Agreement between self-reported and objectively measured sleep duration among white, black, Hispanic, and Chinese adults in the United States: Multi-Ethnic Study of Atherosclerosis FREE

Chandra L Jackson, Sanjay R Patel, W Braxton Jackson, II, Pamela L Lutsey, Susan Redline

PDF   Split View   Cite   Permissions   Share ▾

Abstract

Study Objectives

To identify systematic biases across groups in objectively and subjectively measured sleep duration.

Methods

We investigated concordance of self-reported habitual sleep duration compared with actigraphy- and single-night in-home polysomnography (PSG) across white, black, Hispanic, and Chinese participants in the Multi-Ethnic Study of Atherosclerosis.

Results

Among 1910 adults, self-reported sleep duration, determined by differences between bed and wake times, was overestimated in all racial groups compared with PSG and actigraphy. Compared with whites ($\rho = 0.45$), correlations were significantly lower only in blacks ($\rho = 0.28$). Self-reporting bias for total sleep time compared with wrist actigraphy was 66 min (95% confidence interval [CI]: 61–71) for whites, 58 min (95% CI: 48–69) for blacks, 66 min (95% CI: 57–74) for Hispanics, and 60 min (95% CI: 49–70) for Chinese adults. Compared with PSG, self-reporting bias in whites at 73 min (95% CI: 67–79) was higher than in blacks (54 min [95% CI: 42–65]) and Chinese (49 min [95% CI: 37–61]) but not different from Hispanics (67 min [95% CI: 56–78]). Slight agreement/concordance was observed between self-reported and actigraphy-based total sleep time ($k_w = 0.14$ for whites, 0.10 for blacks, 0.17 for Hispanics, and 0.11 for Chinese) and PSG ($k_w = 0.08$ for whites, 0.04 for blacks, 0.05 for Hispanics, and 0.01 for Chinese) across race/ethnicity.

Conclusions

Self-reported sleep duration overestimated objectively measured sleep across all races, and compared with PSG, overestimation is significantly greater in whites compared with blacks. Larger reporting bias reduces the ability to identify significant associations between sleep duration and health among blacks compared with whites. Sleep measurement property differences should be considered when comparing sleep indices across racial/ethnic groups.

sleep duration, objective, subjective, race, ethnicity

Statement of Significance

Little prior research has investigated the extent to which self-reported measures may misclassify sleep duration. We found that self-reported sleep duration overestimated objectively measured sleep across all racial/ethnic groups. Self-reported sleep duration only moderately correlated ($\rho = 0.38$) with multiday actigraphy-based assessments and even less with single-night polysomnography (PSG) ($\rho = 0.20$). Discrepancies between PSG and actigraphy were largely explained by underestimated wake after sleep onset as measured by actigraphy. Self-reported compared with PSG-measured overestimation of sleep duration was greater for whites compared with blacks and Chinese. However, significantly higher correlations for self-reported sleep duration and objectively measured sleep as well as better calibration (for some measures) among whites compared with blacks suggest a greater potential for random misclassification of sleep in blacks compared with whites.

Introduction

Short sleep duration (<7 hr) is associated with increased risk of obesity, diabetes, hypertension, cardiovascular disease (CVD), and premature mortality [1–3]. Most epidemiological studies rely on self-reported sleep duration, and few studies have investigated the extent to which self-reported sleep reflects average objectively measured sleep. Most of these studies were conducted among predominately white populations [4–11], with a recent study examining this question in a Hispanic population [12].

Two common sleep disorders—insomnia and sleep apnea—each may influence the relationship between self-reported and objectively measured sleep duration. In particular, insomnia may influence perceived sleep duration and sleep apnea may reduce the accuracy of actigraphy-based sleep duration. However, the influence of these disorders on measurement error (or self-reporting bias) across population samples has not been addressed. No study has yet also compared sleep duration measures as derived from a questionnaire versus from a polysomnography (PSG) among white, black, Latino, and Chinese participants.

To address these important gaps in the literature, we analyzed wrist actigraphy, PSG, and sleep questionnaires obtained in the Multi-Ethnic Study of Atherosclerosis (MESA), with the following aims: to (1) investigate reporting bias of self-reported sleep duration (based on reported bed and wake times) on weekdays compared with three objective measures of sleep duration (i.e. average nightly sleep duration over weekdays and average time in bed, from 5 day wrist actigraphy; single-night sleep duration from in-home PSG) within and between white, black, Hispanic, and Chinese study participants; (2) determine the concordance/agreement of self-reported vs. multiday actigraphy compared with in-home PSG (the standard approach for directly measuring electroencephalography (EEG)-defined sleep periods); (3) determine the concordance of multiday actigraphy compared with single-night in-home PSG; (4) across races, determine whether there is differential measurement error or bias for concurrently measured single night actigraphy vs. in-home PSG; and (5) evaluate the influence of prespecified subgroups and health conditions on reporting bias differences between subjective and objective sleep duration. We hypothesized that (1) self-reported data from questionnaires will overestimate sleep duration compared with both actigraphy (for total sleep time and average time in bed) and PSG; (2) self-reporting biases of sleep duration will vary across racial/ethnic groups, but the direction of the reporting biases will not differentially vary (i.e. either under- or overestimation) between the various racial/ethnic groups; (3) self-reported sleep duration will more closely agree with actigraphy-determined time in bed than actigraphy- or PSG-based total sleep time and will not significantly vary between racial/ethnic groups; and (4) reporting biases for each racial group will not significantly vary by self-reported symptoms of insomnia and sleep apnea. These findings will help us to interpret the literature on sleep duration as well as inform future research aimed at understanding sleep health as objective sleep data may be necessary if subjective, self-reported data are shown to produce self-reporting bias or mask true associations.

Methods

Multi-Ethnic Study of Atherosclerosis

The data utilized in these analyses come from the MESA, a longitudinal study which at baseline (2000–2002) recruited 6814 men and women who were between 45 and 84 years old and from six US cities. The study was designed to prospectively examine risk factors for subclinical CVD and its progression to clinical disease. Each site recruited between 1066 and 1319 eligible (e.g. free of clinical CVD at baseline) participants, equally divided between men and women, and according to specified race/ethnicity proportions. Although the cohort was community-based, the emphasis of MESA

sampling was to obtain balanced recruitment across strata defined by sex, ethnicity, and age group rather than to represent the demographic distribution of the source communities. Approximately, 38.5 per cent of the cohort is white, 27.7 per cent black/African-American, 22 per cent Hispanic, and 11.8 per cent Asian, predominantly of Chinese descent. This analysis utilized data on sleep measurements made as part of an ancillary study that was conducted shortly after the fifth examination (2010–2013). All MESA participants except those reporting regular use of oral devices, nocturnal oxygen, or nightly positive airway pressure devices were invited to participate in the MESA Sleep Ancillary Study, which consisted of PSG, actigraphy, and sleep questionnaire data collected during an in-home examination. Of 4077 participants approached, 147 (6.5%) were ineligible (95 due to a history of the use of positive airway pressure [2%], 4 due to use of an oral appliance, and 4 due to oxygen use) and 141 participants lived too far away to participate. Of the remaining 3789 participants, 2261 participated in the sleep exam (59.7%). Of these, 1910 participants were eligible for these analyses as they had complete sleep questionnaire, wrist actigraphy, and PSG data. This subcohort was very similar to the full sample in terms of sex and race, whereas the subcohort was older and more educated. The Partners Health Care Research Committee and the Institutional Review Board at each study site approved our study. Written informed consent was obtained from all participants.

Measurements

Sleep duration measures

SLEEP QUESTIONNAIRE. Participants completed a self-reported questionnaire adapted from the Hispanic Community Health Study/Study of Latinos questionnaire, administered in English or Chinese [13]. Total sleep time or duration was computed by asking participants what time they wake-up and what time they go to bed on weekdays (work/school days) and, separately, weekends (free days). We did not include weekends in this analysis because degree of self-reporting bias or measurement error may differ. For some analyses, sleep duration was categorized into short sleep (<7 hr), recommended sleep (7 and 8 hr), and long sleep (≥9 hr).

Wrist actigraphy.

Participants were asked to wear the Actiwatch Spectrum (PA, USA) on the nondominant wrist for seven consecutive days. A minimum of four (out of a maximum of 5) weekdays were required for analyses other than for single night comparisons between actigraphy and PSG measurements. Actigraphic data were scored in 30 s epochs as sleep or wake using Actiware-Sleep version 5.59 software as previously described by a central Sleep Reading Center (Brigham and Women's Hospital, Boston, MA) by scorers blinded to other data [14]. Total sleep duration was calculated as the sum of epochs scored as sleep in each main sleep interval (manually identified based on a self-actuated event marker, sleep diary, and light sensor) averaged over all days of valid recording. Inter- and intrascorer reliability was assessed; intraclass correlation coefficients for sleep duration exceeded 0.90. The per cent of the sleep period that was spent sleeping (sleep maintenance efficiency, and number as well as duration of awakenings per night) was also calculated. Average time in bed was calculated as the total time between "lights off" and "lights on" (including sleep and wake episodes during the sleep period).

Polysomnography.

PSG was conducted in participants' homes using a 15-channel Type 2 monitor (Somte System; Compumedics Ltd., Abbotsville, Australia). A two-person research team attached the sensors in the late evening, before the participants' usual bedtime. Participants were asked to follow their usual bedtime sleep practices. The recording montage included EEG, bilateral electrooculograms, a chin electromyography (EMG), bipolar electrocardiography (ECG), thoracic and abdominal respiratory inductance plethysmography, airflow measured by thermistry and nasal pressure, finger pulse oximetry, and bilateral limb movements, as previously described [15]. Sleep stages were scored on an epoch by epoch basis using standard scoring criteria centrally, by scorers blinded to all other data [16]. PSG sleep time was defined as the total minutes scored as sleep during the sleep period.

Potential moderators of measurement error/self-reporting bias

RACE/ETHNICITY. Participants self-identified their race/ethnicity as non-Hispanic white, non-Hispanic black/African-American, Hispanic/Latino, or Chinese (hereafter, white, black, Hispanic, or Chinese).

Age group and sex.

We investigated potential differences in whether self-reporting bias differed for participants who were 45–64 and 65–84 years, and between men and women.

Insomnia and sleep apnea.

Insomnia was based on self-report using the Women's Health Initiative Insomnia Rating Scale score of ≥9 (range: 0–20) [17]. From PSG, the Apnea-hypopnea index (AHI) was derived as the sum of the number of obstructive apneas and hypopneas associated with >4 per cent desaturation divided by sleep duration. Obstructive sleep apnea (OSA) was defined using cutoff for "moderate" OSA of AHI ≥ 15 events per hour.

Depression.

Depression was defined as a score of ≥16 on the Center for Epidemiological Studies of Depression (CES-D), a 20-item tool with high internal validity (α = 0.85) [18].

Educational attainment.

Educational attainment was based on a nine-point scale ranging from 0 to 8 and grouped into four educational categories (less than high school, high school, some college [including vocational/technical training], college or greater).

Primary spoken language.

Primary spoken language (Spanish vs. English or Chinese vs. English) of participants was determined through self-report.

Measures of sociodemographic characteristics

Marital status was classified as married/living with partner, divorced/separated/widowed, or never married. Income, based on total gross family income, was categorized as <$25,000, $25,000 to 74,999, and ≥$75,000. Smoking status and lifetime alcohol consumption were classified as "never," "current," or "former." Measured weight and height were used to calculate BMI and categorized as follows: 15–<18.5, 18.5–<25, 25–<30, and ≥30 kg/m^2, to represent underweight, normal, overweight, and obese individuals [19]. We also defined underweight, normal, overweight, and obese categories for Chinese participants as follows: 15–<18.5, 18.5–<23, 23–<25, and ≥25 kg/m^2 based on the International Association for the Study of Obesity and the International Obesity Task Force guidelines for Asians [20]. The prevalence of hypertension (yes vs. no) was defined as systolic blood pressure (BP) ≥ 140 mm Hg and/or diastolic BP ≥ 90 mm Hg or self-reported treatment for hypertension.

Analytic sample

Participants were excluded if they had missing data for sleep duration on the sleep questionnaire, less than 4 weekdays of wrist actigraphy, or missing PSG. Of the 2240 participants, 330 (15%) did not have complete data for the sleep variables and other important covariates, which resulted in a final analytic sample of 1910 participants.

Statistical analysis

Spearman correlation coefficients were calculated between self-reported and measured sleep duration overall and by age and sex groups using multiday wrist actigraphy for total sleep time and average time in bed as well as in-home PSG–measured total sleep time. The error in self-reported sleep duration was calculated as the differences between self-reported and objectively measured sleep duration along with the corresponding 95% confidence interval (CI) for whites, blacks, Hispanics, and Chinese adults. Linear regression models were employed to regress subjective sleep onto the objectively measured sleep for whites, blacks, Hispanics, and Chinese adults, separately. Among each racial/ethnic group separately, we evaluated the model's intercept for potential bias (centering on 7 hr), the regression slope for degree of calibration, and weighted/unweighted κ statistics for the degree of statistical discrimination between the subjective and each objective sleep duration measure. Bias captures the degree to which, on average, participants over- (if intercept > 0) or under-estimate (if <0) sleep. Where there is no bias, the intercept would be zero. The regression slope represents calibration, which is the relationship between the two scales, or the change in the subjective (self-reported sleep duration) relative to change in the objective measure (PSG or wrist actigraphy measured sleep duration). A slope of 1 indicates perfect synchrony between, for example, self-report and PSG: a minute on either scale would equal a minute on an absolute measure, e.g. an atomic clock. Self-report and PSG could also be perfectly calibrated with a slope other than 1, e.g. 0.50: in this case, 1 min in self-reported sleep would equal a half-minute in PSG, and the latter would agree with a half-minute on an atomic clock. In other words, poorly calibrated models have large CIs around the slope. But unlike calibration in a field like physics, there is uncertainty in the calibration of psychological time (self-report), e.g. PSG. The CIs around the slope are larger when there is more uncertainty in the calibration (more variability in the relationship between subjective and objective sleep). When the CIs do

not overlap between two calibration values (e.g. slopes for black participants versus whites), the calibration (relationship of the measures to each other) differs between the two groups. Furthermore, discrimination measures the degree to which individuals with higher objective measures also tend to be those with higher subjective measures, regardless of bias or calibration. It is captured via κ statistics measuring agreement, over and beyond chance, between categories of subjective and objective measures of sleep duration. We considered a less than chance agreement ≤ 0, slight agreement = 0.01–0.20, fair agreement = 0.21– 0.40, moderate agreement = 0.41–0.60, substantial agreement = 0.61–0.80, and almost perfect agreement = 0.81–0.99. We further stratified by a priori variables hypothesized to modify the association between subjective and objective sleep duration, which included sociodemographic (age, sex, educational attainment, and language spoken) and health conditions (depression, insomnia, and sleep apnea) (Supplementary Table 1).

Bland–Altman plots were employed to further analyze the agreement between the subjective and objective measurements. The mean difference represents the estimated bias, and the standard deviation (*SD*) of the differences measures the random fluctuations around the mean.

We tested the differences between Spearman correlations (rho: ρ) using bootstrap estimation. For a given pair of Spearman correlations (e.g. ρ [self-report vs. PSG] for black vs. white participants), we sampled with replacement from the black dataset and white dataset 1000 times, each time calculating the difference in ρ values for black vs. whites, to obtain the standard error of this difference, and thence the probability of the corresponding *z* value based on bootstrap difference and standard error. We also used Levene's test for equality of variances using individual night sleep data from self-report and actigraphy for the 5 weekdays to examine racial differences in nightly sleep duration variability. All tests for significance were two-sided and a *p*-value of <0.05 was considered statistically significant. STATA 14 statistical software (Stata Corporation, College Station, Texas, USA) was used.

Results

Study population characteristics

Table 1 shows sociodemographic, health behavior, and clinical characteristics among 1910 white, black, Hispanic, and Chinese participants. Their mean age was 68.3 ± 9.1 years, and 54 per cent were women. Forty per cent had at least a college education and 26 per cent had an annual household income of <$25,000. Thirty-seven per cent were white, 28 per cent black, 24 per cent Hispanic/Latino, and 11 per cent were Chinese. Hispanics were most likely to have less than a high school education and Chinese participants were the most likely to be married and normal weight. Black participants were the most likely to be widowed, separated, or divorced, obese, and current smokers. Whites were more likely to have a college education or greater and a household income ≥$75,000. The prevalence of moderate or more severe sleep apnea was 33 per cent and insomnia symptoms were 36 per cent.

Table 1.

Sociodemographic, health behavior, and clinical characteristics among 1910 participants of the Multi-Ethnic Study of Atherosclerosis, 2010–2013

	White *n* = 701 (37%)	Black *n* = 539 (28%)	Hispanic *n* = 457 (24%)	Chinese *n* = 213 (11%)	Total *N* = 1910 (100%)
Age (years ± *SD*)	68.4 ± 9.0	68.5 ± 9.0	68.2 ± 9.3	67.6 ± 8.9	68.3 ± 9.1
45–64 years	40	40	41	44	41
65–84 years	60	60	59	56	59
Women	54	55	53	54	54
Marital status†					
Married	66	50	58	82	61
Divorced/separated/widowed	26	41	37	15	32
Never married	8	9	5	3	7
Educational attainment					

	White n = 701 (37%)	Black n = 539 (28%)	Hispanic n = 467 (24%)	Chinese n = 273 (11%)	Total N = 1910 (100%)
Educational attainment					
<High school	3	8	40	17	15
High school graduate	14	16	20	13	16
Some college (including vocational/tech)	25	40	28	23	29
College or greater	58	36	12	47	40
Annual Household income[†]					
<$25,000	12	23	45	41	26
$25,000–74,999	45	52	45	32	46
≥$75,000	43	25	10	27	28
Weekday sleep duration (hr ± SD)					
Self-reported (subjective)[§]	8.0 ± 1.2	7.7 ± 1.8	7.9 ± 1.6	7.8 ± 1.2	7.9 ± 1.5
Actigraphy total sleep time (objective)[‖]	6.8 ± 1.2	6.0 ± 1.4	6.5 ± 1.4	6.3 ± 1.3	6.4 ± 1.4
Actigraphy time in bed (objective)[¶]	7.5 ± 1.3	6.7 ± 1.6	7.3 ± 1.5	7.1 ± 1.4	7.2 ± 1.5
Polysomnography (objective)[#]	6.2 ± 1.3	5.8 ± 1.5	5.9 ± 1.3	6.0 ± 1.4	6.0 ± 1.4
Variability in sleep duration[††]	99.05	117.89	115.23	100.37	111.56
Health behaviors					
Smoking status					
Never	47	48	60	85	55
Current	6	12	4	3	7
Former	47	40	36	12	38
Current alcohol use					
Yes	62	42	31	15	44
BMI categories					
Underweight	1	0	0	3	1
Normal	30	16	15	37	23
Overweight	39	35	42	24	37
Obese	30	49	43	36	39
Clinical characteristics					
Obstructive sleep apnea (yes)[‡‡]	28	32	38	37	33
Insomnia (yes)[§§]	35	38	38	29	36
Depression (yes)[‡]	13	13	20	10	14
Hypertension (yes)	49	72	56	45	57

Data presented as mean ± SD or %; SD = Standard deviation; BMI = Body mass index; tech = Technical.

View Large

The mean self-reported sleep duration on weekdays was 7.9 ± 1.5 hr for all races combined (8.0 ± 1.2 hr for whites, 7.7 ± 1.8 for blacks, 7.9 ± 1.6 for Hispanics/Latinos, and 7.8 ± 1.2 for Chinese participants). Actigraphy-based estimates of weekday sleep duration averaged 6.4 ± 1.4 hr across all groups and are shown in Table 1. Mean sleep duration based on single-night in-home PSG was 6.0 ± 1.4 hr for all races combined (6.2 ± 1.3 hr for whites, 5.8 ± 1.5 for blacks, 5.9 ± 1.3 for Hispanics/Latinos, and 6.0 ± 1.4 for Chinese participants). Self-reports were overestimates in all racial groups compared with PSG, and variability (expressed as SD) in sleep duration (minutes) was greater for all racial/ethnic minorities (117.89 for blacks; 115.23 Latinos; 100.37 Chinese) than whites (99.05).

Figure 1 further illustrates race-specific distributions of weekday sleep duration from self-report, total sleep time and average time in bed from actigraphy, and total sleep time from in-home PSG. Based on graphs, the most apparent racial differences were observed for self-reported sleep duration with whites appearing the most likely to report getting the recommended amounts of sleep (i.e. 7 to 8 hr), and blacks tending to report both shorter and longer sleep durations than other racial/ethnic groups. Overall distributions of weekday sleep duration for all measures are shown in Supplementary Figure 1. Compared with PSG, self-reported measures overestimated sleep duration the most, followed up by actigraphy-based average time in bed and total sleep time. Within race, higher values of sleep duration were observed for self-report compared with wrist actigraphy, whereas even lower values were generally observed for PSG (Supplementary Figure 2). Bland–Altman plots of the systematic difference between self-reported and measured sleep duration showed measurement error patterns that appeared to differ by race and ethnicity (Supplementary Figure 3). Self-reporting bias appeared to increase with increasing average sleep duration from the two measures among blacks and Hispanics.

Figure 1.
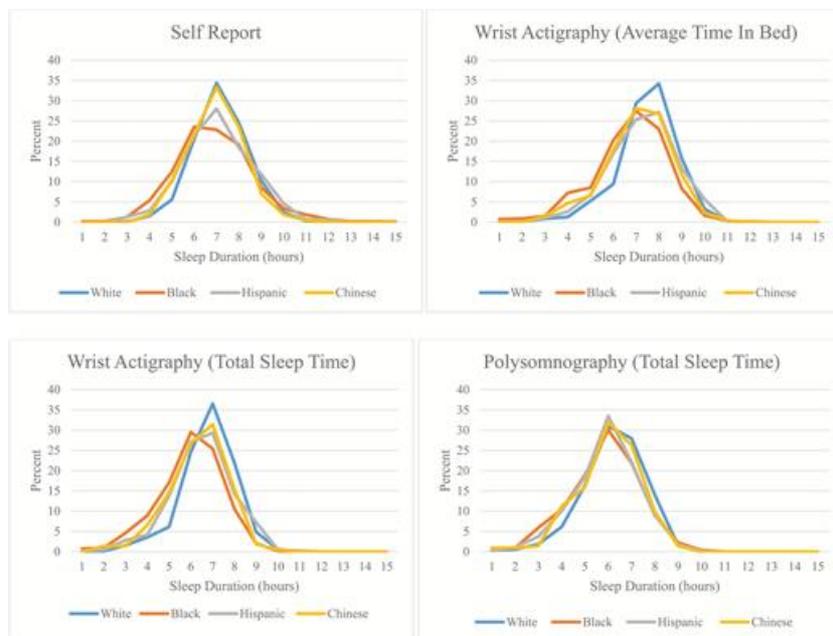
Race-specific distributions of weekday sleep duration from self-report, wrist actigraphy for total sleep time, wrist actigraphy for average time in bed, and in-home polysomnography.

Sleep duration based on self-report vs. multiday wrist actigraphy measuring average weekday total sleep time

Self-reporting bias for total average weekday sleep time based on wrist actigraphy was 66 min (95% CI: 61–71) for whites, 58 min (48–69) for blacks, 66 min (57–74) for Hispanics, and 60 min (49–70) for Chinese adults (Table 2). There were no significant differences in self-reporting bias or measurement error between racial/ethnic groups. Within racial/ethnic groups, Chinese individuals of 45–64 years old had a significantly lower self-reporting bias compared with those 65–84 years (34 min [20–49] vs. 78 min [64–92]); similar age pattern was observed for Hispanics, although nonsignificant. Black men had a lower self-reporting bias than black women (37 min [19–55] vs. 71 min [57–84]).

Table 2.

Unadjusted average differences/bias between self-reported and measured sleep duration (minutes) using wrist actigraphy for both total sleep time and time in bed on weekdays and in-home polysomnography centered to 7 hours of sleep

	Wrist actigraphy (total sleep time [TST])–weekday		Wrist actigraphy (average time in bed [TIB])–weekday		In-home PSG (for total sleep time)–one night (alloyed gold standard)		Between race differences (TST)	Between race differences (TIB)	Between race differences (PSG)
	Inter. (Bias)	Differences/ Calibration (CI)	Inter. (Bias)	Differences/ Calibration (CI)	Inter. (Bias)	Differences/ Calibration (CI)			
Combined	64 (60–68)	0.35 (0.30–0.39)	49 (45–53)	0.33 (0.29–0.38)	64 (59–69)	0.20 (0.19–0.24)	—		—
White	66 (61–71)	0.43 (0.36–0.50)	48 (42–53)	0.43 (0.37–0.49)	73 (67–79)	0.28 (0.21–0.35)	ref	ref	ref
Black	58 (48–69)	0.26 (0.15–0.36)	46 (37–55)	0.22 (0.13–0.31)	54 (42–65)	0.16 (0.06–0.26)	p < 0.001	p < 0.001	p < 0.001
Hispanic	66 (57–74)	0.38 (0.29–0.48)	48 (39–56)	0.39 (0.30–0.48)	67 (56–78)	0.19 (0.09–0.30)	p = 0.001	p = 0.033	p < 0.001

Chinese	Wrist actigraphy (total sleep time [TST])—weekday		Wrist actigraphy (average time in bed [TIB])—weekday		In-home PSG (for total sleep time)—one night (alloyed) gold standard		Between race differences (TST)	Between race differences (TIB)	Between race differences (PSG)
White	Inter. (Bias)	Differences/ Calibration (CI)	Inter. (Bias)	Differences/ Calibration (CI)	Inter. (Bias)	Differences/ Calibration (CI)			
45–64 years	59 (51–67)	0.42 (0.30–0.54)	42 (33–50)	0.42 (0.31–0.53)	63 (55–72)	0.31 (0.21–0.42)	ref	ref	ref
65–84 years	70 (64–77)	0.43 (0.35–0.51)	52 (45–58)	0.43 (0.35–0.51)	80 (72–89)	0.29 (0.20–0.38)	ref	ref	ref
Women	69 (62–76)	0.31 (0.21–0.42)	56 (48–64)	0.32 (0.23–0.42)	73 (65–81)	0.19 (0.09–0.29)	ref	ref	ref
Men	63 (56–71)	0.52 (0.43–0.61)	41 (34–48)	0.51 (0.43–0.59)	74 (65–84)	0.35 (0.26–0.45)	ref	ref	ref
≤HS	72 (60–85)	0.40 (0.22–0.57)	55 (42–69)	0.40 (0.23–0.56)	80 (65–96)	0.21 (0.05–0.37)	ref	ref	ref
>HS	65 (59–70)	0.43 (0.35–0.50)	46 (41–52)	0.43 (0.36–0.50)	72 (65–78)	0.30 (0.22–0.37)	ref	ref	ref
Insomnia (yes)[†]	68 (57–75)	0.32 (0.20–0.44)	52 (43–62)	0.31 (0.20–0.43)	71 (60–81)	0.21 (0.10–0.32)	ref	ref	ref
Insomnia (no)	66 (60–71)	0.50 (0.42–0.58)	45 (38–51)	0.50 (0.42–0.58)	74 (67–82)	0.33 (0.25–0.42)	ref	ref	ref
Sleep apnea (yes)[†]	76 (65–86)	0.39 (0.27–0.52)	59 (49–69)	0.39 (0.28–0.51)	78 (65–91)	0.22 (0.09–0.35)	ref	ref	ref
Sleep apnea (no)	62 (56–67)	0.46 (0.38–0.55)	42 (36–48)	0.46 (0.39–0.54)	71 (64–78)	0.32 (0.24–0.40)	ref	ref	ref
Depression (yes)	77 (60–94)	0.46 (0.25–0.67)	59 (41–76)	0.44 (0.25–0.64)	89 (69–109)	0.35 (0.17–0.53)	ref	ref	ref
Depression (no)	64 (59–69)	0.42 (0.35–0.49)	46 (41–51)	0.43 (0.36–0.49)	70 (64–77)	0.27 (0.19–0.34)	ref	ref	ref
Black									
45–64 years	50 (32–68)	0.31 (0.14–0.48)	36 (22–50)	0.28 (0.14–0.43)	40 (23–58)	0.18 (0.02–0.34)	$p < 0.001$	$p < 0.001$	$p < 0.001$
65–84 years	64 (50–77)	0.21 (0.08–0.34)	54 (42–65)	0.17 (0.05–0.29)	63 (48–78)	0.15 (0.02–0.28)	$p < 0.001$	$p < 0.001$	$p = 0.003$
Women	71 (57–84)	0.28 (0.14–0.42)	58 (46–70)	0.24 (0.11–0.37)	67 (53–81)	0.17 (0.04–0.31)	$p < 0.001$	$p < 0.001$	$p < 0.001$

Men	Wrist actigraphy (total sleep time [TST])–weekday		Wrist actigraphy (average time in bed [TIB])–weekday		In-home PSG (for total sleep time)–one night (alloyed gold standard)		Between race differences (TST)	Between race differences (TIB)	Between race differences (PSG)
	Inter. (Bias)	Differences/ Calibration (CI)	Inter. (Bias)	Differences/ Calibration (CI)	Inter. (Bias)	Differences/ Calibration (CI)			
Men	37 (19–55)	0.18 (0.001–0.32)	29 (13–45)	0.14 (0.001–0.27)	27 (9–47)	0.09 (−0.13–0.20)	p < 0.001	p < 0.001	p = 0.002
≤HS	82 (55–109)	0.41 (0.19–0.64)	63 (40–86)	0.36 (0.16–0.56)	63 (35–94)	0.12 (−0.11–0.35)	p < 0.001	p < 0.001	p = 0.002
>HS	51 (39–62)	0.19 (0.08–0.31)	42 (32–51)	0.16 (0.06–0.27)	52 (40–64)	0.19 (0.08–0.30)	p < 0.001	p < 0.001	p < 0.001
Insomnia (yes)†	54 (35–72)	0.12 (−0.06–0.29)	48 (33–63)	0.11 (-0.05–0.27)	58 (39–77)	0.17 (0.01–0.35)	p < 0.001	p < 0.001	p < 0.001
Insomnia (no)	61 (48–74)	0.34 (0.22–0.47)	45 (34–56)	0.29 (0.17–0.40)	51 (37–66)	0.15 (0.02–0.28)	p < 0.001	p < 0.001	p = 0.014
Sleep apnea (yes)†	59 (37–81)	0.25 (0.06–0.44)	46 (29–64)	0.20 (0.03–0.37)	55 (32–78)	0.17 (−0.02–0.36)	p < 0.001	p < 0.001	p < 0.001
Sleep apnea (no)	58 (46–70)	0.26 (0.14–0.39)	46 (36–56)	0.23 (0.12–0.34)	54 (40–67)	0.15 (0.03–0.27)	p < 0.001	p < 0.001	p = 0.002
Depression (yes)	45 (6–84)	0.11 (−0.21–0.42)	41 (9–72)	0.11 (−0.17–0.40)	40 (3–77)	0.06 (−0.26–0.38)	p < 0.001	p < 0.001	p < 0.001
Depression (no)	60 (49–71)	0.29 (0.18–0.39)	47 (37–56)	0.24 (0.14–0.34)	56 (44–68)	0.18 (0.07–0.28)	p < 0.001	p < 0.001	p = 0.379
Hispanic									
45–64 years	52 (39–66)	0.57 (0.39–0.74)	26 (14–39)	0.55 (0.40–0.71)	47 (31–64)	0.24 (0.07–0.41)	p = 0.005	p = 0.562	p = 0.002
65–84 years	77 (66–88)	0.30 (0.19–0.42)	62 (52–73)	0.31 (0.20–0.41)	83 (68–97)	0.20 (0.07–0.33)	p = 0.043	p = 0.199	p = 0.008
Women	72 (60–83)	0.44 (0.29–0.59)	51 (40–63)	0.45 (0.32–0.59)	74 (61–88)	0.20 (0.06–0.35)	p = 0.085	p = 0.344	p = 0.006
Men	57 (44–71)	0.32 (0.19–0.45)	42 (30–53)	0.32 (0.20–0.44)	54 (35–73)	0.13 (−0.03–0.29)	p = 0.005	p = 0.048	p = 0.004
≤HS	72 (61–82)	0.32 (0.20–0.44)	56 (46–67)	0.32 (0.21–0.43)	78 (64–92)	0.20 (0.07–0.34)	p = 0.053	p = 0.203	p = 0.079
>HS	58 (44–72)	0.48 (0.31–0.64)	36 (24–49)	0.48 (0.34–0.63)	51 (33–68)	0.18 (0.02–0.35)	p < 0.001	p = 0.003	p = 0.005
Insomnia (yes)†	65 (49–81)	0.46 (0.29–0.63)	42 (27–57)	0.47 (0.32–0.63)	73 (54–93)	0.36 (0.18–0.54)	p = 0.043	p = 0.237	p = 0.059
Insomnia (no)	66 (56–76)	0.33 (0.22–0.44)	51 (42–61)	0.32 (0.22–0.43)	61 (49–74)	0.06 (−0.06–0.19)	p = 0.014	p = 0.080	p < 0.001

	Wrist actigraphy (total sleep time [TST])–weekday		Wrist actigraphy (average time in bed [TIB])–weekday		In-home PSG (for total sleep time)–one night (alloyed gold standard)		Between race differences (TST)	Between race differences (TIB)	Between race differences (PSG)
	Inter. (Bias)	Differences/Calibration (CI)	Inter. (Bias)	Differences/Calibration (CI)	Inter. (Bias)	Differences/Calibration (CI)			
Sleep apnea (yes)†	70 (56–84)	0.36 (0.22–0.50)	53 (40–67)	0.35 (0.22–0.48)	75 (65–95)	0.22 (0.04–0.40)	p = 0.314	p = 0.535	p = 0.004
Sleep apnea (no)	63 (52–74)	0.41 (0.28–0.54)	43 (33–54)	0.42 (0.30–0.54)	63 (50–76)	0.19 (0.06–0.32)	p = 0.003	p = 0.052	p = 0.019
Depression (yes)	63 (42–84)	0.44 (0.22–0.65)	42 (22–62)	0.45 (0.26–0.65)	66 (38–94)	0.24 (−0.02–0.50)	p = 0.321	p = 0.655	p = 0.352
Depression (no)	66 (57–76)	0.37 (0.26–0.48)	49 (40–58)	0.36 (0.26–0.46)	67 (55–79)	0.18 (0.07–0.30)	p = 0.002	p = 0.033	p < 0.001
Chinese									
45–64 years	34 (20–49)	0.31 (0.16–0.47)	21 (9–32)	0.29 (0.15–0.44)	20 (5–36)	0.03 (−0.13–0.20)	p < 0.001	p < 0.001	p = 0.038
65–84 years	78 (64–92)	0.21 (0.04–0.37)	68 (56–81)	0.24 (0.10–0.39)	75 (58–92)	0.05 (−0.11–0.20)	p = 0.023	p = 0.082	p = 0.214
Women	62 (49–75)	0.32 (0.16–0.49)	48 (36–61)	0.33 (0.18–0.48)	53 (38–68)	−0.05 (−0.20–0.11)	p = 0.033	p = 0.072	p = 0.050
Men	55 (36–74)	0.23 (0.04–0.42)	44 (29–59)	0.25 (0.09–0.42)	45 (23–66)	0.06 (−0.14–0.26)	p < 0.001	p < 0.001	p = 0.250
≤HS	68 (51–85)	0.23 (0.01–0.44)	57 (40–74)	0.23 (0.02–0.43)	61 (40–81)	−0.05 (−0.28–0.18)	p = 0.182	p = 0.400	p = 0.879
>HS	56 (42–70)	0.29 (0.14–0.44)	43 (32–55)	0.31 (0.18–0.44)	43 (28–59)	0.02 (−0.12–0.17)	p < 0.001	p < 0.001	p = 0.019
Insomnia (yes)†	61 (40–83)	0.19 (−0.03–0.40)	53 (34–71)	0.21 (0.03–0.40)	59 (36–82)	0.11 (−0.10–0.32)	p = 0.006	p = 0.036	p = 0.236
Insomnia (no)	59 (47–71)	0.34 (0.20–0.49)	44 (33–55)	0.35 (0.21–0.48)	45 (30–60)	−0.03 (−0.18–0.12)	p < 0.001	p = 0.003	p = 0.064
Sleep apnea (yes)†	63 (44–82)	0.23 (0.04–0.43)	52 (36–68)	0.24 (0.07–0.41)	56 (31–80)	0.04 (−0.15–0.24)	p = 0.031	p = 0.106	p = 0.003
Sleep apnea (no)	58 (45–71)	0.34 (0.18–0.49)	43 (32–55)	0.35 (0.21–0.48)	47 (33–61)	0.007 (−0.16–0.18)	p < 0.001	p = 0.002	p = 0.869
Depression (yes)	58 (8–109)	0.42 (−0.05–0.89)	39 (1–78)	0.45 (0.08–0.82)	49 (2–97)	0.28 (−0.08–0.63)	p = 0.023	p = 0.083	p = 0.620
Depression (no)	60 (49–71)	0.26 (0.14–0.38)	48 (39–58)	0.26 (0.15–0.38)	47 (34–60)	−0.06 (−0.19–0.07)	p < 0.001	p = 0.001	p = 0.029

Inter. = Intercept; CI = Confidence interval; PSG = Polysomnography; Difference of 0 = no bias or measurement error; whites in

subgroups are the reference group for other racial/ethnic subgroups; Act = Actigraphy; TST = Total sleep time; ≤HS = High school education or less; >HS = More than high school education; Depression = CES-D ≥16.

[†]OSA was measured using one-day in-home polysomnography (number of obstructive apneas and hypopneas per hour of sleep, including events with absent airflow by thermistry but present effort by inductance belts); OSA was defined as an AHI ≥ 15 events per hour; Insomnia was based self-report using a Women's Health Initiative Insomnia Index score of ≥9 (range: 0–20).

Statistically significant difference by race based on confidence intervals. All within-race comparisons between self-reported and measured sleep duration (in minutes) using wrist actigraphy for both total sleep time and time in bed on weekdays and in-home polysomnography were statistically significant at $p \leq 0.01$, with self-reported sleep duration being consistently longer than objective sleep measures. Bias captures the degree to which, on average, participants over (if intercept > 0) or underestimate (if <0) sleep. Where there is no bias, the intercept would be zero. The regression slope represents calibration, which is the relationship between the two scales, or the change in the subjective (self-reported sleep duration) relative to change in the objective measure (polysomnography or wrist actigraphy measured sleep duration).

Sleep duration reporting bias by potential moderators: Table 2 shows unadjusted average differences/bias between self-reported and measured sleep duration (in minutes) using multiday wrist actigraphy for both total sleep time and average time in bed and PSG by a priori sociodemographic (e.g. education, primary language spoken) and health-related (e.g. insomnia, sleep apnea, depression) characteristics. Compared with actigraphy-based total sleep time, a higher self-reporting error among both whites with sleep apnea than those without sleep apnea. However, there were no significant differences in the degree of self-reporting error by status of insomnia, sleep apnea, depression, and educational attainment in all racial/ethnic groups (Table 2). Compared with their white counterparts, measurement error was significantly lower among blacks (with the exception of no significant difference among individuals with depression). Reporting bias was lower among Hispanics compared with whites (except where there were no error differences among individuals without either insomnia, depression, or sleep apnea, and who had greater than a high school education). Lastly, measurement error was lower among Chinese adults with sleep apnea, no depression, and a high school education or greater.

View Large

Spearman correlations based on self-report vs. wrist actigraphy measured total sleep time were $\rho = 0.38$ overall, with correlations significantly lower only among blacks compared with whites ($\rho = 0.45$ for whites, $\rho = 0.28$ blacks, $\rho = 0.38$ Hispanics/Latinos, and $\rho = 0.35$ Chinese) (Table 3). Therefore, we found that 20 per cent of the variation in self-reported sleep duration was explained by actigraphy-estimated sleep among whites, 8 per cent among blacks, 14 per cent among Hispanics, and 12 per cent among Chinese participants. Correlations between self-reported and actigraphy-measured sleep duration were similar across strata of age and sex and were not significantly different for total sleep time as measured by wrist actigraphy.

Table 3.

Spearman correlations between self-reported and examined sleep duration using wrist actigraphy for total sleep time, wrist actigraphy for time in bed, and in-home polysomnography for total sleep time by race/ethnicity, gender, and age

	White N = 701 (37%)	Black n = 539 (28%)	Hispanic n = 457 (24%)	Chinese n = 213 (11%)	Total N = 1910 (100%)
	Wrist actigraphy (total sleep time)				
Total	0.45[B]	0.28[W]	0.38	0.35	0.38
Age					
45–64 years	0.41	0.35	0.45	0.45	0.43
65–84 years	0.47[BC]	0.23[W]	0.34	0.23[W]	0.35
Women	0.41[B]	0.25[WH]	0.43[B]	0.37	0.36
45–64 years	0.45	0.27	0.36	0.51	0.38
65–84 years	0.37	0.24[H]	0.46[B]	0.30	0.34
Men	0.50[BHC]	0.25[W]	0.31[W]	0.30[W]	0.38

	White N = 701 (37%)	Black n = 539 (28%)	Hispanic n = 457 (24%)	Chinese n = 213 (11%)	Total N = 1910 (100%)
45–64 years	0.34[O]	0.40	0.49[O]	0.21	0.45
65–84 years	0.56[BH,Y]	0.16[W]	0.22[W,Y]	0.15[W]	0.34
Wrist actigraphy (total sleep time) Wrist actigraphy (average time in bed)					
Total	0.49[B]	0.28[WH]	0.41[B]	0.39	0.40
Age					
45–64 years	0.45	0.36	0.46	0.47	0.45
65–84 years	0.50[BC]	0.22[W]	0.37	0.29[W]	0.37
Women	0.44[B]	0.25[WH]	0.45[B]	0.40	0.38
45–64 years	0.50	0.27	0.39	0.51	0.40
65–84 years	0.39	0.23[H]	0.48[B]	0.35	0.36
Men	0.52[BH]	0.25[W]	0.34[W]	0.36	0.40
45–64 years	0.37[O]	0.43[O]	0.53[O]	0.24	0.47
65–84 years	0.59[BHC,Y]	0.14[W,Y]	0.25[W,Y]	0.20[W]	0.36
In-home polysomnography (total sleep time)					
Total	0.31[BHC]	0.15[WC]	0.16[WC]	−0.02[WBH]	0.20
Age					
45–64 years	0.37[BC]	0.19[W]	0.21	0.10[W]	0.26
65–84 years	0.31[BC]	0.14[W]	0.17	−0.03[W]	0.19
Women	0.25[C]	0.14	0.21[C]	−0.08[WH]	0.17
45–64 years	0.37[C]	0.15	0.20	−0.03[W]	0.23
65–84 years	0.18	0.15	0.25[C]	−0.08[H]	0.16
Men	0.33[BHC]	0.09[W]	0.08[W]	0.01[W]	0.17
45–64 years	0.29	0.17	0.18	0.13	0.22
65–84 years	0.39[BH]	0.05[W]	0.07[W]	0.13	0.18

Perfect correlation = 1; Strong = 0.99–0.70; Moderate = 0.69–0.30; Weak = 0.29–0.01; No correlation = 0.

[W,B,H,C] Letters indicate significant differences between the races/ethnicities: W = White; B = Black; H = Hispanic; C = Chinese.

[Y,O] Letters indicate significant differences between age groups within race: Y = Young (45–64 years); O = Old (65–84).

Male vs. Female comparisons were all $p \geq 0.05$; comparisons of Spearman correlations between total sleep time using wrist actigraphy and examined sleep duration (PSG)—within race by gender and age: White, Male vs. Female—$p = 0.929$; 45–64 years vs. 65–84 years —$p = 0.515$; Black, Male vs. Female—$p = 0.232$; 45–64 years vs. 65–84 years—$p = 0.406$; Hispanic, Male vs. Female—$p = 0.766$; 45–64 years vs. 65–84 years—$p = 0.267$; Chinese, Male vs. Female—$p = 0.839$; 45–64 years vs. 65–84 years—$p = 0.266$.

View Large

As shown in Table 4, across all races, the weighted κ statistics between self-report sleep duration categories and sleep time measured by actigraphy indicated slight agreement $(k_w = 0.14$ for whites, 0.10 for blacks, 0.17 for Hispanics, and 0.11 for Chinese).

Table 4.

Agreement between sleep duration based on self-reported and wrist actigraphy for total sleep time on weekdays as the alloyed gold standard by race

Whites	Wrist actigraphy (total sleep time)			
Self-reported sleep duration		Short sleep	Recommended sleep	Long sleep	Total
	Short sleep	89 (86)	14 (14)	0	151
	Recommended sleep	240 (56)	185 (43)	3 (1)	538
	Long sleep	50 (29)	111 (65)	9 (5)	12
	Total	379 (54)	310 (44)	12 (2)	701
	Agreement: 40.37%	Expected agreement: 35.36%		κ: 0.0775	$p \leq 0.001$
	(w) Agreement: 66.62%	(w) Expected agreement: 61.00%		(w) κ: 0.1441	(w) $p \leq 0.001$

Blacks	Wrist actigraphy (total sleep time)			
Self-reported sleep duration		Short sleep	Recommended sleep	Long sleep	Total
	Short sleep	146 (91)	13 (8)	2 (1)	241
	Recommended sleep	184 (75)	61 (25)	0 (0)	294
	Long sleep	82 (62)	49 (37)	2 (2)	4
	Total	412 (76)	123 (23)	4 (1)	539
	Agreement: 38.78%	Expected agreement: 33.39%		κ: 0.0809	$p \leq 0.001$
	(w) Agreement: 61.60%	(w) Expected agreement: 57.15%		(w) κ: 0.1037	(w) $p \leq 0.001$

Hispanics	Wrist actigraphy (total sleep time)			
Self-reported sleep duration		Short sleep	Recommended sleep	Long sleep	Total
	Short sleep	91 (89)	11 (11)	0 (0)	150
	Recommended sleep	148 (66)	73 (33)	3 (1)	292
	Long sleep	53 (40)	66 (50)	12 (9)	15
	Total	292 (64)	150 (33)	15 (3)	457
	Agreement: 38.51%	Expected agreement: 31.29%		κ: 0.1051	$p \leq 0.001$
	(w) Agreement: 63.46%	(w) Expected agreement: 56.12%		(w) κ: 0.1672	(w) $p \leq 0.001$

Chinese	Wrist actigraphy (total sleep time)			
Self-reported sleep duration		Short sleep	Recommended sleep	Long sleep	Total
	Short sleep	37 (86)	6 (14)	0 (0)	80
	Recommended sleep	84 (67)	42 (33)	0 (0)	131
	Long sleep	24 (55)	18 (41)	2 (5)	2
	Total	145 (68)	66 (31)	2 (1)	213
	Agreement: 38.03%	Expected agreement: 32.27%		κ: 0.0851	$p = 0.0123$
	(w) Agreement: 63.38%	(w) Expected agreement: 59.01%		(w) κ: 0.1067	(w) $p = 0.001$

Short sleep: <7 hr; Recommended sleep: ≥7–<9 hr; Long sleep: ≥9 hr. *n*(%) = number (percentage) of study participants; κ statistic interpretation: Less than chance agreement ≤ 0; Slight agreement = 0.01–0.20; Fair agreement = 0.21–0.40; Moderate agreement = 0.41–0.60; Substantial agreement = 0.61–0.80; Almost perfect agreement = 0.81–0.99.

View Large

Sleep duration based on self-report vs. multiday wrist actigraphy measuring average weekday time in bed

Self-reporting bias for sleep duration compared with actigraphy-based average time in bed was 48 min (42–53) for whites, 46 min (37–55) for blacks, 48 min (39–56) for Hispanics, and 47 min (38–56) for Chinese adults (Table 3). Older Hispanics (62 min [52–73] vs. 26 min [14–39]) and older Chinese (68 min [56–81] vs. 21 min [9–32]) had a significantly higher self-reporting bias compared with their younger racial counterparts. Furthermore, black women had more self-reporting bias than black men (58 min [46–70] vs. 29 min [15–43]), with similar, albeit nonsignificant sex differences seen for whites. Spearman correlations between sleep duration based on self-report vs. wrist actigraphy average time in bed were ρ = 0.40 overall, and the correlation for blacks was significantly lower than whites and Hispanics (ρ = 0.49 for whites, ρ = 0.28 blacks, ρ = 0.41 Hispanics, and ρ = 0.39 Chinese) (Table 3). Spearman correlations between self-report and actigraphy-based average time in bed were higher for white older than younger men (0.59 vs. 0.37, $p < 0.05$) and white younger than older women (0.50 vs. 0.39, $p > 0.05$). Both black (0.43 vs. 0.14, $p < 0.05$) and Hispanic (0.53 vs. 0.25, $p < 0.05$) men who were younger had higher correlations than their older racial counterparts. There was fair agreement between sleep duration categories based on self-report and actigraphy-measured average time in bed for whites (k_w = 0.27), blacks (k_w = 0.24), Hispanics (k_w = 0.27), and Chinese (k_w = 0.24) participants, with higher κ values than for self-report and actigraphy sleep duration as summarized earlier (Supplementary Table 2). With racial groups combined, calibration between self-reported sleep and average time in bed (0.33 [0.29–0.38]) indicated that a change of 1 min in self-reported sleep time resulted in 20 s (0.33 × 60 s [1min]) of actigraphy average time in bed (Table 2; also, Supplementary Table 3 tests the differences in calibration [slope] by race). For average time in bed using actigraphy, calibration with self-report was lower among black participants (0.22 [0.13–0.31]) than for other groups, meaning that their incremental increases in self-reported sleep duration were associated with lower increases in sleep duration estimates based on actigraphy in comparison to other racial groups.

Sleep duration based on self-report vs. single-night in-home PSG

Self-reporting biases using PSG as the standard showed that compared with whites at 73 min (67–79), self-reporting biases were significantly lower among blacks (54 min [42–65]) and Chinese (49 min [37–61]) but not significantly different for Hispanics (67 min [56–78]) (Table 2). For total sleep time based on PSG, younger Hispanics (47 min [31–64] vs. 83 min [68–97]) and younger Chinese (20 min [5–36] vs. 75 min [58–92]) had a significantly lower self-reporting bias. With regard to sex differences, black women had a higher self-reporting bias than black men (67 min [53–81] vs. 27 min [6–47]). Spearman correlations based on self-report vs. in-home PSG measured total sleep time were weak at ρ = 0.20 for the total sample of participants (ρ = 0.31 for whites, ρ = 0.15 blacks, ρ = 0.16 Hispanics, and ρ = −0.02 for Chinese). There was slight agreement between sleep duration categories based on self-report and single-night in home PSG (k_w = 0.08 for whites, 0.04 for blacks, 0.05 for Hispanics, and 0.01 for Chinese) (Supplementary Table 4), and most races had fair agreement (k_w = 0.33 for whites, 0.28 for blacks, 0.27 for Hispanics, and 0.24 for Chinese) with overestimates of sleep duration across all racial groups comparing wrist actigraphy for total sleep time with one-night of in-home PSG tested using data from the same night (Supplementary Table 5). With race groups combined, calibration between self-reported sleep and PSG (0.20 [0.19–0.24]) indicated that a change of 1 min in self-reported sleep time resulted in only 12 s (0.20 × 60 s [1 min]) of measured PSG time. (Table 2). The calibration between PSG and self-reported sleep was significantly different from the calibration between self-reported sleep and actigraphy for both total sleep time (0.35 [0.30–0.39]) and average time in bed (0.33 [0.29–0.38]). Chinese participants had a calibration of 0.14 (−0.11 to 0.13) for self-report versus PSG. The CI included zero, which indicates that there was no relationship between the two scales.

Sleep duration based on one same-night wrist actigraphy measuring total sleep time vs. concurrent in-home PSG

We compared self-reporting biases of two separate objective measures of sleep duration (wrist actigraphy and PSG), each conducted the same night, across various race/ethnicities (Supplementary Table 5). The overall level of agreement was fair (k_w = 0.29) with a mean over estimation of total sleep time by actigraphy compared with PSG of 45 min (95% CI: 40–50). Discrepancies were largely explained by underestimated wake after sleep time as measured by actigraphy (data not shown).

Comparisons across subjective and objective measures

Self-reported sleep duration compared with PSG generally had greater reporting bias while self-report compared with average time in bed by actigraphy had the least. There was no racial difference in self-reporting error compared with total sleep time and average time in bed from multiday wrist actigraphy (*p* values not shown), but reporting bias differed by race/ethnicity for self-reports compared with PSG. Compared with whites (ρ = 0.45), correlations were significantly lower in blacks (ρ = 0.28) but not in Hispanics and Chinese participants. With race groups combined, the calibration between PSG and self-reported sleep (0.20 [0.19–0.24]) was significantly different from the calibration between actigraphy and self-reported sleep for both total sleep time (0.35 [0.30–0.39]) and average time in bed (0.33 [0.29–0.38]). The self-reported sleep minutes were longer relative to PSG compared with actigraphy. In other words, a change of 1 self-reported minute resulted in only 12 s (0.20 × 60 s [1 min]) of PSG time, whereas a change of 1 self-reported minute resulted in only 20–21 s of actigraphy time (0.35 for total sleep time actigraphy [0.35 × 60 s [1 min]) and 0.33 for average time in bed actigraphy (0.33 × 60 s [1 min]).

Discussion

Although self-reported sleep duration is most commonly used in epidemiological studies and used to infer population differences in sleep, little prior research has addressed the extent to which these measures may be biased or misclassify sleep duration. In this comprehensive analysis of self-reported and objectively measured weekday sleep duration across four racial/ethnic groups, we identified several important overall and race-specific patterns. Overall, self-reported sleep duration, as estimated by the difference between bed and wake times, only moderately correlated (ρ = 0.38) with multiday actigraphy-based assessments, with actigraphy-based sleep explaining only 14 per cent of the variance in self-reported sleep duration. Self-reported sleep correlated even less to sleep measured by single-night PSG (ρ = 0.20). Furthermore, self-reported sleep duration overestimated sleep duration by an average of 64 min when compared with actigraphy-measured total sleep time, 49 min for actigraphy-based time in bed, and 64 min for PSG-defined sleep time. All racial/ethnic groups overestimated their sleep duration when comparing self-reports to all objective measures and had κ statistics that corresponded with only slight-to-fair agreement across the respective comparators. Racial differences of systematic bias in sleep estimation, however, were only observed for self-reported compared with PSG-measured sleep duration (greater for whites compared with blacks and Chinese). However, significantly higher correlations for self-reported sleep duration and objectively measured sleep as well as better calibration (for some measures) among whites compared with blacks suggest a greater potential for random misclassification of sleep in blacks compared with whites.

Furthermore, correlations between actigraphy- and PSG-measured sleep duration were similar across strata of age and sex and were not significantly different for total sleep time as measured by wrist actigraphy and PSG. With regard to sex differences in self-reporting bias based on PSG, black women had a higher self-reporting bias than black men. We observed a higher self-reporting bias among whites with sleep apnea than those without sleep apnea, possibly reflecting low sleep efficiency in this group leading to sleep misperception, which previously has been suggested to result in overestimation of sleep time [21]. It is unclear why this influence would be more apparent in whites. Contrary to expectations, no significant differences in the degree of self-reporting bias were observed according to insomnia, depression, or educational attainment across the racial/ethnic groups. We also observed that Chinese participants speaking Chinese as the primary language compared with those speaking English also overestimated their sleep time, suggesting the potential influence of culture or language on sleep self-report.

This study also provides data on the association between actigraphy and PSG-derived total sleep time in a large sample with concurrent measurements of each (i.e. restricted to the same night of monitoring). These results indicate that there is fair agreement between these two objective measures with a mean bias of 45 min (95% CI: 40–50). Prior studies of only men or women using a different actigraph from the one used in this study showed fair-to-moderate agreement, with bias differing according to subgroups examined and how the actigraphy data were analyzed [22, 23]. These results suggest caution in defining thresholds for identifying "short" and "long" sleep, and need to consider the measurement approach for estimating sleep duration when defining "risky" sleep duration thresholds.

Other studies investigating concordance, with mostly white participants, have also found moderate correlations between self-reported and actigraphy-measured sleep duration, which ranged from 0.31 to 0.47 [8, 9, 24]. The few studies that investigated racial and ethnic differences were consistent with our findings. For instance, Coronary Artery Risk Development in Young Adults study participants with a mean age of 42.9 years, 58 per cent women and 44 per cent blacks had a correlation of 0.47 (0.29 for blacks; 0.56 whites) and mean difference of 1.2 hr between actigraphy-measured sleep duration and sleep duration measured using the question "How many hours of sleep do you usually get at night (or when you usually sleep)?" [5]. With a comparable mean age of 67 years among 53 per cent women and 75 per cent whites, the Sleep Heart Health Study participants reported a 1 hr 1 min difference between in-home PSG and self-report sleep duration based on differences between bed and wake times, with a weak correlation of 0.18 [8]. Cespedes et al. found a 1 hr and 12 min difference between actigraphy and self-report and a correlation of 0.43 among

Hispanics [12].

Furthermore, few studies have considered how self-reported sleep compares with both actigraphy-based sleep and PSG-measured sleep duration. Although single-night PSG may not provide representative data on sleep duration due to night-to-night variability [25], PSG is still considered the gold-standard method for neurophysiological measurement of sleep. Several epidemiological studies have shown that PSG-measured sleep duration significantly associated with hypertension [1] and diabetes [26]. Our findings indicate that data from self-report correspond only weakly to PSG-measured sleep duration, and only moderately with 5-day actigraphy measurements, underscoring the need to very cautiously extrapolate findings from studies using different measurement instruments. Moreover, the difficulty in translating public health recommendations from studies using actigraphy or PSG should be considered within the context that what is readily understood by the public is their self-perceived sleep duration, which better reflects time in bed than actual sleep duration measured by either actigraphy or PSG.

Overall, objective measurements explained only a modest proportion of the variation in self-reported measurements, with a smaller proportion of the variation in self-reported sleep explained by actigraphy in blacks compared with whites. There are several potential explanations. The greater between night variability in sleep duration observed among racial/ethnic minorities may make it more difficult to accurately report habitual sleep duration, which requires estimations of average sleep duration. There may be racial/ethnic differences in actual baseline health status and sleep health literacy in that whites might be more likely than other groups to know the recommended amount of sleep and more likely to provide socially/medically desirable responses. This could result in the higher (albeit still only moderate) correlations but higher bias between subjective and objective measures among whites. There may also be systematic differences—for example, due to culture—in interpreting, conceptualizing, or summarizing the various sleep dimensions (duration; satisfaction; alertness; timing; efficiency) in order to respond to seemingly simple questions like habitual sleep duration [27]. There are notable challenges in answering questions on habitual sleep duration, as it is difficult to know how long it takes to fall asleep or estimate periods of wake after sleep onset. Given that our assessment of self-reported sleep duration used questions regarding bed and wake times, the group differences we observed likely reflect differences in how individuals perceive and report wake after sleep onset rather than sleep latency. Investigating the psychometric properties of sleep instruments across racial/ethnic groups, including evaluating the impact of alternative approaches for collected self-reported data, represents an important area for future research. For instance, if whites overestimate their sleep duration more than blacks, this may lead to an exaggeration of racial differences in sleep duration when relying only on questionnaires. For example, questionnaires, when compared with actigraphy, may misclassify racial differences in sleep duration by 10 to 20 min per night. However, misclassification among minorities may underestimate racial differences in true associations with health outcomes.

This study has several limitations. We had a small sample of participants (especially for Chinese-American participants) once we stratified the sample by race/ethnicity; therefore, results should be interpreted with caution. There were also a large number of tests for statistical significance; so, having a few subgroup comparisons (e.g. age within sex and race) statistically significant is not surprising. For these reasons, results should be interpreted with caution. Our participants were also middle aged or older, and these results could differ in younger age groups. Furthermore, there is great variation in how sleep duration can be assessed as there is no standard, agreed upon way to ask participants about their habitual sleep duration. Simple questions asking for usual sleep time can be limited by math errors made in subtracting bed times from wake times, reporting biases (towards desired durations), and to rounding errors, reducing precision. All questions of sleep duration are limited by problems in identifying actual sleep onset. Although people generally know the time they went to bed, they are poor at estimating sleep latency. Consistent with this observation, we found less difference between self-report and actigraphy-based average time in bed compared with actigraphy-based sleep duration. Nonetheless, actigraphy also is subject to misclassification, with waking rest and sleep difficult to distinguish, with accuracy varying by actigraph model and analysis mode [28].

Despite these limitations, important strengths exist and this study extends beyond the prior studies in several important ways. The racial/ethnic diversity of the cohort is a major strength of the study. We compared self-reporting error and bias by race, which has important implications for the health disparities literature as it is important to distinguish true differences from biases related to measurement. We also compared the self-reporting biases of two separate objective measures of sleep duration (wrist actigraphy and PSG) across various race/ethnicities. In addition to having a fairly large sample size with the ability to stratify by various factors, we had objective data on sleep apnea and validated insomnia symptoms.

Our findings inform future population health and health disparities research aimed at understanding the relationship to sleep across racial/ethnic groups. Overall, self-reported sleep duration overestimates objectively measured sleep across all race and minority groups studied, and compared with PSG, overestimation is significantly greater in whites compared with blacks. Furthermore, self-reported sleep duration less accurately estimates objective sleep in blacks compared with whites; larger reporting bias and poorer calibration of estimates would reduce the ability to identify significant associations between sleep duration and health among blacks compared with whites. More research is needed to improve the measurement of sleep across the population.

Supplementary Material

Supplementary material is available at *SLEEP* online.

Funding

Notes

Author Contributions

Authors: Chandra L. Jackson, Sanjay R. Patel, Braxton Jackson, Pamela L. Lutsey, and Susan Redline. *Study concept and design*: C. Jackson and Redline. *Acquisition of data*: C. Jackson and Redline. *Statistical analysis*: C. Jackson and B. Jackson. *Interpretation of data*: C. Jackson, Patel, B. Jackson, Lutsey, and Redline. *Drafting of the manuscript*: C. Jackson. *Critical revision of the manuscript for important intellectual content*: C. Jackson, Patel, B. Jackson, Lutsey, Redline. *Administrative, technical, and material support:* C. Jackson and Redline. *Obtaining funding and study supervision*: C. Jackson and Redline. *Final approval*: C. Jackson, Patel, B. Jackson, Lutsey, and Redline.

References

1. Gottlieb DJ , et al. Association of sleep time with diabetes mellitus and impaired glucose tolerance. *Arch Intern Med* . 2005;165(8):863–867.

   Google Scholar   Crossref   PubMed

2. Ayas NT , et al. A prospective study of self-reported sleep duration and incident diabetes in women. *Diabetes Care* . 2003;26(2):380–384.

   Google Scholar   Crossref   PubMed

3. Ayas NT , et al. A prospective study of sleep duration and coronary heart disease in women. *Arch Intern Med* . 2003;163(2):205–209.

   Google Scholar   Crossref   PubMed

4. Grandner MA , et al. Who gets the best sleep? Ethnic and socioeconomic factors related to sleep complaints. *Sleep Med* . 2010;11(5):470–478.

   Google Scholar   Crossref   PubMed

5. Lauderdale DS , et al. Self-reported and measured sleep duration: how similar are they?*Epidemiology* .

2008;19(6):838–845.

Google Scholar   Crossref   PubMed

6. Lauderdale DS , et al. Cross-sectional and longitudinal associations between objectively measured sleep duration and body mass index: the CARDIA sleep study. *Am J Epidemiol* . 2009;170(7):805–813.

Google Scholar   Crossref   PubMed

7. Lauderdale DS , et al. Objectively measured sleep characteristics among early-middle-aged adults: the CARDIA study. *Am J Epidemiol* . 2006;164(1):5–16.

Google Scholar   Crossref   PubMed

8. Silva GE , et al. Relationship between reported and measured sleep times: the sleep heart health study (SHHS). *J Clin Sleep Med* . 2007;3(6):622–630.

Google Scholar   PubMed

9. Van Den Berg JF , 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.

Google Scholar   Crossref   PubMed

10. Unruh ML , et al. Subjective and objective sleep quality and aging in the sleep heart health study. *J Am Geriatr Soc* . 2008;56(7):1218–1227.

Google Scholar   Crossref   PubMed

11. McCrae CS , et al. Sleep complaints, subjective and objective sleep patterns, health, psychological adjustment, and daytime functioning in community-dwelling older adults. *J Gerontol B Psychol Sci Soc Sci* . 2005;60(4):P182–P189.

Google Scholar   Crossref   PubMed

12. Cespedes EM , et al. Comparison of self-reported sleep duration with actigraphy: results from the Hispanic Community Health Study/Study of Latinos Sueño Ancillary Study. *Am J Epidemiol* . 2016;183(6):561–573.

Google Scholar   Crossref   PubMed

13. Redline S , et al. Sleep-disordered breathing in hispanic/latino individuals of diverse backgrounds. The Hispanic Community Health Study/Study of Latinos. *Am J Respir Crit Care Med* . 2014;189(3):335–344.

Google Scholar   Crossref   PubMed

14. Oakley N. *Validation With Polysomnography of the Sleepwatch Sleep/Wake Scoring Algorithm Used by the Actiwatch Activity Monitoring System* . Bend: Mini Mitter, Cambridge Neurotechnology; 1997.

15. Chen X , et al. Racial/ethnic differences in sleep disturbances: the multi-ethnic study of atherosclerosis (MESA). *Sleep* . 2015;38(6):877–888.

Google Scholar   PubMed

16. Rechtschaffen A , et al. *A Manual of Standardized Terminology, Techniques, and Scoring System for Sleep Stages of Human Subjects* . Washington DC: Washington Public Health Service, US Government Printing Office; 1968.

17. Levine DW , et al. Reliability and validity of the women's health initiative insomnia rating scale. *Psychol Assess* . 2003;15(2):137–148.

Google Scholar   Crossref   PubMed

18. Radloff LS . The use of the center for epidemiologic studies depression scale in adolescents and young adults. *J Youth Adolesc* . 1991;20(2):149–166.

Google Scholar   Crossref   PubMed

19. National Institutes of Health. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: the evidence report. *Obes Res* . 1998;6:51S–209S.

Crossref   PubMed

20. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet* . 2004;363:157–163.

Crossref   PubMed

21. Lockley SW , et al. Comparison between subjective and actigraphic measurement of sleep and sleep rhythms. *J Sleep Res* . 1999;8(3):175–183.

Google Scholar   Crossref   PubMed

22. Blackwell T , et al. ; Osteoporotic Fractures in Men (MrOS) Study Group. Factors that may influence the classification of sleep-wake by wrist actigraphy: the MrOS sleep study. *J Clin Sleep Med* . 2011;7(4):357–367.

Google Scholar   PubMed

23. Dean DA , et al. A systematic assessment of the association of polysomnographic indices with blood pressure: the multi-ethnic study of atherosclerosis (MESA). *Sleep* . 2015;38(4):587–596.

Google Scholar   Crossref   PubMed

24. Girschik J , et al. Validation of self-reported sleep against actigraphy. *J Epidemiol* . 2012;22(5):462–468.

Google Scholar   Crossref   PubMed

25. Quan SF , et al. ; Sleep Heart Health Study (SHHS) Research Group. Short-term variability of respiration and sleep during unattended nonlaboratory polysomnography–the Sleep Heart Health Study. [corrected]. *Sleep* . 2002;25(8):843–849.

Google Scholar   Crossref   PubMed

26. Vgontzas AN , et al. Insomnia with objective short sleep duration is associated with type 2 diabetes: a population-based study. *Diabetes Care* . 2009;32(11):1980–1985.

Google Scholar   Crossref   PubMed

27. Buysse DJ . Sleep health: can we define it? Does it matter?*Sleep* . 2014;37(1):9–17.

Google Scholar   Crossref   PubMed

28. Mitchell JA , et al. Variation in actigraphy-estimated rest-activity patterns by demographic factors. *Chronobiol Int* . 2017;34(8):1042–1056.

Google Scholar   Crossref   PubMed

Topic:

adult

blacks

Issue Section: Sleep, Health and Disease

- Supplementary data

Supplementary data

Supplementary Tables Figures   - docx file

8   View Metrics

Email alerts

New issue alert
Advance article alerts
Article activity alert
Subject alert

Receive exclusive offers and
updates from Oxford Academic

More on this topic

Interpreting Wrist Actigraphic Indices of Sleep
in Epidemiologic Studies of the Elderly: The
Study of Osteoporotic Fractures

Wrist Actigraphy in Insomnia

Associations between polysomnography and

actigraphy-based sleep indices and glycemic control among those with and without type 2 diabetes: the Multi-Ethnic Study of Atherosclerosis

Racial/Ethnic Differences in Sleep Disturbances: The Multi-Ethnic Study of Atherosclerosis (MESA)

Related articles in

Web of Science

Google Scholar

Related articles in PubMed

Autism and developmental disability caused by KCNQ3 gain-of-function variants.

Cerebral arterial gas embolism in a scuba diver with a primary lung bulla.

Update on hereditary, autosomal dominant cathepsin-A-related arteriopathy with strokes and leukoencephalopathy (CARASAL).

Factors Influencing Sleep Difficulty and Sleep Quantity in the Citizen Pscientist Psoriatic Cohort.

Citing articles via

Web of Science (9)

Google Scholar

CrossRef

Latest | **Most Read** | **Most Cited**

Characterization of the sleep disorder of anti-IgLON5 disease

Actigraphic detection of periodic limb movements: development and validation of a potential device-independent algorithm. A proof of concept study

Simultaneous tonic and phasic REM sleep without atonia best predicts early phenoconversion to neurodegenerative disease in idiopathic REM sleep behavior disorder

Residual symptoms after natural remission of insomnia: associations with relapse over 4 years

Sleep duration and fragmentation in relation to

leukocyte DNA methylation in adolescents

About SLEEP

Editorial Board

Author Guidelines

Facebook

Twitter

Contact Us

Purchase

Recommend to your Library

Advertising and Corporate Services

Journals Career Network

About Us

Contact Us

Careers

Help

Access & Purchase

Rights & Permissions

Open Access

Connect

Join Our Mailing List

OUPblog

Twitter

Facebook

YouTube

Tumblr

Resources

Authors

Librarians

Societies

Sponsors & Advertisers

Explore

Shop OUP Academic

Oxford Dictionaries

Oxford Index

Epigeum

Press & Media

OUP Worldwide

Agents

University of Oxford

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide