intake, in that consumption of sugary beverages can lead to excessive caloric intake that, in turn, can cause weight gain. Further, nutrients might interact with one another in relation to health outcomes. To capture these interactions, nutritional epidemiologists analyze diet patterns, rather than specific foods, to associate food combinations in the diet with disease. These examples highlight the essential role of causal diagrams in diet research as a conceptual framework that guides statistical analysis and collection strategies. A contemporary epidemiologic tool, causal diagrams could also aid investigators in the formulation of research questions, inform collection of sleep measures and other data, guide the statistical analysis, and control for bias. Specifically, causal diagrams allow investigators to assess the role of a sleep variable in the primary study question: a potential confounder (e.g., OSA is an independent predictor of sleep duration and hypertension), a mediator (e.g., sleep quality is on the pathway from shift work to weight gain), or an extraneous source of variation. If sleep confounds the investigated association, adjustment for sleep in the analysis would eliminate the resulting bias. In contrast, if the sleep variable mediates this association, statistical adjustment for this variable will induce overadjustment bias (28). Causal diagrams have been underutilized in epidemiologic sleep research. A recent systematic review (7) has examined the role of diet in the relationship between sleep and cardiometabolic health. This review found numerous studies that considered diet as a confounder rather than a mediator, an analytical strategy that resulted in overadjustment of reported effect estimates.

Validation of sleep instruments

In epidemiologic studies, self-reported instruments are often used because they are quick, affordable, and feasible to administer in large samples. Sleep questionnaires are designed to measure various sleep characteristics, including duration, timing, and sleep quality (e.g., Pittsburgh Sleep Quality Index) (29), or to screen for specific sleep disorders, such as OSA and insomnia. While objective instruments

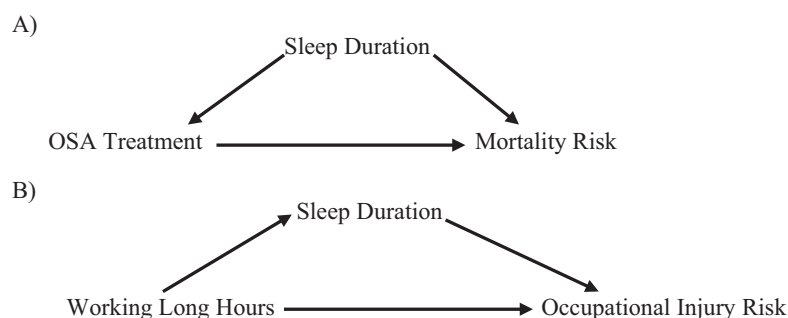


Figure 1. Causal diagrams showing different pathways involving sleep duration. A) Causal diagram representing potential confounding by sleep duration of the association between treatment of obstructive sleep apnea and mortality risk. B) Causal diagram representing potential mediation by sleep duration of the association between working hours per week and occupational injury risk.

are often used for diagnosis of many sleep disorders, self-reported tools are essential for evaluation of some sleep characteristics (e.g., sleep quality) that cannot be measured otherwise. Further, subjective tools have established associations between sleep and chronic disease (30–32), and despite being prone to systematic and random error, these tools have high sensitivity and moderate to high specificity (33).

Objective assessment of total sleep duration is accomplished during polysomnography (PSG) via electroencephalography that measures sleep onset, awakenings, total sleep time, and sleep stages. While PSG is considered the gold standard for measurement of sleep architecture, its utilization in research settings is often not financially or logistically feasible. Moreover, a night in a sleep laboratory might not be typical of a person's usual sleep duration, given that bedtimes and wake times are often related to laboratory practices rather than personal routines. In response to these limitations, wearable devices have been designed and optimized to approximate total sleep duration in clinical and research settings. These devices use accelerometry, in addition to temperature and heart rate in some devices, to quantify sleep duration and sleep timing by recording movement as a proxy for sleep according to the individual device's algorithm. Despite their objective assessment of sleep and wake patterns, these instruments might produce both systematic and random measurement errors in different subpopulations (34, 35). For example, in individuals with insomnia, the use of wearable devices to measure sleep can introduce systematic bias if prolonged inactive wake periods during the night are recorded as sleep time. The resulting overestimation of sleep time could produce systematic misclassification of sleep duration. Further, even in the absence of sleep disorders, occasional inactivity could induce random error when sleep duration is recorded through objective devices.

While wearable devices and questionnaires are useful instruments in sleep research, investigators must interpret them with caution and use validated instruments when available. Yet even "standard" sleep questionnaires are not necessarily validated in all populations, likely due to the dearth of studies that concurrently collected data from questionnaires and gold standard–assessed sleep.

Comparisons of subjective and objective tools to assess sleep duration have demonstrated a moderate correlation between self-reported sleep duration questions (habitual bedtime, wake time on weekdays and weekends), sleep diary, and 7-night actigraphy (36–38). However, these correlations vary by sex, age, and the reported day of the week (36, 38). Associations between sleep duration and self-rated health have been examined across 3 sleep assessment tools (39). For example, a U-shaped relationship has been observed between a single self-reported sleep duration question and self-rated health. Yet the U-shaped association between self-rated health and 3-night sleep logs was significantly attenuated, and it disappeared with actigraphy-based sleep duration (39).

The field of nutritional epidemiology has encountered similar challenges in that dietary questionnaires are not always validated for the specific population under study, and gold standard measurements of diet are often not available. Two of the solutions utilized to address these problems—method of triads and calibration—warrant attention for their potential application to the assessment of sleep duration.

In nutritional epidemiology, the logical parallel to PSG (in which a person's sleep is directly observed and quantified) is researcher-observed food intake (with weighted portions and known nutrient composition) in a laboratory setting. Yet this assessment strategy is infeasible except in small studies. More commonly, self-reported food intake through questionnaires (food frequency questionnaires, 24-hour recall, or food diaries) is used as a diet measure in nutrition research. Food frequency questionnaires face similar issues as with sleep questionnaires: Both sleep and eating are daily occurrences, but people might find it difficult to recall these behaviors with accuracy and precision. Due to these systematic and random measurement errors, which can vary by population, each study population might need its own validated and culturally appropriate questionnaire. To validate self-reported nutrients without the use of a true gold standard (i.e., direct observation of food intake of known nutrient composition), nutritional epidemiologists might use the method of triads.

The method of triads is used primarily in nutritional epidemiology but has been applied to other fields as well

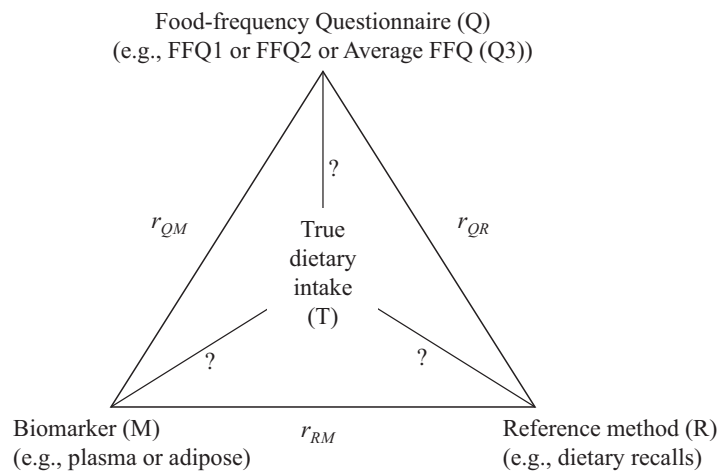


Figure 2. Representation of the method of triads in nutritional epidemiology. Application of the method of triads to evaluate the performance of food frequency questionnaires and biomarkers as indicators of long-term dietary intake. This figure is reprinted with permission from Kabagambe et al. (42).

(e.g., validation of self-reported pubertal stages) (40). In nutritional epidemiology, the method of triads assesses the validity of a self-reported nutrient measurement by conducting 3-way comparisons between nutrient data derived from independent assessments (e.g., biomarker, 24-hour recall, and food frequency questionnaire) (41). The primary utility is that measurement errors between self-reported dietary assessments are often correlated (e.g., participants might misremember their intake of a particular food item in the same way on a 24-hour recall as in a food frequency questionnaire), which can lead to a biased estimation of the correlation coefficients of a given nutrient. In contrast, triangulation analyses evaluate the correlations between nutrients assessed with 3 methods simultaneously, and through a series of equations, the correlation coefficients between each measure and the underlying “true value” of a nutrient can be estimated (Figure 2—figure and equations adapted with permission from Kabagambe et al. (42)).

In sleep research, the method of triads could be similarly used to assess the validity of sleep duration when PSG measurements (the gold standard) are not available; the 3 independent assessment methods could be questionnaires on typical sleep habits, prospective sleep diaries, and sleep assessed by wearable devices (Figure 3). Although many studies evaluate the correlation between 2 independent measures of sleep duration, using 3 independent measures simultaneously will provide more accurate validity coefficients. If data from all 3 methods are not available on all subjects, it is acceptable to perform the triangulation analysis within a randomly selected subset of the population. Of note, utilizing the method of triads might actually be more feasible in epidemiologic sleep studies than in dietary studies, given that self-reported sleep questions and diaries take considerably less time than dietary recalls or prospective dietary records. Further, there might be some epidemiologic sleep studies that have already collected all 3 types of data and could

easily utilize the method of triads (Figure 4). This method might be best suited for assessing the validity of typical sleep duration rather than other aspects of sleep.

Calibration method for correction of systematic errors in sleep duration assessment

Systematic errors (consistent, repeatable errors as opposed to random errors) have often been detected in the assessment of sleep when comparing questionnaire-derived sleep estimates with objective measurements (i.e., PSG or actigraphy).

In comparison with objective amount of sleep, self-reported sleep duration has been over-estimated in adult (43), pediatric (44), pregnant (45), and older (46) populations. In addition, differential misreporting of subjective sleep duration according to the actual amount of sleep obtained has been shown. Short sleepers (≤ 6 hours per night) tended to underestimate sleep duration, and those who slept > 6 hours per night overestimated sleep duration in relation to actigraphy data (36, 47). In contrast to medical-grade devices, such as actigraphy, consumer wearables that estimate heart rate in their sleep assessments have not been well validated against PSG and thus have not been an optimal choice for use in sleep research (48, 49). These wearable devices might have proprietary algorithms that hamper validation efforts (50). However, recent development of mathematical models that incorporate an estimate of circadian phase and utilize raw acceleration and heart rate data from the Apple Watch has shown great promise for the identification of sleep parameters across different device platforms (51). These findings open up the possibility of using consumer-wearable devices to measure sleep on a large scale.

Concurrent sleep disorders, diagnosed or undiagnosed, might also affect self-reported sleep duration in a systematic

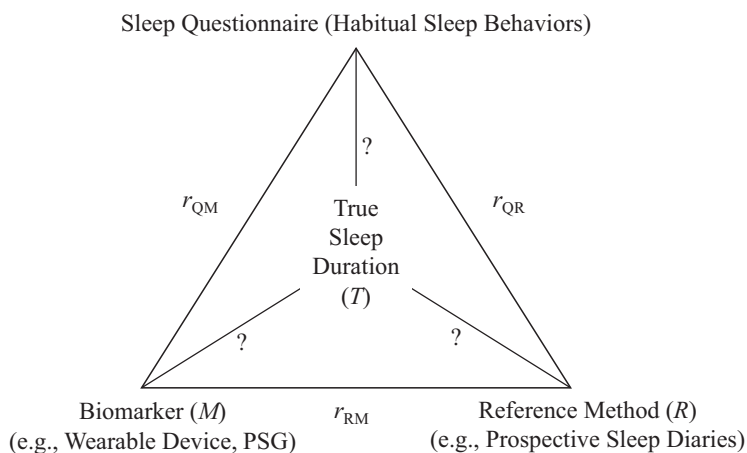


Figure 3. Method of triads adapted for sleep research from Kabagambe et al. (42). Application of the method of triads to evaluate the performance of food frequency questionnaires and biomarkers as indicators of long-term dietary intake. PSG, polysomnography.

way (52). Individuals with OSA or insomnia might underestimate sleep duration; in one study, those with insomnia underestimated total sleep time by 81 minutes (53–55). Finally, an overestimate of sleep duration is expected in responses of employees of industries that regulate sleep, such as the transportation industry (56). While these systematic biases are well-described, we believe that few studies have attempted to correct the bias within the collected data, as has been described in the nutritional epidemiology field (57).

Nutritional epidemiology also uses self-reported instruments, and as in sleep research, systematic errors are often present in self-reported dietary instruments. The source of error could be the instrument itself, such as when a food frequency questionnaire neglects to ask about consumption of an important food, which could happen when short food frequency questionnaires are administered or when a population experiences rapid shifts in the food supply. Systematic error can also be the result of social desirability biases. For example, self-reported total energy intake is often underestimated among overweight study participants (58–60). Overestimation of fruit and vegetable intake and underestimation of fat intake have also been shown to differ according to sex, type of dietary instrument used, and

dietary-intervention treatment status. While a systematic error with the research instrument can be difficult to correct after data collection, predictable under- or over-reporting by the participants could be addressed by calibration studies (61). In these studies, the following steps are undertaken:

1. Using a subset of the data, measurements of the exposure of interest are taken subjectively (self-report) and objectively (biomarker).
2. A linear regression model is fitted, with the objective marker as the continuous dependent variable and the self-report measure as the continuous independent variable.
3. The β estimate from the linear regression model is applied as a “correction factor” to the self-report measures (i.e., each unit of self-reported nutrient is related to x units of the objectively-assessed nutrient).

Once corrected, the new measurements can be used in data analysis as a comparison with the results obtained without correction.

To identify and correct systematic bias in sleep research, small validation studies with objective measures of sleep could be conducted in conjunction with larger cohorts with self-reported measures. For example, a calibration study could proceed as follows:

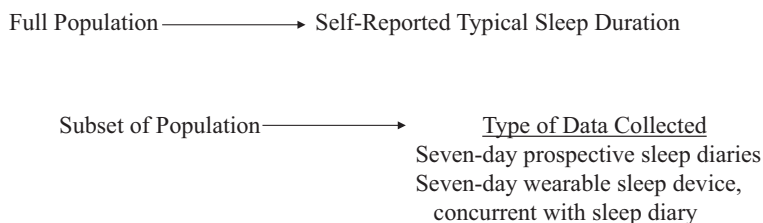


Figure 4. Sample checklist of data for method of triads for sleep research.

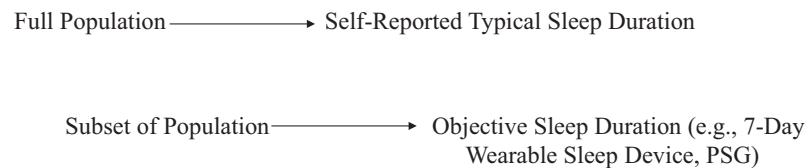


Figure 5. Sample checklist of data for calibration method for sleep research. PSG, polysomnography.

1. Collect self-reported sleep duration for 500 children. For 50 of them, conduct actigraphy at the same time.
2. Perform linear regression with objective sleep duration as the outcome and subjective sleep duration as the exposure.
3. Multiply self-reported sleep duration by the β value from linear regression model to compute corrected sleep duration.

Calibration studies might be an effective strategy to correct bias in already-existing large data sets, if validation studies have been conducted within subsets of the data set or in generalizable populations (Figure 5). Indeed, many validation studies of self-reported sleep duration or timing exist within particular populations (62–64), meaning their estimates could be used to calibrate data sets in similar populations.

FINAL REMARKS

As the field of sleep epidemiology moves forward, we have reviewed and reflected on several methodological challenges it faces in comparison with similar obstacles in nutritional epidemiology. In epidemiologic research, exposures or outcomes are ideally assessed with objective tools, but if such tools are unavailable, self-reported instruments are used. While subjective sleep tools have uncovered the role of poor sleep in morbidity and mortality, they are associated with measurement error. Beyond assessment challenges, analytical errors can stem from an unclear or absent conceptual framework. To minimize these measurement errors and their bias, we propose that investigators pursuing epidemiologic sleep studies borrow methodological approaches used in nutritional epidemiology. Here, we have illustrated the use of causal diagrams, method of triads, and calibration techniques. These methods present a partial view of current research challenges but could inspire future interdisciplinary collaborations among investigators in these fields. For example, self-reported sleep instruments have limited ability to assess sleep physiology, a determinant of healthy aging. This limitation has sparked the development of consumer wearable devices and smartphone applications to mimic the objective sleep assessment of medical devices that are used in sleep laboratories, but methods to analyze these data are constantly evolving. The nutrition field is also working on methods to analyze real-time dietary data from apps.

Thus, as both disciplines advance, assessment and analytical methods developed by sleep researchers could inform

nutrition investigators and strengthen collaborative efforts across these disciplines.

ACKNOWLEDGMENTS

Author affiliations: Division of Sleep Medicine, Department of Neurology, University of Michigan, Ann Arbor, Michigan, United States (Galit Levi Dunietz, Shelley Hershner, Louise M. O'Brien); Department of Nutritional Sciences, University of Michigan School of Public Health, Ann Arbor, Michigan, United States (Erica C. Jansen, Karen E. Peterson, Ana Baylin); Department of Oral and Maxillofacial Surgery, University of Michigan, Ann Arbor, Michigan, United States (Louise M. O'Brien); and Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, Michigan, United States (Ana Baylin).

This work was funded by the National Institute of Neurological Disorders and Stroke (award T32 NS007222), by an F32 National Research Service Award from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (award F32 HD091938), and by a Mentored Research Scientist Development Award from the National Heart, Lung, and Blood Institute (K01 HL144914).

Conflict of interest: none declared.

REFERENCES

1. Steptoe A, Peacey V, Wardle J. Sleep duration and health in young adults. *Arch Intern Med.* 2006;166(16):1689–1692.
2. Liu Y, Wheaton AG, Chapman DP, et al. Prevalence of healthy sleep duration among adults—United States, 2014. *MMWR Morb Mortal Wkly Rep.* 2016;65(6):137–141.
3. Nagai M, Hoshida S, Kario K. Sleep duration as a risk factor for cardiovascular disease—a review of the recent literature. *Curr Cardiol Rev.* 2010;6(1):54–61.
4. Palagini L, Maria Bruno R, Gemignani A, et al. Sleep loss and hypertension: a systematic review. *Curr Pharm Des.* 2013;19(13):2409–2419.
5. Cappuccio FP, Miller MA. Sleep and cardio-metabolic disease. *Curr Cardiol Rep.* 2017;19(11):110.
6. Grandner MA, Chakravorty S, Perlis ML, et al. Habitual sleep duration associated with self-reported and objectively determined cardiometabolic risk factors. *Sleep Med.* 2014; 15(1):42–50.

7. Jansen EC, Dunietz GL, Tsimpanouli ME, et al. Sleep, diet, and cardiometabolic health investigations: a systematic review of analytic strategies. *Current Nutr Rep*. 2018;7(4): 235–258.
8. Perlman CA, Johnson SL, Mellman TA. The prospective impact of sleep duration on depression and mania. *Bipolar Disord*. 2006;8(3):271–274.
9. Cappuccio FP, D'Elia L, Strazzullo P, et al. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. *Sleep*. 2010;33(5): 585–592.
10. Pandi-Perumal SR, Verster JC, Kayumov L, et al. Sleep disorders, sleepiness and traffic safety: a public health menace. *Braz J Med Biol Res*. 2006;39(7): 863–871.
11. Consensus Conference Panel, Watson NF, Badr MS, et al. Recommended amount of sleep for a healthy adult: a joint consensus statement of the American Academy of Sleep Medicine and Sleep Research Society. *J Clin Sleep Med*. 2015;11(6):591–592.
12. Shepard JW Jr, Buysse DJ, Chesson AL Jr, et al. History of the development of sleep medicine in the United States. *J Clin Sleep Med*. 2005;1(1):61–82.
13. Ferrie JE, Kumari M, Salo P, et al. Sleep epidemiology—a rapidly growing field. *Int J Epidemiol*. 2011;40(6): 1431–1437.
14. Frank S, Gonzalez K, Lee-Ang L, et al. Diet and sleep physiology: public health and clinical implications. *Front Neurol*. 2017;8:393.
15. Spaeth AM, Dinges DF, Goel N. Effects of experimental sleep restriction on weight gain, caloric intake, and meal timing in healthy adults. *Sleep*. 2013;36(7):981–990.
16. Dashti HS, Scheer FA, Jacques PF, et al. Short sleep duration and dietary intake: epidemiologic evidence, mechanisms, and health implications. *Adv Nutr*. 2015;6(6):648–659.
17. St-Onge M-P, Mikic A, Pietrolungo CE. Effects of diet on sleep quality. *Adv Nutr*. 2016;7(5):938–949.
18. Castro-Diehl C, Wood AC, Redline S, et al. Mediterranean diet pattern and sleep duration and insomnia symptoms in the Multi-Ethnic Study of Atherosclerosis. *Sleep*. 2018; 41(11):zsy158.
19. Knutson KL. Sleep duration and cardiometabolic risk: a review of the epidemiologic evidence. *Best Pract Res Clin Endocrinol Metab*. 2010;24(5):731–743.
20. Ventura DdA, Fonseca VdM, Ramos EG. Association between quality of the diet and cardiometabolic risk factors in postmenopausal women. *Nutr J*. 2014;13(1):121.
21. Almeneessier AS, Pandi-Perumal SR, Bahammam ASJ. Intermittent fasting, insufficient sleep, and circadian rhythm: interaction and effects on the cardiometabolic system. *Curr Sleep Med Rep*. 2018;4(3):179–195.
22. Osborne ES Jr, Tabor EC. Report of a nutrition demonstration program in Ottawa County, Michigan. *Public Health Reports*. 1949;64(50):1603–1612.
23. Arlinghaus A, Lombardi DA, Willetts JL, et al. A structural equation modeling approach to fatigue-related risk factors for occupational injury. *Am J Epidemiol*. 2012;176(7): 597–607.
24. Young T, Finn L, Peppard PE, et al. Sleep disordered breathing and mortality: eighteen-year follow-up of the Wisconsin Sleep Cohort. *Sleep*. 2008;31(8):1071–1078.
25. Suzuki E, Yorifuji T, Ueshima K, et al. Sleep duration, sleep quality and cardiovascular disease mortality among the elderly: a population-based cohort study. *Prev Med*. 2009; 49(2–3):135–141.
26. Rao MN, Blackwell T, Redline S, et al. Association between sleep architecture and measures of body composition. *Sleep*. 2009;32(4):483–490.
27. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr*. 1997; 65(4 suppl):1220S–1228S.
28. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology*. 2009;20(4):488–495.
29. Buysse DJ, Reynolds CF 3rd, Monk TH, et al. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res*. 1989;28(2):193–213.
30. Cappuccio FP, Cooper D, D'Elia L, et al. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. *Eur Heart J*. 2011; 32(12):1484–1492.
31. Vgontzas AN, Liao D, Bixler EO, et al. Insomnia with objective short sleep duration is associated with a high risk for hypertension. *Sleep*. 2009;32(4):491–497.
32. Dunietz GL, Jansen EC, Chervin RD. On self-reported measurements: an epidemiologic perspective. *Sleep Med*. 2017;38:158–159.
33. Ibañez V, Silva J, Cauli O. A survey on sleep questionnaires and diaries. *Sleep Med*. 2018;42:90–96.
34. Sadeh A. The role and validity of actigraphy in sleep medicine: an update. *Sleep Med Rev*. 2011;15(4):259–267.
35. Marino M, Li Y, Rueschman MN, et al. Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep*. 2013;36(11): 1747–1755.
36. Cespedes EM, Hu FB, Redline S, et al. Comparison of self-reported sleep duration with Actigraphy: results from the Hispanic Community Health Study/Study of Latinos Sueno Ancillary Study. *Am J Epidemiol*. 2016;183(6): 561–573.
37. Jackson CL, Ward JB, Johnson DA, et al. Concordance between self-reported and actigraphy-assessed sleep duration among African-American adults: findings from the Jackson Heart Sleep Study. *Sleep*. 2020;43(3):zsz246.
38. Mallinson DC, Kamenetsky ME, Hagen EW, et al. Subjective sleep measurement: comparing sleep diary to questionnaire. *Nat Sci Sleep*. 2019;11:197–206.
39. Lauderdale DS, Chen JH, Kurina LM, et al. Sleep duration and health among older adults: associations vary by how sleep is measured. *J Epidemiol Community Health*. 2016; 70(4):361–366.
40. Chavarro JE, Watkins DJ, Afeiche MC, et al. Validity of self-assessed sexual maturation against physician assessments and hormone levels. *J Pediatr*. 2017;186:172–178.e3.
41. Ocké M, Kaaks RJ. Biochemical markers as additional measurements in dietary validity studies: application of the method of triads with examples from the European prospective investigation into cancer and nutrition. *Am J Clin Nutr*. 1997;65(4 suppl):1240S–1245S.
42. Kabagambe EK, Baylin A, Allan DA, et al. Application of the method of triads to evaluate the performance of food frequency questionnaires and biomarkers as indicators of long-term dietary intake. *Am J Epidemiol*. 2001;154(12): 1126–1135.
43. Silva GE, Goodwin JL, Sherrill DL, et al. Relationship between reported and measured sleep times: the Sleep Heart Health Study (SHHS). *J Clin Sleep Med*. 2007;3(6):622–630.
44. Dayyat EA, Spruyt K, Molfese DL, et al. Sleep estimates in children: parental versus actigraphic assessments. *Nat Sci Sleep*. 2011;3:115–123.

45. Herring SJ, Foster GD, Pien GW, et al. Do pregnant women accurately report sleep time? A comparison between self-reported and objective measures of sleep duration in pregnancy among a sample of urban mothers. *Sleep Breath.* 2013;17(4):1323–1327.
46. Van Den Berg JF, Van Rooij FJ, Vos H, et al. Disagreement between subjective and actigraphic measures of sleep duration in a population-based study of elderly persons. *J Sleep Res.* 2008;17(3):295–302.
47. Lauderdale DS, Knutson KL, Yan LL, et al. Self-reported and measured sleep duration: how similar are they? *Epidemiology.* 2008;19(6):838–845.
48. Jeon L, Finkelstein J. Consumer sleep tracking devices: a critical review. *Stud Health Technol Inform.* 2015;210:458–460.
49. Ko P-RT, Kientz JA, Choe EK, et al. Consumer sleep technologies: a review of the landscape. *J Clin Sleep Med.* 2015;11(12):1455–1461.
50. Rosenberger ME, Buman MP, Haskell WL, et al. Twenty-four hours of sleep, sedentary behavior, and physical activity with nine wearable devices. *Med Sci Sports Exerc.* 2016;48(3):457–465.
51. Walch O, Huang Y, Forger D, et al. Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device. *Sleep.* 2019;42(12):zsz180.
52. Ram S, Seirawan H, Kumar SKS, et al. Prevalence and impact of sleep disorders and sleep habits in the United States. *Sleep Breath.* 2010;14(1):63–70.
53. McCall WV, Turpin E, Reboussin D, et al. Subjective estimates of sleep differ from polysomnographic measurements in obstructive sleep apnea patients. *Sleep.* 1995;18(8):646–650.
54. Bianchi MT, Williams KL, McKinney S, et al. The subjective-objective mismatch in sleep perception among those with insomnia and sleep apnea. *J Sleep Res.* 2013;22(5):557–568.
55. Manconi M, Ferri R, Sagrada C, et al. Measuring the error in sleep estimation in normal subjects and in patients with insomnia. *J Sleep Res.* 2010;19(3):478–486.
56. Shockey TM, Wheaton AG. Short sleep duration by occupation group—29 states, 2013–2014. *MMWR Morb Mortal Wkly Rep.* 2017;66(8):207–213.
57. Bennett DA, Landry D, Little J, et al. Systematic review of statistical approaches to quantify, or correct for, measurement error in a continuous exposure in nutritional epidemiology. *BMC Med Res Methodol.* 2017;17(1):146.
58. Mendez MA, Wynter S, Wilks R, et al. Under- and overreporting of energy is related to obesity, lifestyle factors and food group intakes in Jamaican adults. *Public Health Nutr.* 2004;7(1):9–19.
59. Stommel M, Schoenborn CA. Accuracy and usefulness of BMI measures based on self-reported weight and height: findings from the NHANES & NHIS 2001–2006. *BMC Public Health.* 2009;9:421.
60. Hebert JR, Hurley TG, Peterson KE, et al. Social desirability trait influences on self-reported dietary measures among diverse participants in a multicenter multiple risk factor trial. *J Nutr.* 2008;138(1):226S–234S.
61. Spiegelman D, McDermott A, Rosner B. Regression calibration method for correcting measurement-error bias in nutritional epidemiology. *Am J Clin Nutr.* 1997;65(4 suppl):1179s–1186s.
62. Tantrakul V, Numthavaj P, Guilleminault C, et al. Performance of screening questionnaires for obstructive sleep apnea during pregnancy: a systematic review and meta-analysis. *Sleep Med Rev.* 2017;36:96–106.
63. Janssen KC, Phillipson S, O'Connor J, et al. Validation of the Epworth Sleepiness Scale for Children and Adolescents using Rasch analysis. *Sleep Med.* 2017;33:30–35.
64. Nagappa M, Liao P, Wong J, et al. Validation of the STOP-Bang questionnaire as a screening tool for obstructive sleep apnea among different populations: a systematic review and meta-analysis. *PLoS One.* 2015;10(12):e0143697.