Sleep duration and health among older adults: associations vary by how sleep is measured

Diane S. Lauderdale, **Jen-Hao Chen**, **Lianne M. Kurina**, **Linda J. Waite**, and **Ronald A. Thisted**

Department of Public Health Sciences (Diane S. Lauderdale, Ronald A. Thisted) and Department of Sociology (Linda Waite), University of Chicago; Department of Health Sciences, University of Missouri (Jen-Hao Chen); General Medical Disciplines, Stanford University School of Medicine (Lianne M. Kurina)

Abstract

Background—Cohort studies have found that short and long sleep are both associated with worse outcomes, compared to intermediate sleep times. While demonstrated biological mechanisms could explain health effects for short sleep, long-sleep risk is puzzling. Most studies reporting the U shape use a single question about sleep duration, a measurement method that does not correlate highly with objectively measured sleep. We hypothesized that the U shape, especially the poor outcomes for long sleepers, may be an artifact of how sleep is measured.

Methods—We examined the cross-sectional prevalence of fair/poor health by sleep hour categories (6, 7, 8, 9, >9 hours) in a national U.S. sample of adults aged 62–90 that included several types of sleep measures (n=727). Survey measures were: a single question; usual bedtimes and waking times; and a three-day sleep log. Actigraphy measures were the sleep interval and total sleep time. Fair/poor health was regressed on sleep hour categories adjusted for demographics, with tests for both linear trend and U shape.

Results—Adjusted odds ratios of fair/poor health across sleep hour categories from the single question were 4.6, 2.2, referent (8 hours), 1.8 and 6.9. There was high prevalence of fair/poor health for 6 hours for all sleep measures, but the long sleep effect was absent for sleep logs and actigraphy measures.

Conclusion—Associations between long sleep and poor health may be specific to studies measuring sleep with survey questions. As cohorts with actigraphy mature, our understanding of how sleep affects health may change.

Keywords

Sleep; measurement; actigraphy; self-rated health

INTRODUCTION

Many cohort studies that included a question about sleep duration, beginning with the first American Cancer Society cohort and the Alameda County Study,[1–3] have provided evidence that sleep duration predicts mortality. Reviews and meta-analyses have concluded that individuals sleeping either less or more than the referent category, usually seven and eight hours, have higher mortality, with generally greater risk for long than short sleep.[4–6] Meta-analyses similarly reported U-shaped associations for incident diabetes and stroke. [7,8] The U shape has been challenging to interpret because the two tails are unlikely to have the same mechanisms or explanation.[9, 10] Researchers point to experimental evidence that total or partial sleep deprivation alters blood pressure, hormones and inflammation in ways that could lead to cardiovascular disease [11], but there is no parallel evidence about a biological mechanism for long sleep. Confounding by health conditions such as obesity and depression, reverse causality, and measurement bias have all been suggested as explanations of the long sleep effect.[12–14]

In 2015, the National Sleep Foundation, the American Academy of Sleep Medicine/Sleep Research Society, and the American Thoracic Society each issued new sleep recommendations. [15–17] Their expert panels reached divergent views about long sleep. The National Sleep Foundation placed more than 10 hours in a "not recommended" category for adults aged 26–64, and more than 9 hours "not recommended" for adults 65 and older. The other two recommendations conveyed uncertainty about whether long sleep was associated with adverse outcomes.

Almost all studies reporting the U-shape have measured sleep using a single survey question, such as "How many hours do you usually sleep at night?"[18] There is no standard wording; questions may ask about "typical" or "average" sleep, at night or over 24 hours. A few studies instead ask for typical bedtime and wake time and then calculate duration. Sleep duration has also been extracted from 24-hour time diaries.[19,20] Objective measurement methods include wrist actigraphy, which estimates sleep from arm motion, and polysomnography, which determines sleep from electrical brain activity. Polysomnography is the gold standard for determining sleep, but there is no gold standard for habitual sleep behavior because polysomnography requires the application of multiple electrodes and may disrupt routine. Epidemiological cohorts have increasingly included actigraphy.[21–27]

When studies measure sleep in more than one way, the correlation between survey responses and an objective estimate is in the low to moderate range.[21,28–30] Some individuals may think about time in bed rather than sleep. While the straightforward wording of a sleep duration question suggests it should be as easy to answer as a question about one's weight, it is likely that many do not know their average sleep duration and respond using strategies other than reporting a known quantity. If that is the case, factors such as believing that short sleep is evidence of a full life or that eight hours is a good answer may introduce systematic bias as well as random noise.

In an ongoing cohort study of older Americans, the National Social Life Health and Aging Project (NSHAP), we collected sleep information using survey questions, a sleep log and actigraphy. Here we compare associations between different measures of sleep duration and self-rated health. Self-rated health ("In general, would you say your health is excellent, very good, good, fair or poor?") is widely included in surveys in part because there is a robust association between the two worst categories (fair and poor) and elevated mortality risk.[31] Self-rated health does not reference a single dimension of health, and research has found that it draws upon respondents' assessments of their illnesses, physical functioning, mental health and health behaviors, in varying degrees for different respondents.[32] Our primary hypothesis was that the association between worse health and sleep duration is sensitive to how sleep is measured, particularly for long sleep. Were we to find this to be the case, there would be three implications: (1) evidence of deleterious health effects for long sleep could be an artifact of measurement method; (2) studies of sleep duration effects on health are subject to confounding by baseline health; and (3) as cohorts with actigraphy mature, they may not confirm associations found for self-reported sleep duration.

METHODS

Data

NSHAP—This longitudinal study of health and social life in the U.S. is based on a national probability sample of community-residing adults born between 1920 and 1947. Wave 1 included 3005 individuals and was conducted with in-home interviews in 2005–2006. Wave 2 in 2010–2011 extended the sample by inviting spouses of respondents, yielding a Wave 2 sample of 3,377. A random one-third of Wave 2 respondents were invited to participate in a sleep substudy, fully described elsewhere.[25] Eighty percent (897/1,117) agreed to participate during the in-home interview. Forty-eight could not be reached by phone or changed their mind when recontacted to arrange delivery of the survey booklet, actigraph and prepaid return mailer. The booklet included instructions, a 3-day sleep log and additional sleep questions. Participants were asked to wear an Actiwatch Spectrum model from Phillips Respironics (Phillips Healthcare, Andover MA), for 72 hours straight.

Usable data were returned from 780 participants, but 53 spouses did not have birth years between 1920 and 1947, resulting in a final sample of 727 individuals aged 62 to 90. Demographic characteristics of the actigraphy subsample were similar to the full age-eligible NSHAP sample, but mean age was 0.4 years younger, the sample was one percent higher female, and the percentage white was two percent greater.

Self-rated Health: Respondents were asked: "Would you say your health is -- excellent, very good, good, fair, or poor?" We dichotomized this into fair/poor versus excellent/very good/good.

Sleep data: *survey measures***:** All sleep measures reference the main sleep period and do not include naps. The time interval between the in-home survey and the mailed sleep substudy allowed us to ask about sleep in different ways without inducing consistency in the same interview. In the sleep booklet, we asked: "How many hours do you usually sleep at night?" We refer to this as *survey sleep hours*. In the core interview, individuals were asked

their usual bedtimes and wake times, separately for weeknights and weekends: "What time do you usually go to bed and start trying to fall asleep?" and "What time do you usually wake up?" Average sleep duration was calculated and weighted for weeknights (5/7) and weekends (2/7). Weeknights averaged 10 minutes shorter than weekends. We refer to this as *survey calculated sleep time*. Finally, individuals concurrently kept a sleep log for the nights when they wore an actigraph, recording hour and minutes in response to: "When you went to bed, what time did you start trying to fall asleep?" and "What time did you wake up?" The average calculated intervals we refer to as *sleep log time*. Sleep logs are considered the criterion standard of self-reported sleep data. [33] We checked whether the sleep log time differed by whether the three nights were all weeknights or included two weekend nights, and they did not; the difference in the means was 0.06 hours (p=0.77), with the weekend average *shorter*. Therefore, we do not take into account the days of the week for the sleep substudy.

Sleep data: *actigraph measures*: The Actiwatch Spectrum records intensity and frequency of movement in epochs (15 second epochs were selected for this study). This model has an event marker, which leaves a time stamp but does not start or stop recording. Participants were instructed to push the event marker when they started trying to fall asleep each night and when they awoke in the morning; not all remembered to push the event marker consistently. The model includes a light sensor. After the Actiwatch was returned the data were downloaded and analyzed using the manufacturer's Actiware software (version 5.59). The rest interval—the period between when the individual starts to try to fall asleep and wakes up— was set initially by the software based on activity pattern and re-examined by the investigators to ensure it captured the full sleep period. The investigators made use of the event marker and the ambient light level, which the software does not use, to revise the interval. The event marker, when used, was considered the best evidence. Investigators did not know self-reported health status when setting the intervals. Using revised intervals, the software estimated sleep parameters based on patterns of activity counts within and across epochs. We used the manufacturer's default cutoff of 40 as the activity count threshold to score each epoch as sleep or wake. Philips Respironics recommends this threshold because it was used for their validation, including setting the sleep start and end times, which are based on contiguous epochs scored as sleep. We also used the default settings of ten minutes of immobile epochs to define sleep onset and sleep offset. We carried out a sensitivity analysis with a lower threshold of 20. It resulted in an average 18 minutes less sleep, but the correlation between sleep times using the two thresholds was greater than 0.99, so the shape of the associations across ordered categories would be similar, but shifted.

We examine two actigraphy durations: *actigraph sleep interval,* the time between the first and last epochs scored as sleep and *actigraph total sleep time,* the summed duration of epochs scored as sleep during the sleep interval. The difference between the two is the total minutes of *wake after sleep onset* (*WASO*).

Analysis

The analysis focuses on the pattern of cross-sectional associations between sleep and fair/poor health. We tested for U-shaped and linear associations.

For all duration measures, we used the same 5 categories: 6 hours, > 6 and 7 hours, >7 and 8 hours, > 8 and 9 hours and > 9 hours. WASO was divided into quintiles.

First, the proportions of individuals reporting fair/poor health were compared across categories of each sleep measure in bar graphs. Then the significance of linear trend or U-shaped association was tested in logistic regression models, adjusted for age, gender, and race/ethnicity (white, black, Hispanic and other). In each model, the outcome was a dichotomous indicator of fair/poor health and the independent variables included indicators for sleep categories. We used a one-degree-of-freedom score test for a U-shaped association for each model, which weights the categories as 2, 1, 0, 1, 2. These weights are most sensitive for detecting growth at a linear rate on either side of the central category. They place less emphasis on the extreme quintiles than quadratic weights would. We used a single ordinal variable to test linear trend. We did not adjust for health behaviors or co-morbidities, because these are likely to be factors considered by respondents in answering the health question. Interaction terms between sleep categories and gender were tested, and none were significant. For actigraphy total sleep time, we examined whether the score test was sensitive to how sleep was categorized by re-categorizing so that the proportion of the sample in each of the five categories matched the proportions for survey sleep hours.

The investigators received de-identified data for this study, and the Institutional Review Board at the University of Chicago considers such data exempt. NSHAP is a complex survey and all analyses take into account the study design and observation weights using the "svy" commands in Stata 14 (StataCorp. 2015. *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP).

RESULTS

Table 1 presents descriptive statistics. Overall, 19.9 percent of participants reported fair/poor health. Mean sleep durations ranged from 7.2 hours for actigraph total sleep time to 8.2 hours for survey calculated sleep time and sleep log time.

Table 2 presents pairwise correlations between duration measures. The highest correlation is 0.56, between sleep log time and both survey calculated sleep time and actigraph sleep interval. Correlations between actigraph total sleep time and the two survey measures were lower.

Percentages of individuals reporting fair/poor health across categories of sleep duration are shown in Figure 1. For survey sleep hours there is a marked U shape, with high percentages reporting fair/poor health for 6 hours (30%) and > 9 hours (39%), compared to 9% in the middle category (8 hours). There is also a U shape for survey calculated sleep time, but it is shallower, including a higher proportion in the middle (14%). For sleep log time, there is a high percentage (41%) reporting fair/poor health in the 6 hour category, and slight variation across the remaining categories (22%, 17%, 16%, 21%). For actigraph duration measures, there are elevated proportions of fair/poor health in the 6 hour category (34% for actigraph sleep interval and 31% for total sleep time), but little pattern across the other

categories. For WASO, there is a monotomic increase in fair/poor health as WASO increases, from 12% up to 30%.

Table 3 presents coefficients from adjusted regression models, confirming patterns seen in the bar graphs. For survey sleep hours, the odds ratios are 4.59 and 6.86 for the shortest and longest categories, and the score test for U shape is highly significant (p < 0.001). The score test is also significant for survey calculated sleep time, but the odds ratios are lower at the ends. For sleep log time, there is a suggestion of a U shape, but the score test is not significant and the odds ratio is just 1.32 in the longest category. For the two actigraph duration measures, there is little variation in the odds ratios greater than 6 hours; but the linear trend tests do suggest a significant trend of *decreasing* odds of fair/poor health with increasing sleep. The score test was also not significant (p=0.53) in a sensitivity analysis where actigraph total sleep time was re-categorized so the proportions in each category matched those for survey sleep hours (data not shown). For WASO, there are increasing odds of fair/poor health with increasing WASO.

DISCUSSION

We found that cross-sectional associations between hours of sleep and fair/poor health varied by how sleep was measured. For the shortest sleep category, here defined as   6 hours, there was a consistently high prevalence of fair/poor health for all five sleep measures. However, poorer health among long sleepers was only observed when survey questions asked about general behavior, and not when sleep data were collected for specific nights, either through sleep logs or by actigraphy. The U shape was especially marked when sleep was measured with a single survey question. There was also a U shape, but it was shallower, when sleep was calculated from survey questions about usual bedtime and waking time. For a three-day average of sleep log times, the elevated prevalence in the longest sleep category was barely discernible, and it was absent for actigraphy measures.

Previous studies have reported U-shaped associations between poorer self-rated health and sleep duration from single questions.[34,35] We are unaware of other studies comparing associations for survey and actigraph sleep duration measures. A study among older adults in Great Britain compared health associations using a single question versus calculated duration from survey questions. They found opposite associations for the two sleep measures, but did not appear to have explored nonlinear patterns.[36] An Australian study directly assessed the longitudinal confounding effect of self-rated health on sleep and mortality. Having found increased odds of poorer health among those reporting short and long sleep, they found no association between sleep hours and mortality for those with better baseline health and a U-shaped association for those with worse baseline health.[37]

Cohort studies provide evidence of the temporal ordering of risk factors and outcomes. However, the cross-sectional U-shaped association for the single question suggests that poor health may precede or bias responses to survey questions about sleep duration. It is also possible that poor sleep has already affected health at baseline in this older population. An association between self-reported sleep hours and baseline health status could lead to a U-shaped association between sleep hours and any outcome that self-rated health strongly

predicts, including mortality, without adequate adjustment for baseline health. However, this potential confounding or bias would be weaker were sleep measured any other way. Sleep assessed with a single question may be subject to same-source bias,[38] with reports of long sleep perhaps conveying information about low energy or dissatisfaction with health or life. Because the factors that lead to individuals endorsing different levels of self-rated health are not well understood, it is impossible to know whether adjusting for self-rated health *per se* at baseline could leave residual confounding by the underlying factors self-rated health signals.

The main limitation to our study is that we have only three nights of actigraphy, because of concern about respondent burden. Most epidemiologic studies with actigraphy collect three to ten nights. Previous research has found that three nights are adequate to estimate mean sleep characteristics for older adults, although not variability.[39] Nonetheless, it is possible that nightly variability diminished our ability to detect associations for sleep log or actigraph measures. However, our finding of *linear* associations between poor health and actigraph measures, especially WASO, makes it less likely that excess variability would explain the absence of a U shape for actigraph total sleep time. While actigraphy is an imperfect measure of sleep duration, studies have generally found a high correlation of about 0.9 compared to concurrent polysomnography.[40]

Nonetheless health factors could be related to sources of inaccuracy in actigraphy estimates. To mask a U shape for actigraphy, there would need to be a correlate of poor health that caused systematic underestimation of sleep for long sleepers. Another limitation is that NSHAP did not ask about sleep apnea. However, apnea diagnosis has generally also been lacking from the cohorts used to examine sleep hours and mortality. WASO may be greater for individuals with apnea. Finally, these results may only generalize to community-dwelling older adults in the United States.

Our data are cross-sectional; in the future, we will be able to compare health outcomes associated with sleep hours measured in different ways. As cohorts with actigraphy mature, we may find different associations with subsequent health events than have been reported to date.

Acknowledgments

References

1. Hammond EC. Some preliminary findings on physical complaints from a prospective Study of 1,064,004 men and women. Am J Public Nations Health. 1964; 54:11–23.

2. Kripke DF, Simons RN, Garfinkel L, et al. Short and long sleep and sleeping pills. Is increased mortality associated? Arch Gen Psychiatry. 1979; 36:103–116. [PubMed: 760693]

3. Wingard DL, Berkman LF. Mortality risk associated with sleeping pattern among adults. Sleep. 1983; 6:102–107. [PubMed: 6878979]

4. Grandner MM, Hale L, Moore M, et al. Mortality associated with short sleep duration: the evidence, the possible mechanisms and the future. Sleep Med Rev. 2010; 14:585–592.

5. Gallicchio L, Kalesan B. Sleep duration and mortality: a systematic review and meta-analysis. J Sleep Res. 2009; 18:148–158. [PubMed: 19645960]

6. 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; 33:585–592. [PubMed: 20469800]

7. Shan Z, Ma H, Xie M, et al. Sleep duration and risk of type 2 diabetes: a meta-analysis of prospective studies. Diabetes Care. 2015; 38(3):529–37. [PubMed: 25715415]

8. Leng Y, Cappuccio FP, Wainwright NW, et al. Sleep duration and risk of fatal and nonfatal stroke: a prospective study and meta-analysis. Neurology. 2015; 84:1072–9. [PubMed: 25716357]

9. Bliwise DL, Young TB. The parable of parabola: what the U-shaped curve can and cannot tell us about sleep. Sleep. 2007; 30(12):1614–5. [PubMed: 18246971]

10. Knutson KL, Turek FW. The U-shaped association between sleep and health: the 2 peaks do not mean the same thing. Sleep. 2006 Jul; 29(7):878–9. [PubMed: 16895253]

11. Mullington JM, Haack M, Toth M, et al. Cardiovascular, inflammatory, and metabolic consequences of sleep deprivation. Prog Cardiovasc Dis. 2009 Jan-Feb;51(4):294–302. [PubMed: 19110131]

12. Grandner MA, Drummond SP. Who are the long sleepers? Towards an understanding of the mortality relationship. Sleep Med Rev. 2007; 11(5):341–360. [PubMed: 17625932]

13. Stamatakis KA, Punjabi NM. Long sleep duration: a risk to health or a marker of risk? Sleep Med Rev. 2007; 11(5):337–9. [PubMed: 17854737]

14. Patel SR, Malhotra A, Gottlieb DJ, et al. Correlates of long sleep duration. Sleep. 2006; 29(7):881–9. [PubMed: 16895254]

15. Mukherjee S, Patel SR, Kales SN, et al. American Thoracic Society ad hoc Committee on Healthy Sleep. An Official American Thoracic Society Statement: The Importance of Healthy Sleep. Recommendations and Future Priorities. Am J Respir Crit Care Med. 2015 Jun 15; 191(12):1450–8. [PubMed: 26075423]

16. Hirshkowitz M, Whiton K, Albert SM, et al. National Sleep Foundation's sleep time duration recommendations: methodology and results summary. Sleep Health. 2015; 1:40–43.

17. Watson NF, Badr MS, Belenky G, 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. SLEEP. 2015; 38(6):843–844. [PubMed: 26039963]

18. Kurina LM, McClintock MK, Chen JH, et al. Sleep duration and all-cause mortality: a critical review of measurement and associations. Ann Epidemiology. 2013; 23:361–370.

19. Basner M, Spaeth AM, Dinges DF. Sociodemographic characteristics and waking activities and their role in the timing and duration of sleep. Sleep. 2014 Dec 1; 37(12):1889–906.10.5665/sleep. 4238 [PubMed: 25325472]

20. Antillón M, Lauderdale DS, Mullahy J. Sleep behavior and unemployment conditions. Econ Hum Biol. 2014 Jul.14:22–32. [PubMed: 24958451]

21. Van Den Berg JF, Miedema HMM, Tulen JHM, et al. Sex differenced in subjective and actigraphic sleep measures: a population-based study of elderly persons. Sleep. 2009; 32:1367–1375. [PubMed: 19848365]

22. Appelhans BM, Janssen I, Cursio JF, et al. Sleep duration and weight change in midlife women: the SWAN sleep study. Obesity (Silver Spring). 2013; 21:77–84. [PubMed: 23505171]

23. Lauderdale DS, Knutson KL, Yan LL, et al. Objectively measured sleep characteristics among early-middle-aged adults: the CARDIA study. Am J Epidemiol. 2006; 164(1):5. [PubMed: 16740591]

24. Fung MM, Peters K, Ancoli-Israel S, et al. Osteoporotic Fractures in Men (MrOS) Research Group. Total sleep time and other sleep characteristics measured by actigraphy do not predict incident hypertension in a cohort of community-dwelling older men. J Clin Sleep Med. 2013; 9(6):585–91. [PubMed: 23772192]

25. Lauderdale DS, Philip Schumm L, Kurina LM, et al. Assessment of sleep in the National Social Life, Health, and Aging Project. J Gerontol B Psychol Sci Soc Sci. 2014; 69(Suppl 2):S125–33. [PubMed: 25360013]

26. Spira AP, Covinsky K, Rebok GW, et al. Poor sleep quality and functional decline in older women. J Am Geriatr Soc. 2012; 60(6):1092–8. [PubMed: 22690985]

27. Anderson KN, Catt M, Collerton J, et al. Assessment of sleep and circadian rhythm disorders in the very old: the Newcastle 85+ Cohort Study. Age Ageing. 2014; 43(1):57–63. [PubMed: 24123786]

28. Silva GE, Goodwin JL, Sherrill DL, et al. Relationship between reported and measured sleep times. J Clin Sleep Med. 2007; 6:622–630. [PubMed: 17993045]

29. Lauderdale DS, Knutson KL, Yan LL, et al. Self-reported and measured sleep duration: how similar are they? Epidemiology. 2008; 19:838–845. [PubMed: 18854708]

30. Mezick EJ, Wing RR, McCaffery JM. Associations of self-reported and actigraphy-assessed sleep characteristics with body mass index and waist circumference in adults: moderation by gender. Sleep Med. 2014 Jan; 15(1):64–70. [PubMed: 24239499]

31. Idler EL, Benyamini Y. Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies. Journal of Health and Social Behavior. 1997; 38:21–37. [PubMed: 9097506]

32. Krause NM, Jay GM. What do global self-rated health items measure? Med Care. 1994 Sep. 32:930–42. [PubMed: 8090045]

33. Miller CB, Gordon CJ, Toubia L, et al. Agreement between simple questions about sleep duration and sleep diaries in a large online survey. complete citation. Sleep Health. 2015; 1(2):133–137.

34. Shankar A, Charumathi S, Kalidindi S. Sleep duration and self-rated health: the national health interview survey 2008. Sleep. 2011; 34:1173–7. [PubMed: 21886354]

35. Kim JH, Kim KR, Cho KH, et al. The association between sleep duration and self-rated health in the Korean general population. J Clin Sleep Med. 2013; 9:1057–64. [PubMed: 24127150]

36. Leng Y, Wainwright NW, Cappuccio FP, et al. Self-reported sleep patterns in a British population cohort. Sleep Med. 2014; 15:295–302. [PubMed: 24503473]

37. Magee CA, Holliday EG, Attia J, et al. Investigation of the relationship between sleep duration, all-cause mortality, and preexisting disease. Sleep Med. 2013; 14:591–6. [PubMed: 23517587]

38. Podsakoff PM, MacKenzie SB, Lee JY, et al. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol. 2003; 88(5):879–903. [PubMed: 14516251]

39. Rowe M, McCrae C, Campbell J, et al. Actigraphy in older adults: comparison of means and variability of three different aggregates of measurement. Behavioral Sleep Medicine. 2008; 6:127–45. [PubMed: 18443950]

40. Ancoli-Israel S, Cole R, Alessi C, et al. The role of actigraphy in the study of sleep and circadian rhythms. Sleep. 2003; 26:342–92. [PubMed: 12749557]

What is already known on this subject?

Numerous studies have found that sleep hours predict subsequent health outcomes, including mortality, in a U shape, with both short and long sleepers having worse outcomes than people reporting seven or eight hours. Almost all of these studies have measured sleep duration with a single survey question. Survey questions about sleep duration do not correlate highly with objective sleep measures.

What does this study add?

We find that short sleep is always associated with poorer concurrent health, no matter how sleep is measured (i.e., different types of survey questions, sleep log and wrist actigraphy), but long sleep is not consistently associated with poor health: the association is strongest using a single survey question and absent using actigraphy. Our findings suggest that U-shaped associations previously reported may be due to baseline confounding by health status when sleep is measured by a single question.
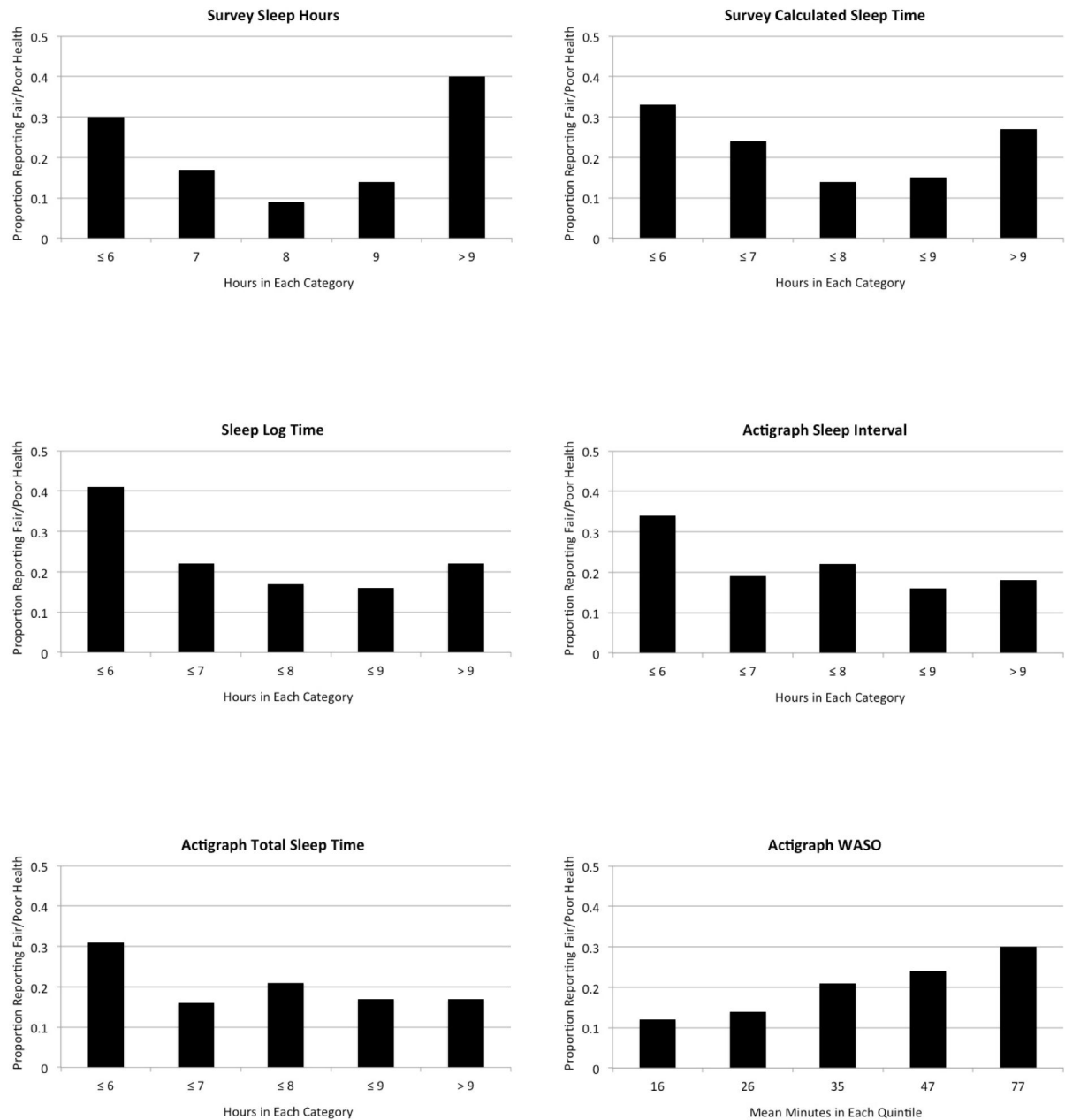
Figure 1.
The prevalence of fair/poor health by sleep hours measured five different ways, and by WASO quintiles. Data are from the NSHAP Sleep Substudy, 2010–2011.

Table 1

Study population from the NSHAP Sleep Substudy, 2010–2011 (N=727), estimated using survey design and weights.

Characteristic	Total	Excellent/Very Good/Good Health 80.1%	Fair/Poor Health 19.9%
Mean age (standard error)	71.8 (.33)	71.7 (.33)	72.4 (.79)
Race/ethnicity (%)			
White	83.2	84.6	78.0
Black	7.1	6.9	7.9
Hispanic, non-Black	6.0	4.5	12.0
Other	3.7	4.1	2.2
Male (%)	46.5	45.8	49.5
Female (%)	53.5	54.2	50.1
Mean of Sleep Measures (standard error)			
Survey Sleep Hours	7.5 (.06)	7.5 (.07)	7.2 (.16)
Survey Calculated Sleep Time (hours)	8.2 (.06)	8.2 (.05)	8.3 (.16)
Sleep Log Time (hours)	8.2 (.06)	8.2 (.07)	8.3 (.16)
Actigraph Sleep Interval (hours)	7.9 (.06)	7.9 (.06)	7.8 (.13)
Actigraph Total Sleep Time (hours)	7.2 (.05)	7.3 (.06)	7.1 (.12)
Actigraph WASO (minutes)	39.1 (1.4)	37.3 (1.6))	46.3 (2.7)

Table 2

Pairwise Pearson correlations between survey-based and actigraph-estimated sleep duration measures. Data are from the NSHAP Sleep Substudy, 2010–2011 (N=727), estimated using survey weights. The statistical significance of each correlation is in parentheses.

	A	B	C	D	E
A. Survey Sleep Hours	1				
B. Survey Calculated Sleep Time	0.39 (<.0001)	1			
C. Sleep Log Time	0.46 (<.0001)	0.56 (<.0001)	1		
D. Actigraph Sleep Interval	0.25 (<.0001)	0.39 (<0.0001)	0.56 (<.0001)	1	
E. Actigraph Total Sleep Time	0.29 (<.0001)	0.38 (<.0001)	0.52 (<.0001)	0.95 (<.0001)	1

Table 3

Logistic regression models that regress fair/poor self-reported health on six sleep metrics. Each sleep metric reported represents a different model and all models are adjusted for age, gender and race/ethnicity (four categories). Data are from the NSHAP Sleep Substudy, 2010–2011 (N=727). Estimates take into account survey design and weights in computing variance.

Fair/Poor Health	6 hours	> 6 and 7 hours	> 7 and 8 hours	> 8 and 9 hours	> 9 hours	Score test for U-shape	p-trend
			Odds Ratios (95% Confidence Interval)				
Survey Sleep Hours	4.59 (2.50,8.43)	2.23 (.95,5.22)	1 ref	1.82 (.87,3.81)	6.86 (2.27, 20.7)	<.001	.14
Survey Calculated Sleep Time	2.77 (1.47,5.21)	1.83 (.91,3.69)	1	1.07 (.53,2.15)	2.28 (1.30,4.04)	.001	.92
Sleep Log Time	3.28 (1.57,6.82)	1.35 .(68,2.70)	1 ref	0.95 (.45,2.00)	1.32 (.76,2.31)	.089	.27
Actigraph Sleep Interval	2.29 (1.00, 5.23)	0.94 (.57,1.57)	1 ref	0.70 (.39, 1.27)	0.80 (.37,1.73)	.87	.021
Actigraph Total Sleep Time	1.92 (1.07,3.48)	.72 (.39,1.33)	1 ref	.78 (.35,1.72)	.80 (.34,1.85)	.39	.051
			Quintiles				
Actigraph WASO	.53 (.24,1.18)	.57 (.29,1.10)	1 ref	1.13 (.62,2.05)	1.53 (.71,3.30)	.85	.016