

Factors that May Influence the Classification of Sleep-Wake by Wrist Actigraphy: The MrOS Sleep Study

Terri Blackwell, M.A.¹; Sonia Ancoli-Israel, Ph.D., F.A.A.S.M.²; Susan Redline, M.D., M.P.H.³; Katie L. Stone, Ph.D.¹; for the Osteoporotic Fractures in Men (MrOS) Study Group

¹California Pacific Medical Center Research Institute, San Francisco, CA; ²Department of Psychiatry, University of California, San Diego, La Jolla, CA; ³Departments of Medicine, Brigham and Women's Hospital and Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

Study Objectives: Total sleep time (TST), sleep efficiency (SE), sleep latency (SOL) and wake after sleep onset (WASO) assessed by actigraphy gathered in 3 different modes were compared to polysomnography (PSG) measurements to determine which mode corresponded highest to PSG. Associations of measurement error for TST (PSG-actigraphy) with demographics, medical history, exam data, and sleep characteristics were examined.

Methods: Participants underwent in-home 12-channel PSG. Actigraphy data were collected in 3 modes: proportional integration mode (PIM), time above threshold (TAT) and zero crossings mode (ZCM). The analysis cohort was a subgroup of 889 men (mean age 76.4 years) from the MrOS Sleep Study with concurrently measured PSG and actigraphy. Intraclass correlation coefficients (ICCs) were used to compare the association between PSG and actigraphy.

Results: The PIM mode of actigraphy corresponded moderately to PSG for all measures (ICCs 0.32 to 0.57), TAT a little lower (ICCs 0.17 to 0.47), and ZCM lower still (ICCs

0.16 to 0.33). The PIM mode corresponded best to PSG (ICCs TST 0.57; SE 0.46; SOL 0.23; WASO 0.54), though the estimations from PSG and PIM mode differed significantly ($p < 0.01$). The PIM mode overestimated TST by 13.2 min on average, but underestimated TST for those in certain subgroups: those with excessive daytime sleepiness, less sleep fragmentation, or more sleep disordered breathing ($p < 0.05$).

Conclusions: Sleep parameters from the PIM and TAT modes of actigraphy corresponded reasonably well to PSG in this population, with the PIM mode correlating highest. Systematic measurement error was observed within subgroups with different sleep characteristics.

Keywords: Actigraphy, polysomnography, total sleep time, sleep efficiency, validation

Citation: Blackwell T; Ancoli-Israel S; Redline S; Stone KL.; 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.

The measurement of sleep in humans has been typically performed using polysomnography (PSG), the “gold standard” for sleep-wake assessment. The use of wrist actigraphy has become more common in past years as an alternative to PSG for sleep-wake assessment.^{1,2} Actigraphy is a noninvasive, objective method for gathering data on movement via an accelerometer. These data are then used to infer sleep from wake using validated algorithms. Use of actigraphy has the benefit over PSG of being less costly and invasive, and sleep can be monitored continuously for long periods of time. Typical parameters gathered by both PSG and actigraphy are total sleep time (TST), sleep efficiency (SE), sleep onset latency (SOL), and wake after sleep onset (WASO).

There have been numerous studies showing good concordance for sleep-wake estimation between actigraphy and PSG.³⁻¹⁰ As noted by Ancoli-Israel and colleagues, different actigraphic devices may have different measurement of activity level and sleep-wake scoring algorithms, which make direct comparison of devices difficult.¹ There can also be differences in the accuracy of actigraphy for estimation of sleep among differing populations. For example, one study of elderly women and one study of adolescents that both used the same actigraph

BRIEF SUMMARY

Current Knowledge/Study Rationale: This study was performed to examine whether total sleep time, sleep fragmentation, and sleep latency measured with wrist actigraphy were comparable to similar assessments using the gold-standard of PSG, and to determine which of 3 modes of data collection from this actigraph model were best. Potential causes of the difference between the 2 devices were examined.

Study Impact: The PIM mode of actigraphy performed reasonably well, and is preferred in older community-dwelling men. A number of sleep related characteristics had a significant impact on the accuracy of actigraphy for measurement of total sleep time, and should be considered when using actigraphy among populations with sleep disorders.

model and scoring algorithms found differing optimal modes of data collection when compared to PSG.^{3,7} Because of these issues it is preferred to validate actigraphic sleep-wake estimation within specific populations for specific actigraphic devices. To our knowledge there have been no studies comparing sleep measurements from actigraphy to PSG in older community-dwelling men.

Because actigraphy is often used to characterize sleep patterns in subgroups with different underlying pathologies, it is

important to understand whether systematic misclassification of coding wake as sleep and vice versa influences the quantification of sleep patterns. Many studies examined homogeneous samples consisting of those with sleep disordered breathing,^{5,6} insomnia,^{9,11-14} depression,¹⁵ or healthy volunteers.^{4,16} Some have had small sample sizes.^{4,11-16} However, both small and/or homogeneous samples prevent the examination of the association of factors related to misclassification of sleep and wake from actigraphy. A few studies have examined the relationship of misclassification among subgroups within a population.^{3,7} These studies found that while actigraphy does reasonably well estimating sleep from wake, there is less accuracy in subsets such as those with sleep disordered breathing or poor sleep quality.^{3,7}

The Outcomes of Sleep Disorders in Men (MrOS Sleep) Study provides a unique opportunity among a large population of community-dwelling older men to investigate whether TST, SE, SOL, and WASO measured using the actigraph model Sleepwatch-O (Ambulatory Monitoring, Inc, Ardsley, NY) were comparable to similar assessments using PSG, and to determine which of 3 different modes of data collection from this actigraph model were optimal for assessment of sleep in older men. The study was large enough to examine whether differences in estimation of the outcome of TST from PSG and the preferred mode of actigraphy are significantly associated to the presence of sleep problems, comorbidities, demographic or anthropometric factors.

METHODS

Participants

The MrOS Sleep Study recruited 3135 men between December 2003 and March 2005 from a larger study of 5994 men, the Osteoporotic Fractures in Men (MrOS) Study. Community-dwelling men aged 65 and older were recruited to participate in the MrOS study at 6 US clinical centers in Birmingham, AL; Minneapolis, MN; Palo Alto, CA; the Monongahela Valley near Pittsburgh, PA; Portland, OR; and San Diego, CA. At the time of the baseline visit men were ineligible for the study if they had a history of a bilateral hip replacement or were unable to walk without the assistance of another person.^{17,18} Of the 2859 men who did not participate in the ancillary MrOS Sleep Study, 344 died before the sleep visit, 36 had already terminated the study, 332 were not asked because recruitment goals had already been met, and 1997 refused. One hundred fifty men were not eligible for the study due to an open tracheotomy or use of CPAP, bilevel PAP, sleeping with a mouthpiece for snoring or sleep apnea, or using oxygen therapy in the past 3 months with the inability to forgo use of these devices for the PSG recording. However, 49 men who reported intermittent use of one of these devices are in the MrOS Sleep Study, and 17 of these men were able to forgo use of the device on the night of the in-home polysomnography.

There were 3058 men with actigraphy data gathered and 2911 men with PSG data gathered, with 2865 having data for both measures. The actigraph recording was typically started the day of the clinic exam. The study protocol did not specifically require that the participants wear the actigraph while the

PSG recording was performed, but rather that the PSG was performed within one month of the clinical examination. The PSG recording was often done later due to scheduling issues or availability of equipment. Therefore, not all men with a PSG recording had an actigraphic recording done concurrently. There were 896 men with concurrent PSG and actigraphy recordings. Of these, 7 were not included in the analysis because they reported use of CPAP or bilevel PAP in the last 3 months, leaving 889 in this analysis subset. The institutional review boards at each clinic site approved the study, and written informed consent was obtained from all participants.

In-Home Polysomnography

All clinic staff that collected PSG data were required to go through formal, centralized training and pass a certification test before being allowed to oversee collection of sleep study data. PSG data were collected in the participant's home using the Compumedics Safiro Unit (Melbourne, Australia). Channels monitored included 2 central electroencephalographic (EEG) leads (C₃/A₂, C₄/A₁), bilateral electrooculogram (EOG), a bipolar submental electromyogram (EMG), thoracic and abdominal respiratory inductance plethysmography, airflow (by a nasal-oral thermocouple and nasal pressure cannula), finger pulse oximetry, electrocardiogram (ECG), body position, and bilateral leg movements (with piezoelectric sensors). After studies were downloaded, they were transferred to the Central Sleep Reading Center (Cleveland, OH) for centralized scoring by a trained technician. Sleep stages and arousals in the PSG data were scored by certified scorers using standard criteria.^{19,20} Sleep was staged in 30-sec scoring epochs. Scorers were blinded to the results of the actigraphy data. The sleep period was defined as the time from reported lights off to morning awakening. Apneas were defined by the absence or near absence of airflow on thermistor for ≥ 10 sec. Hypopneas were defined by a decrease in breathing amplitude of $\geq 30\%$ for ≥ 10 sec with an oxygen desaturation of 4% or more. The apnea-hypopnea index (AHI) was defined as apneas plus hypopneas/h of sleep time. Sleep onset latency was defined as the first 2 continuous min scored as sleep, and wake after sleep onset was defined as the min awake during the sleep period after sleep onset latency. Sleep efficiency was defined as the percent of time scored as sleep during the sleep period. Nocturnal hypoxia was calculated from PSG as the percent of time during sleep with oxygen saturation (SpO₂) $< 90\%$. Leg movements were scored according to the American Academy of Sleep Medicine (AASM) criteria (> 4 consecutive 0.5 to 5-sec movements, each separated by 5-90 sec).²¹ Leg movements that occurred at the termination of respiratory events were not considered in the calculation of periodic leg movements during sleep (PLMS) unless they were part of a cluster of ≥ 4 leg movements in which ≥ 2 leg movements occurred independently of respiratory event termination. Those PLMS associated with EEG arousals were also calculated. These PLM counts were summarized as an index per hour of sleep (PLMI and PLMAI).

The outcomes of TST, SE, SOL, and WASO were used as continuous variables for comparison to the corresponding actigraphic measures. For summary purposes categorizations of continuous variables were made as follows: TST using the categories ≤ 5 h, > 5 to 7 h, and > 7 h; SE using the categories $< 70\%$, 70 to $< 80\%$, 80 to $< 85\%$, and $\geq 85\%$; SOL using the

categories < 15 min, 15 to < 30 min, \geq 30 min; WASO in the categories < 45 min, 45 to < 90 min, \geq 90 min; AHI using the cut points < 5, 5 to < 15, 15 to < 30, and \geq 30; nocturnal hypoxia was categorized as < 1%, 1 to < 3.5%, 3.5% to < 10% and \geq 10%; both PLMI and PLMAI were categorized as 0, > 0 to < 5, 5 to < 15, and \geq 15.

Wrist Actigraphy

The Sleepwatch-O (Ambulatory Monitoring, Inc, Ardsley, NY) was used. This actigraph, which looks like a wristwatch, measures movement using a piezoelectric biomorph-ceramic cantilevered beam, which generates a voltage each time the actigraph is moved. These voltages are gathered continuously and stored in 1-min epochs. The term “mode” is used to refer to the technique with which different measures were obtained. Data were collected in the 3 modes of zero crossings (ZCM), proportional integration mode (PIM), and time above threshold (TAT). In ZCM mode the conditioned transducer signal is compared with a sensitivity threshold of zero. The number of times the signal voltage crosses zero voltage is summed over the epoch. The ZCM mode is a measure of frequency of movement. The PIM mode is a high-resolution measurement of the area under the rectified conditioned transducer signal (area under the curve). The PIM mode is a measure of activity level or vigor of motion. In TAT mode the amount of time in tenths of a second spent above the sensitivity threshold is gathered over the epoch. The TAT mode measures time spent in motion or duty-cycle (the time spent in an active state).^{1,22}

Actigraphy data were transferred to the San Francisco Coordinating Center (San Francisco, CA) for centralized processing. Centralized training and certification were also required for clinic staff gathering actigraphy data. Action W-2 software was used to score the data.²³ Sleep scoring algorithms available in this software were used to determine sleep from wake times. The Cole-Kripke algorithm was used for data collected in the ZCM mode, and the University of California, San Diego (UCSD) scoring algorithm was used for data collected in the PIM and TAT modes.^{24,25} These algorithms calculate a moving average, which takes into account the activity levels immediately prior to and after the current minute to determine if each time point should be coded as sleep or wake. Although the UCSD algorithm is available for the ZCM mode and the Cole-Kripke algorithm is available for the PIM and TAT modes, previous comparisons using the same model of actigraph have shown a very high rate of agreement between the scoring algorithms within mode.^{3,7} Therefore, the default algorithm for each mode that was selected by the software was used in this analysis. SE, SOL, and WASO were defined similarly to PSG for comparison.

The men wore the actigraphs for a minimum of 4 consecutive 24-h periods. For those 889 men who wore the actigraph concurrent with their PSG recording, the actigraphy files were edited to include only the time period that was assessed by both methods. None of the men removed their actigraphs during this time period. Clock times for PSG and actigraphy were not synchronized, so there may be slight differences in machine times.

Sleep Parameters

The outcome measures of TST, sleep fragmentation (WASO and SE), and SOL were examined. The differences between the

TST, SOL, WASO, and SE as measured by PSG, and each of the 3 modes of actigraphy data collected are presented to show the direction of any bias. Absolute differences are presented to quantify the overall magnitude of differences among measurements without taking into account the direction of the difference (positive or negative).

Other Measurements

At the time of the Sleep Visit all participants completed questionnaire data, which included questions about demographics, medical history, self-reported health, alcohol use, and smoking status. Caffeine consumption was estimated based on self-report of the average daily number of cups of caffeinated coffee, tea, or soda consumed.²⁶ Physical activity was assessed using the Physical Activity Scale for the Elderly (PASE).²⁷ The Geriatric Depression Scale (GDS) was used to assess depressive symptoms, with the standard cutoff of \geq 6 symptoms used to define depression.²⁸ Information to compute the Goldberg anxiety scale was also collected, and the cutpoint of \geq 5 was used to define clinically significant anxiety.²⁹ The Pittsburgh Sleep Quality Index (PSQI) was completed.^{30,31}

At the Sleep clinic visit, the Epworth Sleepiness Scale (ESS) was completed with the cutpoint of ESS > 10 used to define excessive daytime sleepiness.^{32,33} The Teng Modified-Mental State Examination (3MS) was administered to assess cognitive function, with higher scores on a scale of 0 to 100 representing better cognition.³⁴ Functional status was assessed by collecting information on 5 instrumental activities of daily living (IADL), which included walking 2 to 3 blocks on level ground, climbing up to 10 steps, preparing meals, doing heavy housework, and shopping for groceries or clothing.^{35,36} Tests of physical function included walking speed (time in seconds to walk 6 meters at usual pace expressed as meters/sec). A comprehensive examination included measurements of resting blood pressure, body weight and height. Body mass index (BMI) was calculated as weight in kilograms divided by the square of height in meters. Obesity was defined as a BMI \geq 30 kg/m². Circumference at the waist was measured in a standardized fashion.³⁷

All prescription and nonprescription medications taken 24 h prior to and during the PSG recording were collected by the clinics and stored in an electronic medications inventory database (San Francisco Coordinating Center, San Francisco, CA). Each medication was matched to its ingredient(s) based on the Iowa Drug Information Service (IDIS) Drug Vocabulary (College of Pharmacy, University of Iowa, Iowa City, IA).³⁸

Statistical Analysis

Characteristics of this subset of 889 men and the remaining 2246 in the MrOS Sleep Study cohort were summarized by means and standard deviations (SD) for continuous variables, and counts and percentages for categorical variables. Differences in these characteristics between the 2 groups were analyzed using *t*-tests for continuous normally distributed variables, Wilcoxon rank-sum tests for continuous data with skewed distributions, and χ^2 tests for categorical characteristics.

The differences between TST, WASO, SOL, and SE as assessed by the gold standard PSG measurement and those from actigraphy were examined using paired *t*-tests. Agreement between the 2 methods of sleep assessment was examined with

simple correlations and with intraclass correlation coefficients (ICC) and 95% confidence intervals (CI) which were computed using a 2-way analysis of variance.³⁹ Bland and Altman plots were presented to assess systematic bias in the differences in measurement of TST.⁴⁰ Formal tests of systematic bias were performed using simple linear regression models to examine whether the scatter in the Bland and Altman plots was heteroscedastic.⁴¹

Continuous covariates that could potentially explain the differences between the PSG and actigraphy measurements of the outcome of TST were classified into categories. Quartiles were used for age, caffeine, cognition, walking speed, physical activity (PASE), and waist circumference. Other continuous covariates were categorized using cut points from previous analyses or standard definitions as described above. Presence of statistically significant differences between the misclassification of TST by the preferred mode of actigraphy and these covariates were examined using ANOVA. The mean difference between TST assessed by PSG and the preferred mode of actigraphy and the ICC (95% CI) are shown by category of the covariate.

Linear regression models were performed to examine if associations of the difference between TST calculated by PSG and by the preferred mode of actigraphy (PSG - actigraphy) and a given covariate still held after adjustment for clinic site, PSG TST, AHI, and PLMI. The dependent variable in these models was the difference between the 2 devices rather than the absolute difference between the 2 devices in order to examine the direction of association.

All significance levels reported were 2-sided, and all statistical analyses were performed using SAS software, version 9.2 (SAS Institute, Inc, Cary, NC).

RESULTS

Characteristics of the Study Population

Of the 2865 men with both PSG and actigraphy data, 889 (31%) had concurrent data from both methods. There were few men from the Birmingham, AL, or Palo Alto, CA, clinic sites with concurrent measurement of PSG and actigraphy. These 889 men were 76.4 years old on average, with Caucasians making up 93% of this analysis subset (**Table 1**). The mean time spent in bed during the PSG recording was 7.9 ± 1.2 h, with an average TST from PSG of 5.9 h, mean WASO of 112.0 min, sleep efficiency averaging 74.5%, an average sleep onset latency of 12.53 ± 20.53 min, and a mean AHI of 12.0 (**Table 1**). In comparison to the remaining 2246 participants in the MrOS Sleep Study cohort, this subset of 889 men were more likely to be Caucasian, had lower rates of stroke, had slightly higher levels of cognitive function, had slightly higher systolic and diastolic blood pressure, had lower rates of excessive daytime sleepiness, and had less TST and time in bed ($p < 0.05$; **Table 1**).

Comparison of TST, SE, SOL and WASO Calculated by PSG and Actigraphy

There was a statistically significant difference between the estimation of the outcome of TST by the gold standard PSG and all 3 modes of actigraphy ($p < 0.01$ for paired *t*-test, **Table 2**). Higher levels of agreement with PSG were observed for the

PIM mode than other actigraphic modes, with an average overestimation of TST of 13.2 min (range -280 to 329 min) and an absolute difference of 52.9 min on average. The TAT mode overestimated sleep on average by 22.1 min (range -276 to 386 min), while the ZCM mode underestimated sleep by an average of 59.0 min (range -235 to 436 min). While these differences are statistically significant, the intraclass correlation coefficients of the PSG measurement of TST and the data from the PIM and TAT mode of actigraphy were moderate (0.57 for the PIM mode, 0.47 for the TAT mode). Examining the Bland and Altman plots comparing PSG to actigraphic TST showed a systematic bias towards overestimation for the PIM and TAT modes which increased with decreasing sleep duration ($p < 0.0001$, **Figure 1**). The plots also show the actigraphic TST measurement corresponded more closely to PSG when TST was longer, showing a systematic bias in misclassification for short sleepers. When subset to those with PSG sleep duration of 6-8 h, there is no significant systematic bias for the actigraphic estimation of TST by the PIM or TAT modes ($p > 0.29$). The mean difference between actigraphic and PSG measurement of TST was closer to 0 (agreement) for the PIM mode, which had a more compact clustering of differences, a more compact 95% confidence interval for the mean difference, and less points falling outside of these 95% confidence intervals.

The PIM mode of actigraphy also corresponded better with PSG for measures of sleep fragmentation and SOL. For all 3 modes, there was a significant difference between the 2 procedures in estimation of WASO, SOL, and SE ($p < 0.01$). For the PIM mode, there was an average underestimation of 11.6 minutes for WASO (range -300 to 278 minutes) and an ICC of 0.54 (**Table 2**). SE estimation was slightly overestimated, on average, by the PIM mode (2.6%), while SOL was slightly underestimated (2.8 min on average).

Associations of Participant Characteristics and the Differences between TST from PSG and Actigraphy (PIM Mode)

The variables associated with misclassification of TST by actigraphy are shown in **Table 3**. The 2 indices related to sleep disordered breathing (SDB), AHI and nocturnal hypoxia, had a linear relationship to actigraphic misclassification of TST ($p < 0.001$). Those categories with low levels of SDB showed a mean overestimation of TST by actigraphy of about 25 min (23.3 min for time with nocturnal hypoxia $< 1\%$; 24.9 min for AHI < 5). For those with the highest levels of SDB the relationship is reversed: actigraphy underestimated TST, on average by about 16 min (16.7 min for nocturnal hypoxia $\geq 10\%$; 16.3 min for AHI ≥ 30). Of note, the ICC for those with an AHI ≥ 30 was 0.36, which was considerably lower than what is seen in the entire study population. Short sleepers (≤ 5 h) had an average overestimation of TST by actigraphy of 29 min ($p < 0.01$). Those with TST > 7 h showed very little misclassification (mean overestimate of 7.2 min). Those with higher rates of sleep fragmentation as measured by PSG SE and WASO showed a systematic overestimation of TST by actigraphy ($p < 0.001$). For those with PSG WASO ≥ 90 min, there was an average overestimation of TST by 30.9 min, while those with PSG SE $< 70\%$ had an average overestimation of 44.9 min. For those with lower levels of sleep fragmentation there was

Table 1—Comparing characteristics of the analysis subset to the remaining MrOS sleep study population

Characteristic	Remaining MrOS Participants (N = 2246)	Participants in Analysis Subset (N = 889)	p-value
Clinic			< 0.01
Birmingham, AL	498 (22.17)	17 (1.91)	
Minneapolis, MN	322 (14.34)	212 (23.85)	
Palo Alto, CA	463 (20.61)	41 (4.61)	
Pittsburgh, PA	308 (13.71)	212 (23.85)	
Portland, OR	312 (13.89)	201 (22.61)	
San Diego, CA	343 (15.27)	206 (23.17)	
Age, years	76.48 ± 5.60	76.28 ± 5.47	0.37
Race/Ethnicity			< 0.01
Caucasian	1988 (88.51)	828 (93.14)	
African American	101 (4.50)	20 (2.25)	
Asian	82 (3.65)	19 (2.14)	
Hispanic/other	75 (3.34)	22 (2.47)	
Self-reported health status good/excellent	1944 (86.67)	770 (86.61)	0.97
1 or more IADL impairments	491 (21.87)	176 (19.80)	0.20
Alcohol intake, drinks/week			0.84
0-2 drinks/week	1329 (59.60)	522 (58.78)	
3-13 drinks/week	775 (34.75)	318 (35.81)	
≥ 14 drinks/week	126 (5.65)	48 (5.41)	
Average caffeine intake, mg/day	234.39 ± 242.49	237.69 ± 256.62	0.57
Current smoker	51 (2.27)	13 (1.46)	0.15
Antidepressant use during PSG recording	159 (7.08)	57 (6.41)	0.50
Benzodiazepine use during PSG recording	77 (3.43)	30 (3.37)	0.94
Sleep medication use during PSG recording	34 (1.51)	8 (0.90)	0.18
Any of the selected medical conditions	928 (41.35)	364 (40.94)	0.83
History of stroke	93 (4.14)	24 (2.70)	0.05
History of diabetes	294 (13.10)	123 (13.84)	0.59
History of Parkinson disease	33 (1.47)	7 (0.79)	0.12
History of COPD/emphysema	121 (5.39)	43 (4.84)	0.53
History of coronary heart disease*	742 (33.14)	292 (32.88)	0.89
History of hypertension	1102 (49.11)	458 (51.52)	0.22

Table 1 continues on the following page

an average underestimation of about 19 min. For those without periodic leg movements during sleep or those with no PLMs causing arousals there was an overestimation of actigraphic TST by about half-hour ($p < 0.001$). As levels of PLM without arousal rose, these overestimations were reduced, bringing the average overestimation closer to 0. For those with ≥ 15 PLMs causing arousals per hour of sleep the direction of misclassification was reversed: the actigraph underestimates the TST an average of 25 minutes. All associations held after adjustment for clinic site, PSG TST, AHI, and PLMI. For example, the adjusted means of the difference in PSG and actigraphy-measured TST for categories of PSG WASO were 24.3 of underestimation for those with PSG WASO < 45 min, an underestimation of 5.6 min for those with PSG WASO 45 to < 90 min, and an adjusted average overestimation of 31.6 min for those with PSG WASO ≥ 90 min.

A few non-PSG characteristics were also associated to actigraphic misclassification of TST. Those who did not take antidepressants on the night the concurrent PSG and actigraphy

recordings had a rate of overestimation similar to that of the entire study population, while users of antidepressants had an underestimation of 20.5 minutes, on average ($p < 0.001$). Those with higher BMI showed better agreement of TST measured by actigraphy and PSG ($p < 0.001$). Those classified as obese (BMI ≥ 30 kg/m²) showed a slight underestimation of TST by actigraphy (2.2 min on average). Those in the lowest quartile of waist circumference had an average overestimation of 23 minutes. This misclassification was reduced as the quartiles increase. Those who met the criteria for excessive daytime sleepiness (ESS > 10) showed the best agreement of actigraphy and PSG ($p = 0.03$). All associations but that of BMI held after adjustment for clinic site, PSG TST, AHI, and PLMI.

No other characteristics examined, including age, race, medical conditions, depression, functional status, sleep onset latency, and cognition were significantly associated to the difference in PSG and actigraphic PIM mode estimation of TST ($p > 0.10$, data not shown).

Table 1 (continued)—Comparing characteristics of the analysis subset to the remaining MrOS sleep study population

Characteristic	Remaining MrOS Participants (N = 2246)	Participants in Analysis Subset (N = 889)	p-value
Depression (Geriatric Depression Score ≥ 6)	153 (6.83)	58 (6.52)	0.76
Goldberg anxiety disturbance (score ≥ 5)	192 (8.57)	85 (9.56)	0.38
Teng 3MS score (range 0-100)	92.41 \pm 6.67	93.14 \pm 5.65	< 0.01
Walking speed, m/s	1.14 \pm 0.23	1.14 \pm 0.23	0.79
PASE score	146.01 \pm 72.61	144.57 \pm 70.23	0.61
Body mass index, kg/m ²	27.12 \pm 3.85	27.36 \pm 3.86	0.11
Waist circumference, cm	99.48 \pm 10.96	100.22 \pm 11.04	0.09
Diastolic blood pressure, mm Hg	67.44 \pm 9.48	68.18 \pm 9.53	0.05
Systolic blood pressure, mm Hg	126.50 \pm 16.47	128.11 \pm 16.11	0.01
Pittsburgh sleep quality index (range 0-21)	5.64 \pm 3.28	5.58 \pm 3.27	0.67
Epworth sleepiness scale (range 0-24)	6.27 \pm 3.77	5.91 \pm 3.49	< 0.01
Excessive daytime sleepiness	314 (13.99)	92 (10.35)	< 0.01
Total sleep time from PSG, min	357.20 \pm 71.13	351.66 \pm 65.21	0.04
Time in bed, min	488.99 \pm 77.59	475.73 \pm 73.71	< 0.01
Sleep efficiency from PSG, %	73.51 \pm 12.34	74.51 \pm 12.12	0.05
Sleep latency from PSG, min	14.53 \pm 25.78	12.53 \pm 20.53	0.81
WASO from PSG, min	117.26 \pm 67.52	111.53 \pm 64.76	0.04
Apnea-hypopnea index	11.64 \pm 12.81	11.97 \pm 13.37	0.80
% of sleep time spent with SpO ₂ < 90%			0.90
< 1	971 (48.02)	439 (49.38)	
1 to < 3.5	540 (26.71)	227 (25.53)	
3.5 to < 10	261 (12.91)	113 (12.71)	
≥ 10	250 (12.36)	110 (12.37)	
Periodic leg movements/h sleep	35.73 \pm 37.80	35.71 \pm 36.92	0.80
Periodic leg movements causing arousal/h sleep	3.92 \pm 5.69	4.30 \pm 5.80	0.08

Mean \pm SD or n (%). *Coronary heart disease includes a history of myocardial infarction, angina, congestive heart failure, bypass surgery, angioplasty, or pacemaker placement. p-values for continuous normally distributed variables are from a *t*-test, skewed from a Wilcoxon Rank-sum test. p-values for categorical variables from a χ^2 test.

DISCUSSION

The results from this study of community-dwelling older men suggest a moderate correlation between the outcome of TST measured concurrently with PSG and actigraphy with both the PIM and TAT modes of data collection, and a low correlation with the ZCM mode. The PIM mode corresponded more closely to the PSG estimation of TST than the estimations from the other two actigraphic modes. The results also suggest some systematic misclassification of sleep and wake by actigraphy. Overall, there was an overestimation of TST by actigraphy, which corresponded to results from other studies consisting of participants that were not selected on the basis of sleep complaints.^{3,4,10,16} The average difference in measurement of TST between the PIM mode and PSG was only 13 minutes, with an average error rate of 4.6%. It is likely this average measurement error would not be considered clinically significant or effect diagnosis and treatment decisions.

The results from the outcomes of WASO, SOL, and SE were largely similar to the TST results, with moderate correlations between the measurements from PSG and the actigraphic modes of PIM and TAT, and low correlations for the ZCM mode. Once again, the PIM mode of actigraphy corresponded better to PSG for these measurements. As with TST, the average differences in

measurement between the PIM mode of actigraphy and the gold standard of PSG were relatively small (SE 2.6%, SOL 2.8 min, WASO 11.6 min).

The misclassification of sleep and wake by actigraphy was magnified by underlying sleep related problems. On average, those with short sleep periods as estimated by PSG tended to have an overestimation of TST by actigraphy. Those with higher levels of sleep fragmentation also had an overestimation of TST by actigraphy. There are previous studies with populations of insomniacs and those with other sleep disorders that have also found an overestimation of sleep by actigraphy.^{9,11,13} This overestimation may be due to long periods of wakefulness that were unaccompanied by movement, causing the sleep scoring algorithms to detect sleep rather than wake. Those with higher levels of sleep disordered breathing had an average underestimation of sleep by actigraphy, possibly caused by the sleep scoring algorithm classifying as wake those slight SDB-associated movements which are not accompanied by an EEG arousal. This average underestimation of TST among those with SDB is consistent with previous results.⁵⁻⁷ While the magnitude of measurement error was larger in these subgroups with sleep problems and may be clinically significant, the inaccuracy of measurement occurs most in these more severe subgroups that would be easier to identify. The

Table 2—Comparison of total sleep time, sleep efficiency, sleep onset latency and WASO calculated by PSG and actigraphy (N = 889 pairs)

	mean ± SD	Difference (PSG-Actigraphy) mean ± SD	Absolute Difference mean ± SD	rho	ICC (95% CI)
Total sleep time, min					
PSG	351.66 ± 65.21				
Actigraphy Mode					
PIM	364.85 ± 86.81	-13.18 ± 70.02	52.85 ± 47.76	0.61	0.57 (0.53, 0.62)
TAT	373.79 ± 97.81	-22.12 ± 83.64	67.14 ± 54.52	0.53	0.47 (0.42, 0.52)
ZCM	292.65 ± 125.37	59.01 ± 116.67	94.96 ± 89.83	0.39	0.21 (0.15, 0.27)
Sleep efficiency, %					
PSG	74.51 ± 12.12				
Actigraphy Mode					
PIM	77.05 ± 15.67	-2.55 ± 14.33	10.93 ± 9.61	0.49	0.46 (0.41, 0.51)
TAT	78.99 ± 18.34	-4.48 ± 7.16	13.94 ± 10.97	0.42	0.36 (0.31, 0.42)
ZCM	61.99 ± 25.72	12.52 ± 24.24	19.92 ± 18.64	0.35	0.16 (0.10, 0.22)
Sleep onset latency, min					
PSG	12.53 ± 20.53				
Actigraphy Mode					
PIM	9.76 ± 17.27	2.77 ± 22.03	10.58 ± 19.52	0.44	0.32 (0.26, 0.38)
TAT	10.11 ± 22.58	2.43 ± 27.82	12.37 ± 25.04	0.39	0.17 (0.10, 0.23)
ZCM	29.88 ± 52.78	-17.56 ± 51.37	24.26 ± 48.56	0.36	0.12 (0.06, 0.19)
WASO, min					
PSG	111.53 ± 64.76				
Actigraphy Mode					
PIM	100.07 ± 73.52	11.61 ± 65.40	49.52 ± 44.23	0.56	0.54 (0.50, 0.59)
TAT	90.10 ± 81.16	21.79 ± 76.59	62.02 ± 49.91	0.47	0.42 (0.37, 0.48)
ZCM	143.11 ± 102.62	-31.06 ± 95.33	74.69 ± 66.85	0.42	0.33 (0.27, 0.39)

All p-values from a *t*-test on the paired data for difference were $p < 0.0001$. Rho = correlation coefficient for TST from a Pearson correlation, SOL, WASO and SE from a Spearman correlation.

benefit of actigraphy is the ability to measure sleep for multiple nights, which allows examination of sleep patterns. This study found no association between the misclassification of TST by actigraphy with SOL, demographic, lifestyle, depression, anxiety, cognition, blood pressure, physical activity level or medical history data, which was similar to findings of another similar study.³ There was a relationship found between the misclassification of TST and body size, with less misclassification among the obese and those in the highest levels of waist circumference.

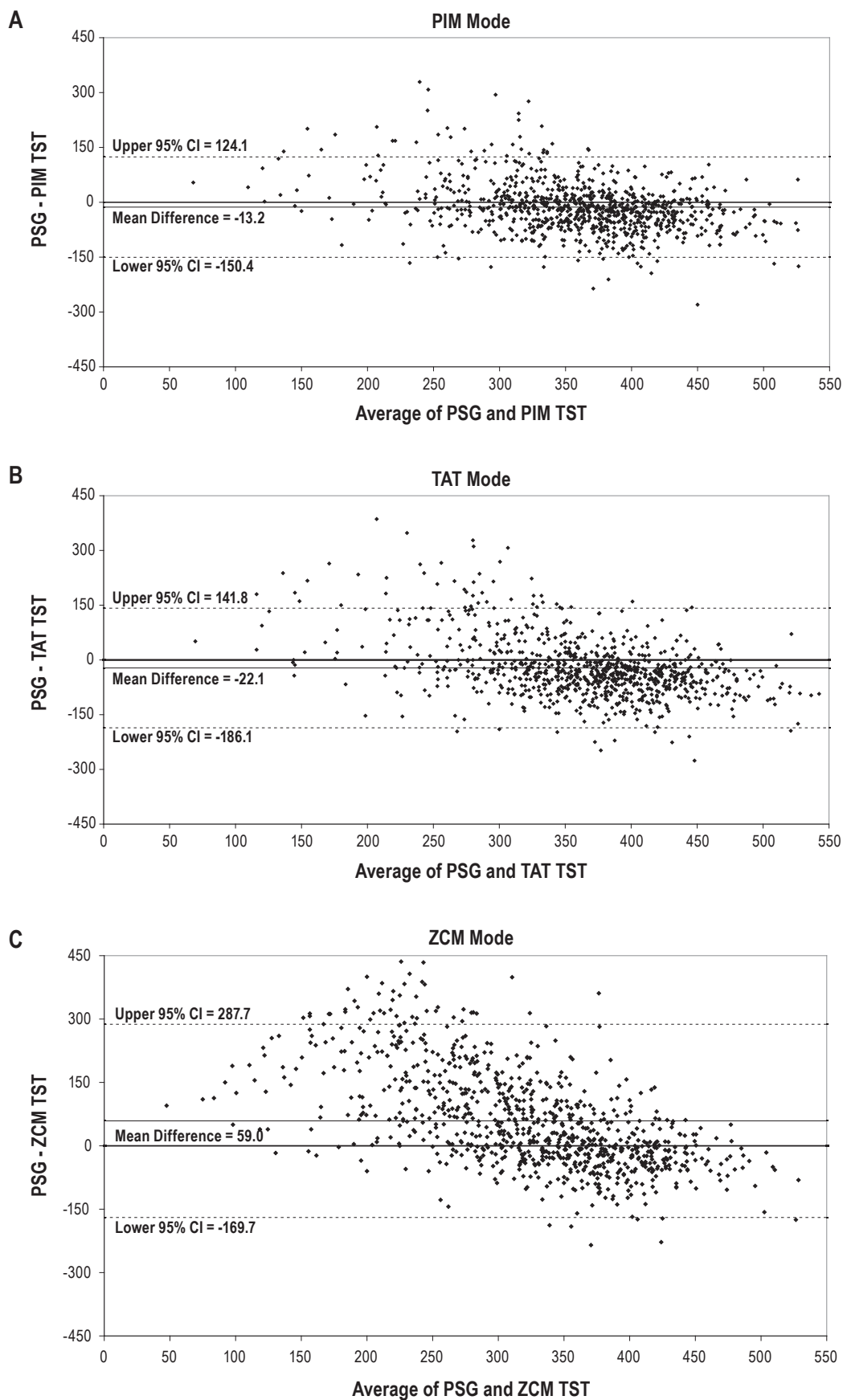
The study has several strengths. To our knowledge, this is the largest study examining the estimation of sleep and wake from wrist actigraphy compared to polysomnography. The community-dwelling men in this large population were not selected for inclusion based on sleep problems. The size and heterogeneous population allowed us to have the statistical power to examine the potential underlying causes of misclassification of sleep and wake by actigraphy. We compared three different modes of actigraphy to polysomnography. Data were collected in the home rather than a sleep laboratory, so disruption of sleep in an unfamiliar environment was minimized.

This study also had several limitations. The findings may not be generalizable to populations other than community-dwelling

older men, or to other to actigraphic devices. The study protocol did not require concurrent PSG and actigraphy recording, making it impossible to include all men with data from both measures. Those men not included in our analysis did differ from our analysis subset on some characteristics, most notably levels of excessive daytime sleepiness, time spent in bed, and total sleep time. Clock times for PSG and actigraphy were not synchronized, so there may be slight differences in machine times. PSG data were collected in 30-sec epochs, actigraphy data in 1-min epochs. Because of the lack of synchronization and differing epoch lengths, direct comparison of each epoch was not possible.

In conclusion, data from actigraphy estimated TST, SE, SOL, and WASO reasonably well when compared to PSG estimations. The PIM mode of actigraphy correlated best to PSG estimations of sleep and wake in this population of community-dwelling older men, and could be used as a proxy measure for the sleep parameters examined. A number of sleep related characteristics had a significant impact on the accuracy of actigraphy to measure total sleep time. Actigraphy underestimated total sleep time for those with excessive daytime sleepiness, those with lower levels of sleep fragmentation, and those with higher levels of sleep disordered breathing. Actigraphy overestimated total sleep time for

Figure 1—Bland and Altman plots of total sleep time



Dashed lines represent 95% confidence intervals, solid lines represent the mean difference, the solid line at zero represents perfect agreement.

Table 3—Those characteristics associated with the difference between measurement of total sleep time from PSG and actigraphy (in PIM Mode)

Characteristic	Category	N	Difference (PSG-PIM)		p-value
			mean \pm SD	ICC(95% CI)	
Antidepressant use	Nonuser	832	-15.49 \pm 67.99	0.59 (0.54, 0.63)	< 0.001
	User	57	20.53 \pm 88.88	0.35 (0.10, 0.55)	
Body mass index, kg/m ²	< 30	695	-17.22 \pm 65.59	0.57 (0.52, 0.62)	0.001
	\geq 30	194	1.28 \pm 82.57	0.56 (0.45, 0.65)	
Waist circumference, cm	Quartile 1	223	-22.97 \pm 61.63	0.57 (0.47, 0.65)	0.002
	Quartile 2	224	-19.44 \pm 64.84	0.58 (0.48, 0.66)	
	Quartile 3	221	-10.78 \pm 73.99	0.54 (0.44, 0.63)	
	Quartile 4	221	0.62 \pm 76.72	0.58 (0.49, 0.66)	
Epworth Sleepiness Scale	\leq 10	797	-14.98 \pm 68.85	0.58 (0.53, 0.62)	0.025
	> 10	92	2.34 \pm 78.09	0.53 (0.36, 0.66)	
Total sleep time from PSG	\leq 5 h	165	-28.75 \pm 71.94	0.37 (0.24, 0.50)	< 0.001
	> 5 to 7 h	614	-12.65 \pm 66.88	0.22 (0.14, 0.29)	
	> 7 h	110	7.17 \pm 78.84	0.12 (-0.07, 0.30)	
PSG sleep efficiency, %	< 70	262	-44.93 \pm 79.43	0.50 (0.40, 0.58)	< 0.001
	70 to < 80	307	-10.62 \pm 61.18	0.55 (0.47, 0.63)	
	80 to < 85	149	1.32 \pm 58.34	0.55 (0.42, 0.65)	
	\geq 85	171	18.20 \pm 59.00	0.58 (0.47, 0.67)	
PSG WASO, min	< 45	101	19.37 \pm 54.67	0.76 (0.66, 0.83)	< 0.001
	45 to < 90	289	6.04 \pm 59.06	0.57 (0.49, 0.65)	
	\geq 90	499	-30.91 \pm 73.35	0.53 (0.46, 0.59)	
Apnea-hypopnea index	< 5	349	-24.85 \pm 59.25	0.60 (0.52, 0.66)	< 0.001
	5 to < 15	301	-9.73 \pm 65.05	0.60 (0.53, 0.67)	
	15 to < 30	144	-11.58 \pm 71.57	0.60 (0.49, 0.70)	
	\geq 30	95	16.27 \pm 102.93	0.36 (0.17, 0.52)	
% of sleep time spent with SpO ₂ < 90%	< 1	439	-23.31 \pm 58.77	0.60 (0.54, 0.66)	< 0.001
	1 to < 3.5	227	-16.99 \pm 71.01	0.55 (0.45, 0.63)	
	3.5 to < 10	113	4.75 \pm 77.05	0.54 (0.39, 0.66)	
	\geq 10	110	16.66 \pm 88.48	0.50 (0.35, 0.63)	
Periodic leg movements/h sleep	0	108	-37.16 \pm 61.32	0.60 (0.47, 0.71)	< 0.001
	> 0 to < 5	147	-21.50 \pm 61.01	0.59 (0.48, 0.69)	
	5 to < 15	86	-14.93 \pm 58.70	0.62 (0.47, 0.73)	
	\geq 15	548	-5.96 \pm 74.23	0.55 (0.49, 0.61)	
Periodic leg movements causing arousal/h sleep	0	167	-33.59 \pm 61.06	0.64 (0.54, 0.72)	< 0.001
	> 0 to < 5	456	-15.71 \pm 61.67	0.55 (0.48, 0.61)	
	5 to < 15	209	-1.70 \pm 76.16	0.56 (0.46, 0.65)	
	\geq 15	57	24.74 \pm 104.84	0.41 (0.17, 0.60)	

p-values are from ANOVA.

those with higher levels of sleep fragmentation, lower levels of sleep disordered breathing, and short sleepers. The average misclassification for all four sleep parameters examined was relatively small and would likely not effect clinical diagnosis or treatment.

ABBREVIATIONS

PSG, polysomnography

TST, total sleep time
 SE, sleep efficiency
 SOL, sleep onset latency
 WASO, wake after sleep onset
 EEG, electroencephalographic
 EOG, electrooculogram
 EMG, electromyogram
 ECG, electrocardiogram

AHI, apnea-hypopnea index
 AASM, American Academy of Sleep Medicine
 ZCM, zero crossings mode
 PIM, proportional integration mode
 TAT, time above threshold
 UCSD, University of California, San Diego
 PASE, Physical Activity Scale for the Elderly
 GDS, Geriatric Depression Scale
 PSQI, Pittsburgh Sleep Quality Index
 ESS, Epworth Sleepiness Scale
 3MS, Teng Modified-Mental State Examination
 IADL, instrumental activities of daily living
 BMI, body mass index
 SD, standard deviation
 ICC, intraclass correlation coefficient
 CI, confidence interval
 SDB, sleep disordered breathing

REFERENCES

1. Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep* 2003;26:342-92.
2. Morgenthaler T, Alessi C, Friedman L, et al. Standards of Practice Committee; American Academy of Sleep Medicine. Practice parameters for the use of actigraphy in the assessment of sleep and sleep disorders: an update for 2007. *Sleep* 2007;30:519-29.
3. Blackwell T, Redline S, Ancoli-Israel S, et al. Study of Osteoporotic Fractures Research Group. Comparison of sleep parameters from actigraphy and polysomnography in older women: the SOF study. *Sleep* 2008;31:283-91.
4. de Souza L, Benedito-Silva AA, Pires ML, Poyares D, Tufik S, Caill H. Further validation of actigraphy for sleep studies. *Sleep* 2003;26:81-5.
5. Hedner J, Pillar G, Pittman SD, Zou D, Grote L, White DP. A novel adaptive wrist actigraphy algorithm for sleep-wake assessment in sleep apnea patients. *Sleep* 2004;27:1560-6.
6. Hyde M, O'Driscoll DM, Binette S, et al. Validation of actigraphy for determining sleep and wake in children with sleep disordered breathing. *J Sleep Res* 2007;16:213-6.
7. Johnson NL, Kirchner HL, Rosen CL, et al. Sleep estimation using wrist actigraphy in adolescents with and without sleep disordered breathing: a comparison of three data modes. *Sleep* 2007;30:899-905.
8. Jean-Louis G, Kripke DF, Cole RJ, Assmus JD, Langer RD. Sleep detection with an accelerometer actigraph: comparisons with polysomnography. *Physiol Behav* 2001;72:21-8.
9. Lichstein KL, Stone KC, Donaldson J, et al. Actigraphy validation with insomnia. *Sleep* 2006;29:232-9.
10. Mullaney DJ, Kripke DF, Messin S. Wrist-actigraphic estimation of sleep time. *Sleep* 1980;3:83-92.
11. Hauri PJ, Wisbey J. Wrist actigraphy in insomnia. *Sleep* 1992;15:293-301.
12. Jean-Louis G, Zizi F, von Gizycki H, Hauri P. Actigraphic assessment of sleep in insomnia: application of the Actigraph Data Analysis Software (ADAS). *Physiol Behav* 1999;65:659-63.
13. Sivertsen B, Omvik S, Havik OE, et al. A comparison of actigraphy and polysomnography in older adults treated for chronic primary insomnia. *Sleep* 2006;29:1353-8.
14. Vallières A, Morin CM. Actigraphy in the assessment of insomnia. *Sleep* 2003;26:902-6.
15. Jean-Louis G, Mendlowicz MV, Gillin JC, et al. Sleep estimation from wrist activity in patients with major depression. *Physiol Behav* 2000;70:49-53.
16. Paquet J, Kawinska A, Carrier J. Wake detection capacity of actigraphy during sleep. *Sleep* 2007;30:1362-9.
17. Blank JB, Cawthon PM, Carrion-Petersen ML, et al. Overview of recruitment for the osteoporotic fractures in men study (MrOS). *Contemp Clin Trials* 2005;26:557-68.
18. Orwoll E, Blank JB, Barrett-Connor E, et al. Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study—a large observational study of the determinants of fracture in older men. *Contemp Clin Trials* 2005;26:569-85.
19. Rechtschaffen A, Kales A, eds. *A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects*. Washington DC: National Institutes of Health; 1968. NIH publication 204.

20. American Sleep Disorders Association. EEG arousals: scoring rules and examples: a preliminary report from the Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. *Sleep* 1992;15:173-84.
21. ASDA Atlas Task Force. Recording and scoring leg movements. The Atlas Task Force. *Sleep* 1993;16:748-59.
22. Motionlogger User's Guide: Act Millennium, Ambulatory Monitoring, Inc. Ardsley NY.
23. Action-W User's Guide, Version 2.0. Ambulatory Monitoring, Inc. Ardsley NY.
24. Cole RJ, Kripke DF, Gruen W, Mullaney DJ, Gillin JC. Automatic sleep/wake identification from wrist activity. *Sleep* 1992;15:461-9.
25. Jean-Louis G, Kripke DF, Mason WJ, Elliot JA, Youngstedt SD. Sleep estimation from wrist movement quantified by different actigraphic modalities. *J Neurosci Methods* 2001;105:185-91.
26. Barone JJ, Roberts HR. Caffeine consumption. *Food Chem Toxicol* 1996;34:119-29.
27. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): Development and Evaluation. *J Clin Epidemiol* 1993;46:153-62.
28. Sheikh J, Yesavage J. Geriatric Depression Scale: recent evidence and development of a shorter version. *Clin Gerontol* 1986; 5:165-73.
29. Goldberg D, Bridges K, Duncan-Jones P, Grayson D. Detecting anxiety and depression in general medical settings. *BMJ* 1988;297:897-9.
30. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989;28:193-213.
31. Buysse DJ, Reynolds CF, Monk TH, Hoch CC, Yeager AL, Kupfer DJ. Quantification of subjective sleep quality in healthy elderly men and women. *Sleep* 1991;14:331-8.
32. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 1991;14:540-5.
33. Johns MW. Sensitivity and specificity of the multiple sleep latency test (MSLT), the maintenance of wakefulness test and the Epworth sleepiness scale: failure of the MSLT as a gold standard. *J Sleep Res* 2000;9:5-11.
34. Teng EL, Chui HC. The Modified Mini-Mental State (3MS) examination. *J Clin Psychiatry* 1987;48:314-8.
35. Fitti JE, Kovar MG. The supplement on aging to the 1984 National Health Interview Survey. Vital & Health Statistics-series 1: Programs & collection procedures. 1987;21:1-115.
36. Pincus T, Summey JA, Soraci SA Jr, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. *Arthritis Rheum* 1983;26:1346-53.
37. Callaway CW, Buchard C. Circumferences. In: Lohman TG, Roche AF, Martorell R, eds. *Anthropometric standardization reference manual*. Champaign, IL: Human Kinetic Books, 1988:41-5.
38. Pahor M, Chrischilles EA, Guralnik JM, Brown SL, Wallace RB, Carbonin P. Drug data coding and analysis in epidemiologic studies. *Eur J Epidemiol* 1994;10:405-11.
39. Shrout PE, Fleiss LJ. Interclass correlations: Uses in assessing rater reliability. *Psychol Bull* 1979;86:420-8.
40. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-10.
41. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med* 2000;30:1-15.

ACKNOWLEDGMENTS

The Osteoporotic Fractures in Men (MrOS) Study is supported by National Institutes of Health funding. The following institutes provide support: the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), the National Institute on Aging (NIA), the National Cancer Institute (NCI), the National Center for Research Resources (NCRR) and NIH Roadmap for Medical Research under the following grant numbers: U01 AR45580, U01 AR45614, U01 AR45632, U01 AR45647, U01 AR45654, U01 AR45583, U01 AG18197, U01-AG027810, UL1 RR024140, and AG08415.

The National Heart, Lung, and Blood Institute (NHLBI) provides funding for the MrOS Sleep ancillary study "Outcomes of Sleep Disorders in Older Men" under the following grant numbers: R01 HL071194, R01 HL070848, R01 HL070847, R01 HL070842, R01 HL070841, R01 HL070837, R01 HL070838, and R01 HL070839.

This work was performed at the San Francisco Coordinating Center. Investigators in the Outcomes of Sleep Disorders in Older Men study (MrOS Sleep):

Coordinating Center (California Pacific Medical Center Research Institute and University of California, San Francisco): K.L. Stone (Principal Investigator), D.C. Bauer (co-Investigator), S.R. Cummings (co-Investigator), N. Goldschlager (co-Investigator), P. Varosy (co-Investigator), K. Yaffe (co-Investigator), P.M. Cawthon

(co-Investigator), R. Fullman (Project Director), R. Benard, T. Blackwell, L. Conception, J. Diehl, S. Ewing, C. Fox, M. Jaime-Chavez, E. Kwan, S. Litwack, W. Liu, L.Y. Lui, J. Schneider, R. Scott, D. Tanaka, J. Ziarno; **Administrative Center (Oregon Health & Sciences University):** E. Orwoll (Principal Investigator), K. Phipps (co-Investigator), L. Marshall (co-Investigator), J. Babich Blank (Project Director), L. Lambert, B. Chan, D. Neevel; **University of Alabama, Birmingham:** C.E. Lewis (Principal Investigator), J. Shikany (co-Investigator), P. Johnson (Project Director), C. Oden, S. House, N. Webb, K. Hardy, S. Felder, J. Wilkoff, J. King, T. Johnsey, M. Young, J. Smith, C. Sassaman, C. Collier, C. Atkins; **University of Minnesota:** K. Ensrud (Principal Investigator), H. Fink (co-Investigator), D. King (Program Manager), N. Michaels (Asst. Program Manager), N. Nelson (Clinic Coordinator), C. Bird, D. Blanks, F. Imker-Witte, K. Moen, M. Paudel, M. Slindee; **Stanford University:** M. Stefanick (Principal Investigator), A. Hoffman (co-Investigator), K. Kent, B. Malig, S. Wong; **University of Pittsburgh:** J. Cauley (Principal Investigator), J. Zmuda (co-Investigator), M. Danielson (Study Administrator), L. Harper (Project Director), L. Buck (Clinic Coordinator), M. Nasim, D. Cusick, M. Gorecki, N. Watson, C. Bashada, C. Newman; **University of California, San Diego:** E. Barrett-Connor (Principal Investigator), S. Ancoli-Israel (co-Investigator), T. Dam (co-Investigator), ML Carrion-Petersen (Project Director), P. Miller, N. Kamantigue; **Case Western Reserve University:** S. Redline (Principal Investigator), S. Surovec (Project Administrator), N. Scott (Chief Polysomnologist), N. Johnson (Programmer Analyst), J. Arnold (Polysomnologist), R. Nawabit (Polysomnologist), J. Romaniuk (Polysomnologist), S. Seacian (Polysomnologist).

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication December, 2010

Submitted in final revised form February, 2011

Accepted for publication February, 2011

Address correspondence to: Terri Blackwell, M.A., San Francisco Coordinating Center, 185 Berry Street, Lobby 5, Suite 5700, San Francisco, CA 94107; Tel: (415) 600-7412; Fax: (415) 514-8150; E-mail: tblackwell@sfcc-cpmc.net

DISCLOSURE STATEMENT

This was not an industry supported study. Dr. Ancoli-Israel has consulted for or is on the advisory board of Ferring Pharmaceuticals, GlaxoSmithKline, Merck, NeuroVigil, Pfizer, Philips Respironics, Sanofi-Aventis, Sepracor, Schering-Plough, and Perdue and has received research support from Sepracor and Litebook. Dr. Redline is the incumbent of an endowed chair professorship donated to Harvard Medical School by Dr. Peter Farrell, the founder and board chairman of ResMed, Inc, has received research support from Dymedix, Inc, and has received equipment for use in research from Philips Respironics. The other authors have indicated no financial conflicts of interest.