Exercise effects on night-to-night fluctuations in self-rated sleep among older adults with sleep complaints

MATTHEW P. BUMAN[1], ERIC B. HEKLER[1], DONALD L. BLIWISE[2] and ABBY C. KING[3]

[1]Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA, [2]Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA and [3]Department of Health Research & Policy and Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA

SUMMARY   Sleep interventions have rarely explored reductions in night-to-night fluctuations [i.e. intra-individual variability (IIV)] in sleep, despite the negative impacts of such fluctuations on affective states and cognitive and physical symptoms. In a community-based randomized controlled trial we evaluated whether physical exercise reduced IIV in self-rated sleep outcomes among middle-aged and older adults with sleep complaints. Under-active adults 55 years and older ($n = 66$, 67% women) with mild to moderate sleep complaints were randomized to 12 months of a moderate-intensity endurance exercise ($n = 36$) or a health education control group ($n = 30$). Daily sleep logs, Pittsburgh Sleep Quality Index (PSQI) and in-home polysomnographic sleep recordings (PSG) were collected at baseline, 6 months and 12 months. Sleep log-derived means and IIV were computed for sleep-onset latency (SOL), time in bed, feeling rested in the morning, number of nighttime awakenings, and wake after final awakening (WAFA). Using intent-to-treat methods, at 6 months no differences in IIV were observed by group. At 12 months, SOL-based IIV was reduced in the exercise group compared with the control (difference = 23.11, 95% CI: 3.04–47.18, $P = 0.025$, Cohen's $d = 0.57$). This change occurred without mean-level or IIV changes in sleep–wake schedules. For all sleep variables, except SOL and WAFA, IIV changes and mean-level changes in each variable were negatively correlated ($r = -0.312$ to -0.691, $P < 0.05$). Sleep log-derived IIV changes were modestly correlated with mean-level PSQI and PSG-based changes at 12 months. Twelve months of moderate-intensity exercise reduced night-to-night fluctuations in self-rated time to fall asleep, and this relationship was independent of mean-level time to fall asleep.

KEYWORDS   intervention, intra-individual variability, physical activity, sleep, sleep-onset latency, unpredictability

INTRODUCTION

Night-to-night fluctuations in sleep, or intra-individual variability (IIV), has received limited attention in the sleep literature. IIV is an indicator of unpredictability in sleep

Correspondence: Matthew P. Buman, PhD, Stanford Prevention Research Center, Stanford University School of Medicine, Medical School Office Building, 251 Campus Drive, Stanford, CA 94305-5411, USA. Tel.: 650-723-3342; fax: 650-725-6247; e-mail: mbuman@stanford.edu

(Vallieres *et al.*, 2005a), often accompanies insomnia (Levitt *et al.*, 2004; McCrae *et al.*, 2003; Vallieres *et al.*, 2005a), and may be a primary reason individuals seek treatment for poor sleep (Espie, 1991). Elevated sleep IIV has important implications for health and well-being as well. Early cross-sectional studies indicated individuals with highly variable sleep–wake schedules had poorer sleep architecture, physiological arousal, psychomotor performance and subjective mood states compared with individuals with more regular sleep schedules (Taub, 1978, 1981). Nightly fluctuations in sleep have also

been shown to impact daily fluctuations in mood states, cognitive function and physical health complaints (McCrae *et al.*, 2008; Totterdell *et al.*, 1994). Further still, elevated sleep IIV appears to be highly prevalent, even among early middle-aged adults without sleep disorders, where individuals fluctuated more night-to-night (i.e. IIV) in sleep-onset latency (SOL) than was observed between individuals (Knutson *et al.*, 2007).

Pharmacological or non-pharmacological interventions reporting changes in sleep IIV have been rare. Espie *et al.* (1989) compared 8 weeks of three behavioral conditions (tension-release progressive relaxation, stimulus control or paradoxical intention) with placebo and no treatment control conditions in physician-referred chronic insomniacs. The greatest improvements in sleep IIV were for self-rated SOL and total sleep time, and were specific to the stimulus control and paradoxical intention groups. These changes were maintained after a 17-month follow-up. Edinger *et al.* (1992) found similar results in older adults with chronic insomnia in response to 4 weeks of cognitive-behavioral therapy (CBT), but not 4 weeks of relaxation therapy. Sleep IIV reductions were found for self-rated SOL, total sleep time, time in bed (TIB) and sleep efficiency, and an objective measure of nocturnal awakenings. Finally, in a small multiple-baseline study of adults with primary insomnia combining medication with CBT (Vallieres *et al.*, 2005b), the authors noted reductions in sleep efficiency IIV in response to conditions involving CBT, but not conditions of medication alone. These observations were not tested statistically, due likely to small sample sizes.

While these preliminary studies indicate that CBT (either singly or combined) may be effective at reducing sleep IIV, other behavioral treatments may also be effective. Physical exercise appears to be a prime treatment candidate given its modest, yet significant impact on subjective and objective sleep parameters (Buman and King, in press; Youngstedt, 2005; Youngstedt *et al.*, 1997). Additionally, exercise impacts numerous other physical and mental health outcomes simultaneously (Physical Activity Guidelines Advisory Committee, 2008), along with its relative ease of delivery in the community in relation to traditional treatments for poor sleep (Youngstedt, 2005).

There are a number of potential mechanisms whereby exercise produces mean-level sleep improvements that may also account for changes in night-to-night fluctuations in sleep. First, the restorative function of sleep suggests that during sleep, reparative effects related to daily metabolism and growth hormone-releasing hormone optimally occur (Driver and Taylor, 2000). The increased energy expenditure occurring during exercise and subsequent muscle repair are thought to stimulate this restorative process (Shapiro *et al.*, 1981), and mean-level sleep parameters indicative of non-restorative sleep (i.e. SOL, rated sleep quality, sleep duration and sleep efficiency) are known to improve in response to exercise (King *et al.*, 1997, 2002, 2008; Singh *et al.*, 1997, 2005). This restorative process may also impact IIV in these same sleep parameters through greater predictability of autonomic response. Second, thermoregulatory effects (i.e. temperature elevation) following an acute bout of exercise may prompt

deeper forms of sleep through improved temperature downregulation (Van Someren, 2000; Youngstedt, 2005). From an acute perspective, it is unlikely, however, that exercise would impact bed time body temperatures unless exercise was performed less than 4 h prior to sleep (Youngstedt *et al.*, 1997). However, Youngstedt (2005) has suggested that chronic exercise could have a conditioning effect on the temperature downregulation process that would result in improved and more consistent sleep, regardless of the time of day when exercise was performed. Finally, the regular scheduling aspects attendant with many forms of exercise may also represent an important behavior that increases consistency in daytime behaviors and routines (i.e. daily social rhythms or lifestyle regularity; Monk *et al.*, 1990), and could consequently 'carry over' into reduced IIV in sleep behaviors and schedules. If true, this would appear to mimic the effects of other behavioral treatments that stress consistency in sleep–wake schedules (Morin *et al.*, 1999).

While a number of studies have documented that elevated sleep IIV is associated cross-sectionally with poor subjective (Rowe *et al.*, 2008b) and objective sleep (Rowe *et al.*, 2008a; Van Hilten *et al.*, 1993), to our knowledge the extant literature has not addressed if IIV and mean-level sleep (either objective or subjective) respond similarly to treatment. Moreover, IIV and mean-level changes in the same sleep parameters are seldom reported together, making it difficult to understand how night-to-night fluctuations may be influenced by mean-levels, and *vice versa*. Borrowing from other health domains, measures of variability and modification in variability has resulted in similar levels and changes in analogous mean-level parameters. For example, heart rate variability (HRV) is known to be closely related to mean-level heart rate (i.e. pulse), and both are improved through regular exercise. However, HRV has additional predictive value as an indicator of parasympathetic and sympathetic activity (Lahiri *et al.*, 2008), and increased HRV through chronic exercise has resulted in improvements in cardiovascular disease risk factors, such as hypertension, diabetes and cholesterol (Thayer *et al.*, in press).

The purpose of this study was two-fold. The primary research question was whether 12 months of moderate-intensity exercise designed to meet public health recommendations (Physical Activity Guidelines Advisory Committee, 2008) would reduce night-to-night fluctuations in self-rated sleep in community-recruited middle-aged and older adults with mild to moderate sleep complaints. A secondary question was whether observed changes in night-to-night fluctuations were correlated with, or reasonably independent of, mean-level changes in the same sleep outcomes as well as other subjective and objective sleep parameters.

MATERIALS AND METHODS

Design

The study was a 12-month randomized controlled trial of sedentary adults aged 55 years or older with chronic mild to

moderate sleep complaints who were recruited from the community at large. The study methods are described in detail elsewhere (King *et al.*, 2008) and are summarized here. The primary focus of the exercise program was on increasing moderate-intensity endurance exercise to a level that meets or exceeds public health recommendations (Physical Activity Guidelines Advisory Committee, 2008). Exercise intervention participants were instructed to attend exercise classes 2 days per week for 60 min (30–45 min of which was aimed at moderate-intensity endurance exercise, including brisk walking and aerobic movement) and home-based exercise an additional 3 days per week for 30 min throughout the 12-month intervention period. The endurance exercise was targeted at an intensity of 60–85% of treadmill-based peak heart rate. All exercise classes were scheduled during mornings or afternoons, and participants were instructed to complete their home-based exercise during this time as well. Control arm participants received weekly health education classes for 90 min focusing on non-exercise age-related topics, such as nutrition, home safety and foot care. Both arms received brief standard sleep hygiene that included a recommendation for establishing a regular sleep–wake cycle.

Potential enrollees responded to community-based health promotion recruitment strategies focused on enhancing overall health-related quality of life (and not specifically focused on sleep quality or physical activity). Recruitment was delivered through local media outlets. The primary eligibility criteria included: (a) age 55 years or older; (b) underactive (defined as < 60 min per week of moderate or more vigorous physical activity over the previous 6 months); (c) body mass index ≤35; (d) free of sleep apnea, defined by a multivariate apnea prediction score ≤0.8 (Maislin *et al.*, 1996), and overnight pulse oximetry standardized criteria for oxygen saturation (< 10% cumulative time at < 90% SaO_2) and desaturation index (> 3% falls in SaO_2 per sleep hour, < 10 per hour; George *et al.*, 1988; Series *et al.*, 1993); and (e) mild to moderate sleep complaints (defined by scores ≥3 on at least two of three items: focused on getting to sleep; waking up during the night; and waking up in the morning of the Sleep Questionnaire and Assessment of Wakefulness; Miles, 1982). The major study results are described in detail elsewhere (King *et al.*, 2008). The study protocol was approved by the appropriate university institutional review boards.

Measurement

Sleep logs

The five primary sleep outcomes for this study were derived from sleep logs. Sleep logs were used to monitor self-rated sleep (Lichstein *et al.*, 2004) for 14 days at baseline, and 7 days at 6 and 12 months after baseline. This measurement period has been found to be adequate to obtain reliable estimates of IIV in sleep (Rowe *et al.*, 2008b). Sleep logs were collected following each measurement period and verified for completeness by research staff. Participants recorded daily ratings of illness,

pain, exercise, medication use and alcohol/caffeine use prior to sleep, and ratings of nighttime sleep quality upon awakening in the morning. Five sleep–wake parameters were computed from the daily logs: (a) SOL in minutes (i.e. 'How long did it take you to fall asleep last night?'); (b) TIB in hours, computed as (time out of bed – time in bed); (c) feeling rested in the morning [rested AM; i.e. 'How rested do you feel this morning?' using a nine-point Likert-type scale from 0 (not at all rested) to 8 (extremely rested)]; (d) number of awakenings (i.e. 'How many times last night did you wake up?'); and (e) wake after final awakening (WAFA) in minutes (time out of bed – wake time).

Individual-level means and standard deviations at each time point were obtained, and a mean-referenced IIV coefficient was calculated as follows: SD/mean × 100%. This metric represents the ratio of IIV in a variable to its individual-level mean. The IIV coefficient is considered a more conservative measure of variability, relative to a raw IIV score, as it adjusts for the greater likelihood of high variability in persons with high individual-level mean scores and the influence of mean-level changes on changes to raw individual standard deviations. Recent studies addressing sleep IIV have advocated for the IIV coefficient as the preferred measure of variability in sleep research (Knutson *et al.*, 2007; Rowe *et al.*, 2008a; Van Hilten *et al.*, 1993).

In-home polysomnographic (PSG) sleep recordings

Nine-channel PSG recordings were performed in participants' homes using the Oxford Medilog MR95 digital recording system (Oxford Instruments, Oxford, UK) and consisted of the following channels: three channels of electroencephalography (C4-A1, C3-A2, Pz-A1), one channel of bipolar electrooculography, one channel of surface mentalis electromyography (EMG), one channel of modified lead II electrocardiography, two channels (left and right, recorded unilaterally) of surface anterior tibialis EMG, and one channel of ambient light, recorded with the MR95 light sensor. In-home PSG was measured for three nights at baseline, and two nights at 6 and 12 months. Data were averaged for the second and third nights at baseline, and the two nights at 6 and 12 months to avoid potential first-night effects (Woodward *et al.*, 1996). A full explanation of the PSG data collection and scoring methods are presented in detail elsewhere (King *et al.*, 2008). PSG results presented in the main outcome paper (King *et al.*, 2008) indicated that at 12 months exercisers showed: (a) significantly lower percentages of time in Stage 1 sleep relative to controls; (b) significantly greater percentages of time in Stage 2 sleep relative to controls; and (c) significantly fewer awakenings during the first third of the sleep period relative to controls. These PSG mean results were used to examine correlated change with the IIV effects.

Pittsburgh Sleep Quality Index (PSQI)

The PSQI was used as a secondary self-rated sleep outcome, and was administered on a single occasion at baseline, 6 months

and 12 months (Buysse *et al.*, 1991). The PSQI is made up of seven 'component' scores (using a 0–3 scale) and a summary global sleep quality score (range = 0–21; lower score = better sleep quality). Study results on the PSQI are reported elsewhere (King *et al.*, 2008), with significant intervention effects for the sleep disturbance subscale and marginal effects for the global sleep quality score. These two subscores were used in the analyses here to examine correlated change with the IIV effects.

Analyses

An initial step in IIV analyses is to examine how much variance can be attributed at the between- and intra-individual levels. Between-subject variation (i.e. the standard deviation between persons after averaging the sleep outcome across days) and within-subject variation (i.e. standard deviation within each person across days) were compared at baseline. In addition, intra-class correlation coefficients (ICC) were computed for all sleep outcomes at baseline to estimate between-subject variability relative to total variability. These types of analyses are an important step in IIV analyses to assess whether the level of within-person variability is substantial enough to warrant study (Bryk and Raudenbush, 1992; McCrae *et al.*, 2008).

To address our primary research question concerning whether IIV in the sleep outcomes of interest changed in response to moderate-intensity exercise, analysis of covariance (ANCOVA) was conducted with baseline values as covariates, and 6-month and 12-month outcomes examined in separate models. Pearson correlations were computed for the residualized IIV change scores (baseline to 12 months) for all variables to address our secondary questions concerning whether IIV changes were correlated with mean-level changes in each sleep log outcome of interest as well as the previously described PSG and PSQI outcomes. Finally, additional exploratory ANCOVA analyses were performed to determine whether mean-level or IIV changes occurred in bed time or wake time at 6 months or 12 months due to exercise. This assessed one of our hypothesized mechanisms concerning whether exercise produced more regular and consistent sleep behaviors that may also be associated with improved consistency in self-rated sleep outcomes.

Intent-to-treat principles were used for all sleep outcomes (IIV and mean-level sleep log, PSG and PSQI variables) such that when data were missing at 6 or 12 months, baseline values were used. Effect size estimates were calculated for all IIV outcomes. Magnitude of effect size estimates were based on established conventions (Cohen, 1988). Alpha was set at $P < 0.05$ and all analyses were two-tailed. The analyses were carried out using SAS Enterprise guide version 4.1 (SAS Institute, Cary, NC, USA).

RESULTS

Participant and data completion descriptions

Of 201 persons initially responding to the study promotional announcements, 66 individuals [44 women, 22 men; M (SD) age = 61.42 (6.72)] were eligible and randomized (36 intervention, 30 control). Twenty-four of the initial 201 respondents (11.9%) were excluded based on the oximetry/apnea screen. Eighty-nine percent completed the 12-month trial (32 intervention, 27 control), with no significant differences in dropout rates by study arm. There were no baseline group-level differences for demographic variables (age, gender, education, race/ethnicity), over-the-counter sleep medication use, physical activity or fitness levels, body mass index, or mean-level sleep diary, PSG or PSQI outcomes (*P*-values > 0.10). Adherence to the exercise prescription was good; participants attended 74% of exercise classes, reported significantly greater energy expenditure at 12 months relative to control, and significantly improved cardiorespiratory fitness at 12 months relative to control (King *et al.*, 2008). Health education control attendance was 80%.

Sleep log completion was high for each time point. The mean sleep log completion rate was 13.21 (SD = 2.63; out of 14) days at baseline, 6.20 (SD = 2.13; out of 7) days at 6 months, and 6.29 (SD = 2.19; out of 7) days at 12 months. Out of 9240 possible data points (5 sleep log variables × 28 days × 66 participants), only 160 were missing (1.7%). Two individuals (1 intervention, 1 control) had partially completed logs at baseline that precluded IIV-level calculations for TIB and WAFA. The PSG data were complete for 88 and 89% of participants at 6 and 12 months, respectively. The PSQI data were complete for 92 and 89% of participants at 6 and 12 months, respectively.

Relative amounts of IIV and mean-level variability at baseline

The ratio of IIV (i.e. within-person) to between-person variability is a useful index for assessing the relative magnitude of variance from each of these two sources for a given outcome. For all five sleep log outcome variables at baseline, the ratio of within-person variability to between-person variability was approximately equal to 1 or > 1: SOL = 0.93; TIB = 1.46; rested AM = 1.16; number of awakenings = 0.92; WAFA = 1.17). Similarly, ICC estimates the ratio of between-person variance to total variance. ICC calculations showed that 15–35% of the total variability in the five outcome variables occurred at the between-person level, leaving 65–85% as within-person variability. These initial analyses thus revealed that there was a significant amount of IIV that was worthy of study and not typically accounted for in traditional mean-level analyses.

Changes in 6- and 12-month IIV sleep outcomes

Because there were 14 consecutive days of sleep logs collected at baseline (compared with 7 days at 6 months and 12 months), we conducted preliminary analyses to explore whether there were IIV differences between the first (i.e. Days 1–7) and second (i.e. Days 8–14) baseline weeks. Our analyses suggested there were no observable differences ($P > 0.10$). All subsequent analyses therefore combined baseline weeks of

sleep logs, as 14 consecutive days of measurement have been found to produce more stable and conservative estimates of IIV than 7 consecutive days (Rowe *et al*., 2008a).

We conducted ANCOVA analyses to determine if the exercise intervention significantly reduced IIV for the five self-rated (log-based) sleep outcomes. Using intent-to-treat methods, Table 1 presents group means and effect sizes at 6 months and 12 months for the five IIV outcomes. The exercise group showed significantly less variability at 12 months relative to the control group for SOL ($F_{2,63}$ = 5.30, P = 0.025; group difference = 23.11, 95% CI, 3.04–47.18). No other statistically significant group differences were found at 6 or 12 months. Effect size estimates indicated the following: (a) moderate effect sizes for SOL at 6 and 12 months; (b) small effects for number of awakenings and WAFA at 6 months; and (c) small effects for TIB and feeling rested in the AM at 12 months. Fig. 1 shows the individual change scores for the raw IIV data for SOL at 12 months (a). For the exercise and control groups, 81 and 33% reduced their variability in SOL, respectively. On average, the exercise group reduced their night-to-night variability in SOL by 10 min, while the control group increased by 6 min.

Relationships between mean-level changes and IIV changes in sleep outcomes of interest

Table 2 displays interrelationships between 12-month residualized change in mean-level and IIV sleep parameters. The strongest relationships tended to be observed among mean-level/IIV pairs (on the diagonal). IIV in feeling rested in the AM and number of nighttime awakenings also tended to be correlated with mean-level changes in other sleep log measures. Perhaps of greatest interest was that SOL-based IIV (the only

sleep log variable for which a significant IIV reduction was observed) was not correlated with mean-level changes in SOL. Fig. 1 corroborates this correlation by comparing IIV change (a) to mean-level change (b) in SOL for each participant individually. A careful examination of this figure indicates that for most participants in the exercise group, the direction of change in SOL was similar for IIV and mean-level variables; however, the magnitude of change appeared unrelated. For the control group, both direction of change and magnitude of change were similar for IIV and mean-level changes for most participants.

Relationships between IIV changes and PSG and PSQI sleep outcomes

Table 3 displays interrelationships between 12-month residualized change in IIV sleep outcomes from the sleep logs and those PSG and PSQI outcomes that were significantly improved by the exercise intervention relative to controls (King *et al*., 2008). A direct correlation was observed between change in IIV for rested in the AM and the PSQI global sleep quality score, indicating that as variability in feeling rested in the morning increased, global sleep quality worsened (lower score = better sleep quality).

Changes in self-reported sleep–wake schedules as a potential mechanism underlying IIV changes

The final analysis addressed the question of whether exercise produced more regular and consistent sleep behaviors, evidenced by more consistent sleep–wake scheduling (i.e. bed time, wake time). Using intent-to-treat, the ANCOVA analyses revealed no significant changes in IIV in bed time or wake time

Table 1 IIV coefficient in sleep diary outcomes at baseline, 6 months, and 12 months, sample size (N), mean (SD), effect size (ES) by study arm								
			Baseline	*6 months*		*12 months*	
Sleep diary outcomes	*N*	*Study Arm*	*M (SD)*	*M (SD)*	*ES*	*M (SD)*	*ES*
SOL	36	Exercise	71.80 (44.90)	53.61 (33.47)	0.41	56.57 (41.90)*	0.57
	30	Control	79.03 (37.04)	67.31 (27.02)		83.38 (49.34)	
TIB	35	Exercise	12.08 (6.50)	12.44 (6.97)	−0.29	10.39 (6.27)	0.18
	29	Control	10.76 (7.32)	10.69 (7.00)		9.82 (4.34)	
Rested AM	36	Exercise	36.35 (15.65)	32.44 (23.17)	−0.03	29.62 (21.02)	0.14
	30	Control	37.03 (15.64)	32.21 (15.01)		32.38 (13.15)	
Awakenings	35	Exercise	66.11 (52.51)	69.15 (44.58)	0.35	66.78 (46.74)	0.05
	30	Control	71.83 (53.73)	89.54 (68.43)		76.18 (63.16)	
WAFA	35	Exercise	143.15 (79.83)	109.30 (56.61)	0.17	111.60 (50.38)	−0.19
	29	Control	113.48 (64.02)	98.41 (42.04)		96.55 (60.16)	
Sleep-wake schedule							
Bed time	35	Exercise	3.05 (1.38)	3.07 (2.16)	0.15	2.77 (1.83)	0.23
	30	Control	3.34 (1.59)	2.98 (1.67)		3.31 (1.79)	
Wake time	36	Exercise	16.69 (9.06)	13.02 (9.95)	0.04	10.98 (5.69)	0.25
	30	Control	16.05 (10.36)	12.42 (7.45)		12.57 (10.02)	

Notes. Intent-to-treat values are imputed. Effect size is Cohen's *d*.
SOL, sleep-onset latency; IIV, intra-individual variability; TIB, time in bed; WAFA, wake after final awakening.
*Between-arm difference (ANCOVA), P < 0.05, two-tailed.

Table 2 Baseline scores in mean-level sleep diary outcomes and correlations of 12-month residualized change (controlling for baseline) in IIV coefficient and mean sleep diary outcomes

		IIV Coefficient				
	Baseline (SD)	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
Mean-level Outcomes						
(1) SOL (min)	31.23 (22.70)	0.062	−0.200	0.395***	−0.143	0.129
(2) TIB (h)	8.16 (0.66)	0.018	−0.315*	0.003	−0.250*	0.074
(3) Rested AM	3.69 (1.11)	0.007	−0.058	−0.691***	0.092	−0.130
(4) Awakenings (no)	2.61 (2.13)	−0.042	0.110	0.305*	−0.312*	−0.054
(5) WAFA (min)	36.78 (32.42)	0.028	0.024	−0.109	0.027	−0.149

Notes. IIV values are on the vertical, means are on the horizontal, IIV/mean pairs are located on the diagonal.
SD, standard deviation; SOL, sleep-onset latency; TIB, time in bed; WAFA, wake after final awakening; IIV, intra-individual variability.
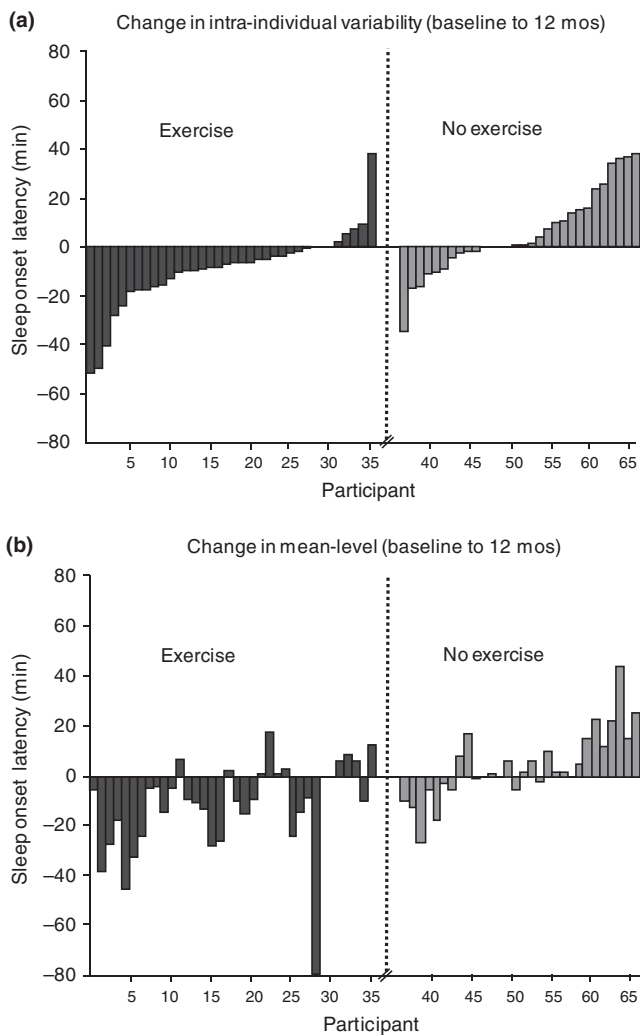*$P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Figure 1. Individual change scores in intra-individual variability (IIV; a) and mean-level (b) sleep onset latency (SOL) from baseline to 12 months by intervention group. Notes: IIV and mean-level changes are in parallel order by participant; Raw IIV change is used in (a) instead of the IIV coefficient to aid in comparison with mean-level change (i.e. retain identical metric).

in the exercise arm relative to the control arm at 6 months or 12 months (Table 1); nor were differences observed for mean-level changes in bed time and wake time.

DISCUSSION

Main findings

The results of this study indicate that 12 months of moderate-intensity exercise at a level equal to or exceeding public health recommendations (Physical Activity Guidelines Advisory Committee, 2008) reduced night-to-night fluctuations in self-rated time to fall asleep in initially underactive mild to moderately sleep-impaired middle-aged and older adults. These changes occurred without concurrent changes in mean-level or IIV of scheduled bed time or wake time. No other significant reductions in fluctuations of self-rated sleep variables were observed, with small effect size estimates found for awakenings and WAFA at 6 months, and TIB and feeling rested in the AM at 12 months.

Importance of exercise and sleep IIV

This research contributes to the limited but growing body of literature suggesting that night-to-night fluctuations in sleep, independent of mean-level sleep, are an important outcome in sleep research in need of further attention. The current results, along with what is already known regarding IIV in sleep, demonstrate this need in two important ways. First, the amount of within-person night-to-night variability was roughly equal to between-subject mean-level variability in the sleep outcomes reported in our study. This has consistently been found with both middle-aged (Knutson *et al.*, 2007) and older adults (Rowe *et al.*, 2008a). For example, Knutson *et al.* (2007) found in the CARDIA study that middle-aged adults showed similar proportions of within- and between-person variability in sleep variables as measured with actigraphy that we found with sleep logs. Traditional analyses of objective and subjective measures of sleep tend to either measure outcomes based on a single night or average multiple nights, both of

Table 3 Baseline scores in objective PSG and subjective PSQI sleep outcomes and correlations of 12-month residualized change (controlling for baseline) in sleep log IIV coefficients and PSG and PSQI outcomes ($N = 66$)

		IIV Coefficient				
	Baseline (SD)	SOL	TIB	Rested AM	Awakenings	WAFA
Polysomnography outcomes						
% Sleep time in Stage 1	8.69 (3.82)	−0.009	0.011	0.157	−0.048	−0.152
% Sleep time in Stage 2	52.95 (8.38)	−0.128	0.053	−0.229	0.019	0.086
Number of Awakenings: 1st third of sleep period	4.03 (2.48)	−0.220	−0.160	0.007	−0.191	−0.117
Pittsburgh Sleep Quality Index						
+Sleep disturbance subscale	1.47 (0.50)	0.134	−0.180	0.075	−0.073	0.181
++Global sleep quality score	8.12 (3.28)	−0.113	0.019	0.381**	−0.210	−0.001

Notes: +0–3 scale, lower score = less disturbance; ++0–21 scale, lower score = better sleep quality.
SD, standard deviation; PSG, polysomography; PSQI, Pittsburg Sleep Quality Index; SOL, sleep-onset latency; TIB, time in bed; WAFA, wake after final awakening; IIV, intra-individual variability.
**$P < 0.01$.

which ignore what appears to be systematic and potentially large amounts of within-person variation. In a meta-analysis of non-pharmacological interventions for insomnia treatment in older adults, Pallesen *et al.* (1998) encouraged future outcome studies to report parameters of IIV. Since their recommendation more than 10 years ago, few studies have reported on such effects.

Second, the current results suggest that exercise-induced effects on sleep IIV were localized to SOL, and not observed for other self-rated sleep outcomes. Interestingly, one of the better designed studies in this area found prospective sequential associations (at the daily level) between subjective SOL and mood changes, where the time to fall asleep for the previous night predicted the following day's mood (Totterdell *et al.*, 1994). McCrae *et al.* (2008) found similar results for a self-reported sleep quality rating and positive and negative effect. Collectively, these results indicate that SOL-based IIV is associated with affective states, and the sequential nature of this relationship raises the potential that SOL-based IIV may be causally linked to affective states. Controlled trials are needed to empirically test this hypothesis. Our study adds a useful contribution for future study of the sleep IIV–effect relationship, and perhaps other important health outcomes as well, by demonstrating an efficacious, non-pharmacological alternative (exercise) to CBT for systematically reducing SOL-based IIV in future studies.

Finally, the analyses indicated that the vast majority of the exercise participants (and a minority of the control participants) reduced their SOL-based IIV, and on average the exercise participants' reduction was approximately 21% (i.e. 10 min). This reduction is lower than what has been reported in other non-pharmacological trials, where Espie *et al.* (1989) reported 70% and 55% reduction in IIV following 8 weeks of paradoxical intention and stimulus control, respectively. Likewise, Edinger *et al.* (1992) reported a 41% reduction in IIV following 4 weeks of CBT. Differences between these studies and the current study are worth noting. First, both of these studies focused on chronic insomniacs who at baseline had much higher mean- and IIV-levels of sleep latency.

Additionally, both of these studies reported raw changes to individual standard deviations, and therefore did not adjust for concurrent mean-level changes in their analyses. By using the IIV coefficient, where the IIV outcome is in reference to changes in the mean, we took a more conservative approach to ensure that our conclusions were made free of influence of mean-level changes. If raw values would have been used in our analyses, we would have shown significant reductions in IIV-based SOL at 6 months, with a 40% reduction in the exercise group by 12 months.

Potential mechanisms of exercise-induced effects on IIV

There are a number of potential mechanisms that could account for changes in SOL-based IIV in response to regular exercise. Given that the exercise intervention focused on pre-scheduling of exercise on a daily basis, this may have in turn aided the establishment of a more consistent schedule of activities in other portions of the day, including sleep schedules, similar to recommendations given for sleep hygiene and stimulus control (Morin *et al.*, 1999). This hypothesis, however, was not supported by the data, as log reports of bed time and wake time schedules were not changed in response to exercise. This may be due, at least in part, to observations suggested in the literature that older adults as a group may exhibit less IIV in habitual sleep timing compared with younger adults, suggesting a potential floor effect (Monk *et al.*, 1991). This was evidenced by similar mean-levels and variations for bed time and wake time compared with the older adults in the report by Monk *et al.* (1991), as well as observably low baseline IIV coefficients for bed time and wake time relative to the other sleep diary outcomes. It also should be noted that timing and scheduling represent only one aspect of behavioral routines. Exercise is also likely to impact additional contexts related to routines, including exercise setting (e.g. active transport, nature hiking, indoor gym facility) and actions preceding or following exercise (e.g. meal-timing, work–leisure balance) that could impact biological rhythms and sleep behaviors. Future research examining

IIV changes should comprehensively document accompanying changes in consistencies of daytime behaviors and routines, such as the Social Rhythm Metric (Monk *et al.*, 1990, 2002), to further explore this potential mechanism.

In addition to this behavioral mechanism, other biological mechanisms can be postulated that apply more broadly to the exercise–sleep relationship, but perhaps have special considerations for SOL-based IIV in particular – body temperature regulation and autonomic response. Recent literature suggests that more rapid declines in core temperature (Driver and Taylor, 2000; Murphy and Campbell, 1997), and even skin temperature (Raymann *et al.*, 2007), can trigger sleep onset. In regards to exercise, Youngstedt *et al.* (1997) concluded in a meta-analysis that the acute effect of exercise is unlikely to alter body temperature during sleep unless exercise was performed less than 4 h prior to sleep. This conclusion likely applies to our study as well, as participants were instructed to complete exercise during morning or early afternoon hours. Youngstedt (2005) has suggested, however, that chronic exercise could improve efficiency in this temperature downregulation over time. This notion is supported by other research suggesting that autonomic response (i.e. HRV) can be improved with chronic exercise (Lahiri *et al.*, 2008), and persons with disturbed sleep have blunted temperature decline at night adversely affecting sleep onset (Lack *et al.*, 2008). Our results are also in line with this hypothesis. If these effects were due solely to the acute lowering of body temperature derived from the exercise-induced elevation of temperature or acute changes to HRV, then one might expect equal or even greater variability at 6 months and 12 months relative to baseline, given that improved SOL would be expected to occur only on exercise days during the measurement period. The fact that SOL-based IIV decreased overall during these periods suggests a conditioning effect whereby the thermoregulatory process became more efficient at temperature downregulation and/or greater autonomic control triggered sleep onset more rapidly and consistently. Future studies should formally test these hypotheses by examining changes in the temperature downregulation process and HRV in response to chronic exercise, and whether these changes directly mediate impacts on SOL-based sleep IIV.

Our results that log-based IIV changes were uncorrelated with changes in mean-level PSG outcomes were not surprising given that subjective and objective measures of sleep are often observed to be unrelated in middle-aged and older adults (Buysse *et al.*, 1991; Vitiello *et al.*, 2004). This may be especially true for our sample as differences in subjective and objective measures are amplified among older and impaired sleepers compared with younger and better sleepers (Espie, 1991; Means *et al.*, 2003). Also, because only 2 nights of PSG measurements were used to estimate mean-level outcomes, it is unclear how this may have impacted correlations with 14 consecutive nights (and 7 nights at 6 months and 12 months) of IIV sleep log measure. It is important to note, however, that this lack of relationship with objective outcomes does not undermine the importance of these findings. Of note, McCrae

et al. (2008) found, using sleep logs, daily associations between IIV for subjective sleep quality and total wake time (which included SOL) and affective states, but not for wrist actigraphy-derived IIV and affective states. Espie (1991) contends that perceptions of nightly fluctuations in sleep likely drive many complaints related to poor sleep and insomnia, and may result in poorer quality of life and increased health care visits in older adults (Foley *et al.*, 1995). Nevertheless, an important next step would be to study whether exercise is able to reduce objectively measured sleep IIV as well, especially given that elevated levels of IIV have been shown in older and poorer sleepers in both wrist actigraphy (Rowe *et al.*, 2008b) and PSG (Edinger *et al.*, 1991). This would require multiple consecutive days of objective measurement of at least 7 days, but preferably 14 days based upon reliability estimates for sleep IIV (Rowe *et al.*, 2008b), along with potentially larger sample sizes than those available in the present study.

Strengths and limitations

There are a number of strengths and limitations of this study worth noting. First, our sample of middle-aged and older adults with moderate sleep complaints represents an excellent population to study IIV in sleep, given: (a) higher levels of IIV relative to younger and better sleepers; (b) the likelihood that such complaints may lead to other health complaints and medical costs; and (c) the salutary benefits of exercise in this population. This does limit the generalizability of our results to this population, however, as other exercise–sleep researchers have noted, the overall effects of exercise on sleep are likely underestimated given that most studies include younger and less impaired sleepers (Youngstedt, 2003). Indeed, Youngstedt (2003) concluded that after controlling for baseline sleep characteristics, exercise has a relatively equal impact on mean-level sleep outcomes as does hypnotics. Such comparisons currently are more difficult to make for sleep IIV given that few pharmacological or non-pharamacological interventions to date report IIV effects.

A second study constraint relates to the fact that we were unable to fully control for light exposure during exercise and the time of day exercise occurred. While at least two exercise sessions were supervised and completed under indoor domestic light conditions, and participants were instructed to avoid nighttime exercise, precise data are not available regarding when and under what light conditions exercise typically occurred. Despite this important limitation, it should be noted that the single and combined effects of light exposure and exercise on sleep remain understudied (Guilleminault *et al.*, 1995; Youngstedt *et al.*, 2002), and controlling such factors potentially limits the external validity and public health impact of the exercise prescription. Also, light exposure and time of day effects would be expected to produce acute exercise-related effects on sleep. While it is plausible that these acute effects could account for at least some of the SOL-based IIV changes observed, reduced IIV may also require chronic effects to occur in order to improve night-to-night consistency, regardless of

where exercise occurred during the previous day. This issue deserves further examination in future studies.

Our study may not have been adequately powered to detect the reasonably small effects of exercise on IIV in total sleep time, feeling rested in the AM, number of awakenings and sleep efficiency. Also, it should be noted that 12 months of moderate-intensity exercise was needed to observe significant IIV changes. The main objectives of this trial were mean-level objective and subjective sleep parameters, and not necessarily variability in these outcomes. Because there are no known intervention studies that report changes in IIV, the reporting of these IIV effects in response to exercise, and the length of time needed to observe such effects, represents a valuable contribution to the literature and could serve as useful effect size estimates to guide future trials in this area. Finally, because sleep logs in this study were collected after each measurement period, it is possible that bias was introduced; however, this method for collecting sleep log data is consistent with other studies in the behavioral sleep literature.

CONCLUSIONS

Moderate-intensity exercise at a level equal to or exceeding public health recommendations (Physical Activity Guidelines Advisory Committee, 2008) reduces night-to-night fluctuations in self-rated time to fall asleep (SOL) in mild to moderately sleep-impaired middle-aged and older adults. These changes appear to function independent of mean-level changes in SOL and without concurrent changes in wake time or bed time schedules. In light of the effects of regular moderate-intensity exercise on a wide range of health parameters, and the relationships between sleep IIV and sleep complaints as well as negative health outcomes, continued investigation of the potential benefits of exercise on sleep variability is warranted.

ACKNOWLEDGEMENTS

REFERENCES

Bryk, A. S. and Raudenbush, S. W. *Hierarchical Linear Models for Social and Behavioral Research: Applications and Data Analysis Methods*. Sage, Newbury Park, CA, 1992.

Buman, M. P. and King, A. C. Exercise as treatment to enhance sleep. *Am. J. Lifestyle Med.*, in press.

Buysse, D. J., Reynolds, C. F., Monk, T. H., Hoch, C. C., Yaeger, A. L. and Kupfer, D. J. Quantification of subjective sleep quality in health elderly men and women using the Pittsburgh Sleep Quality Index (PSQI). *Sleep*, 1991, 14: 331–338.

Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. Erlbaum, Hillsdale, NJ, 1988.

Driver, H. S. and Taylor, S. R. Exercise and sleep. *Sleep Med. Rev.*, 2000, 4: 387–402.

Edinger, J. D., Marsh, G. R., Mccall, W. V., Erwin, C. W. and Lininger, A. W. Sleep variability across consecutive nights of home monitoring in older mixed DIMS patients. *Sleep*, 1991, 14: 13–17.

Edinger, J. D., Hoelscher, T. J., Marsh, G. R., Lipper, S. and Ionescu-Pioggia, M. A cognitive-behavioral therapy for sleep-maintenance insomnia in older adults. *Psychol. Aging*, 1992, 7: 282–289.

Espie, C. A. The Psychological Treatment of Insomnia. J. Wiley, Chichester, UK, 1991.

Espie, C. A., Lindsay, W. R., Brooks, D. N., Hood, E. M. and Turvey, T. A controlled comparative investigation of psychological treatments for chronic sleep-onset insomnia. *Behav. Res. Ther.*, 1989, 27: 79–88.

Foley, D. J., Monjan, A. A., Brown, S. L., Simonsick, E. M., Wallace, R. B. and Blazer, D. G. Sleep complaints among elderly persons – an epidemiological study of 3 communities. *Sleep*, 1995, 18: 425–432.

George, C. F., Millar, T. W. and Kryger, M. H. Identification and quantification of apneas by computer-based analysis of oxygen-saturation. *Am. Rev. Resp. Dis.*, 1988, 137: 1238–1240.

Guilleminault, C., Clerk, A., Black, J., Labanowski, M., Pelayo, R. and Claman, D. Nondrug treatment trials in psychophysiologic insomnia. *Arch. Intern. Med.*, 1995, 155: 838–844.

King, A. C., Oman, R. F., Brassington, G. S., Bliwise, D. L. and Haskell, W. L. Moderate-intensity exercise and self-rated quality of sleep in older adults. A randomized controlled trial. *JAMA*, 1997, 277: 32–37.

King, A. C., Baumann, K., O'Sullivan, P., Wilcox, S. and Castro, C. Effects of moderate-intensity exercise on physiological, behavioral, and emotional responses to family caregiving: a randomized controlled trial. *J. Gerontol. A Biol. Sci. Med. Sci.*, 2002, 57: M26–M36.

King, A. C., Pruitt, L. A., Woo, S. *et al.* Effects of moderate-intensity exercise on polysomnographic and subjective sleep quality in older adults with mild to moderate sleep complaints. *J. Gerontol. A Biol. Sci. Med. Sci.*, 2008, 63: 997–1004.

Knutson, K. L., Rathouz, P. J., Yan, L. J. L., Liu, K. and Lauderdale, D. S. Intra-individual daily and yearly variability in actigraphically recorded sleep measures: the CARDIA study. *Sleep*, 2007, 30: 793–796.

Lack, L. C., Gradisar, M., Van Someren, E. J. W., Wright, H. R. and Lushington, K. The relationship between insomnia and body temperatures. *Sleep Med. Rev.*, 2008, 12: 307–317.

Lahiri, M. K., Kannankeril, P. J. and Goldberger, J. J. Assessment of autonomic function in cardiovascular disease. *J. Am. Coll. Cardiol.*, 2008, 51: 1725–1733.

Levitt, H., Wood, A., Moul, D. E. *et al.* A pilot study of subjective daytime alertness and mood in primary insomnia participants using ecological momentary assessment. *Behav. Sleep Med.*, 2004, 2: 113–131.

Lichstein, K. L., Durrene, H. H., Riedel, H., Taylor, D. J. and Bush, A. J. *Epidemiology of Sleep: Age, Gender, and Ethnicity*. Lawrence Erlbaum, Mahwah, NJ, 2004.

Maislin, G., Dinges, D. F., Pack, F. M. and Pack, A. I. Validity study of the Multivariate Apnea Index (MAP) in an elderly population. *Sleep Res.*, 1996, 25: 288.

McCrae, C. S., Wilson, N. M., Lichstein, K. L. *et al.* 'Young old' and 'old old' poor sleepers with and without insomnia complaints. *J. Psychosom. Res.*, 2003, 54: 11–19.

McCrae, C. S., Mcnamara, J. P., Rowe, M. A. *et al.* Sleep and affect in older adults: using multilevel modeling to examine daily associations. *J. Sleep Res.*, 2008, 17: 42–53.

Means, M. K., Edinger, J. D., Glenn, D. M. and Fins, A. I. Accuracy of sleep perceptions among insomnia sufferers and normal sleepers. *Sleep Med.*, 2003, 4: 285–296.

Miles, L. E. A sleep questionnaire. In: C. Guilleminault (Ed.) *Sleeping and Waking Disorders: Indications and Techniques*. Addison-Wesley, Menlo Park, CA, 1982: 383–413.

Monk, T. H., Flaherty, J. F., Frank, E., Hoskinson, K. and Kupfer, D. J. The social rhythm metric – an instrument to quantify the daily rhythms of life. *J. Nerv. Ment. Dis.*, 1990, 178: 120–126.

Monk, T. H., Reynolds, C. F., Buysse, D. J. *et al.* Circadian characteristics of healthy 80-year-olds and their relationship to objectively recorded sleep. *J. Gerontol.*, 1991, 46: M171–M175.

Monk, T. H., Frank, E., Potts, J. M. and Kupfer, D. J. A simple way to measure daily lifestyle regularity. *J. Sleep Res.*, 2002, 11: 183–190.

Morin, C. M., Hauri, P. J., Espie, C. A., Spielman, A. J., Buysse, D. J. and Bootzin, R. R. Nonpharmacologic treatment of chronic insomnia. *Sleep*, 1999, 22: 1134–1156.

Murphy, P. J. and Campbell, S. S. Nighttime drop in body temperature: a physiological trigger for sleep onset? *Sleep*, 1997, 20: 505–511.

Pallesen, S., Nordhus, I. H. and Kvale, G. Nonpharmacological interventions for insomnia in older adults: a meta-analysis of treatment efficacy. *Psychotherapy*, 1998, 35: 472–482.

Physical Activity Guidelines Advisory Committee. Physical Activity Guidelines Advisory Committee Report. In: US Department of Health and Human Services, Washington, DC, 2008.

Raymann, R., Swaab, D. F. and Van Someren, E. J. W. Skin temperature and sleep-onset latency: changes with age and insomnia. *Physiol. Behav.*, 2007, 90: 257–266.

Rowe, M., McCrae, C., Campbell, J. *et al.* Actigraphy in older adults: comparison of means and variability of three different aggregates of measurement. *Behav. Sleep Med.*, 2008a, 6: 127–145.

Rowe, M. A., McCrae, C. S., Campbell, J. M., Benito, A. P. and Cheng, J. Sleep pattern differences between older adult dementia caregivers and older adult noncaregivers using objective and subjective measures. *J. Clin. Sleep Med.*, 2008b, 4: 362–369.

Series, F., Marc, I., Cormier, Y. and Laforge, J. Utility of noctural home oximetry for case-finding in patients with suspected sleep-apnea hypopnea syndrome. *Ann. Int. Med.*, 1993, 119: 449–453.

Shapiro, C. M., Bortz, R., Mitchell, D., Bartel, P. and Jooste, P. Slow-wave sleep: a recovery period after exercise. *Science*, 1981, 214: 1253–1254.

Singh, N. A., Clements, K. M. and Fiatarone, M. A. Sleep, sleep deprivation, and daytime activities – a randomized controlled trial of the effect of exercise on sleep. *Sleep*, 1997, 20: 95–101.

Singh, N. A., Stavrinos, T. M., Scarbek, Y., Galambos, G., Liber, C. and Fiatarone Singh, M. A. A randomized controlled trial of high versus low intensity weight training versus general practitioner care for clinical depression in older adults. *J. Gerontol. A Biol. Sci. Med. Sci.*, 2005, 60: 768–776.

Taub, J. M. Behavioral and psychophysiological correlates of irregularity in chronic sleep routines. *Biol. Psychol.*, 1978, 7: 37–53.

Taub, J. M. Nocturnal electrographic features of frequently changing irregular sleep-wakefulness rhythms. *Int. J. Neurosci.*, 1981, 14: 227–237.

Thayer, J. F., Yamamoto, S. S. and Brosschot, J. F. The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors. *Int. J. Cardiol.*, 2010, 141: 122–131 .

Totterdell, P., Reynolds, S., Parkinson, B. and Briner, R. B. Associations of sleep with everyday mood, minor symptoms and social interaction experience. *Sleep*, 1994, 17: 466–475.

Vallieres, A., Ivers, H., Bastien, C. H., Beaulieu-Bonneau, S. and Morin, C. M. Variability and predictability in sleep patterns of chronic insomniacs. *J. Sleep Res.*, 2005a, 14: 447–453.

Vallieres, A., Morin, C. M. and Guay, B. Sequential combinations of drug and cognitive behavioral therapy for chronic insomnia: an exploratory study. *Behav. Res. Ther.*, 2005b, 43: 1611–1630.

Van Hilten, J. J., Braat, E. M., Van Der Velde, E. A., Middelkoop, H. M., Kerkhof, G. A. and Kamphuisen, H. C. Ambulatory activity monitoring during sleep: an evaluation of internight and intrasubject variability in healthy persons aged 50–98 years. *Sleep*, 1993, 16: 146–150.

Van Someren, E. J. W. More than a marker: interaction between the circadian regulation of temperature and sleep, age-related changes, and treatment possibilities. *Chronobiol. Int.*, 2000, 17: 313–354.

Vitiello, M. V., Larsen, L. H. and Moe, K. E. Age-related sleep change – gender and estrogen effects on the subjective-objective sleep quality relationships of healthy, noncomplaining older men and women. *J. Psychosom. Res.*, 2004, 56: 503–510.

Woodward, S. H., Bliwise, D. L., Friedman, M. J. and Gusman, F. First night effects in PTSD inpatients. *Sleep*, 1996, 19: 312–317.

Youngstedt, S. D. Ceiling and floor effects in sleep research. *Sleep Med. Rev.*, 2003, 7: 351–365.

Youngstedt, S. D. Effects of exercise on sleep. *Clin. Sports Med.*, 2005, 24: 355–365.

Youngstedt, S. D., O'Connor, P. J. and Dishman, R. K. The effects of acute exercise on sleep: a quantitative synthesis. *Sleep*, 1997, 20: 203–214.

Youngstedt, S. D., Kripke, D. F. and Elliott, J. A. Circadian phase-delaying effects of bright light alone and combined with exercise in humans. *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, 2002, 282: R259–R266.