Actigraphy and Sleep Diary Measurements in Breast Cancer Survivors: Discrepancy in Selected Sleep Parameters

Camille M. Moore, PhD [Student],
University of Colorado Department of Biostatistics and Informatics, Colorado School of Public Health, Aurora, Colorado, U.S.A. 80045 camille.moore@ucdenver.edu

Sarah J. Schmiege [Assistant Professor], and
University of Colorado Department of Biostatistics and Informatics, Colorado School of Public Health, Aurora, Colorado, U.S.A. 80045 Phone: 303-724-8080; Fax: 303-724-8560 sarah.schmiege@ucdenver.edu

Ellyn E. Matthews[*] **[Associate Professor of Nursing]**
University of Colorado Denver, College of Nursing 13120 East 19th Ave, Aurora, Colorado, U.S.A. 80045

Keywords

actigraphy; sleep diary; insomnia; breast cancer; Bland-Altman plot

Measuring longitudinal sleep patterns presents unique challenges in research and clinical practice. Polysomnography (PSG) is the current gold standard for detecting sleep based on "time in bed" period, usually in a sleep lab. Yet PSG is impractical for long-term sleep studies and may not represent sleep in a usual environment (Sanchez-Ortuno, Edinger, Means, & Almirall, 2010; Tilmanne, Urbain, Kothare, Wouwer, & Kothare, 2009). Thus, researchers have turned to alternative, cost-effective ways of evaluating sleep, such as sleep diaries/logs and wrist actigraphy. Sleep diary data is self-reporting of sleep periods; actigraphs provide longitudinal, 24-hour assessment of sleep patterns in a natural environment.

The purpose of the present analysis was to compare actigraphy and diary measurement in breast cancer survivors with insomnia. The etiology of insomnia in these women is complex. Hormonal treatments, persistent fatigue, pain, and other nighttime symptoms lead to frequent arousals and reduced sleep duration (Dhruva et al., 2012; Savard, Ivers, Villa, Caplette-Gingras, & Morin, 2011; Van et al., 2013). The ability to detect sleep periods by actigraphy is substantially reduced in those with frequent arousals and short sleep duration

(Martin & Hakim, 2011). Moreover, accurate self-report of sleep is likely to be complicated by variability in circadian rhythms and chemotherapy-related cognitive impairment in breast cancer survivors. Reduced insight or poor recall of sleep may influence subjective reporting (Bender et al., 2013; Koppelmans, Breteler, Boogerd, Seynaeve, & Schagen, 2013; Minton & Stone, 2013). Actigraphic measurement may be particularly problematic in breast cancer survivors because of frequent nighttime awakenings, which provides the rationale for examining actigraphy and diary measurement discrepancies in this population.

Lack of concordance between measures of sleep may impose clinical dilemmas for those who treat insomnia in medical populations with restless or fragmented sleep. Trustworthy estimates of sleep improvement are required to determine the efficacy of insomnia treatments. Reliability, validity, and agreement among sleep measures have been the focus of recent reviews (Berger et al., 2008; Sadeh, 2011). Additional investigation, however, is needed to evaluate performance of sleep measurements in individuals with insomnia *and* medical conditions.

Background

Actigraphy

Actigraphy, a method of inferring sleep from the presence or absence of wrist movement, uses a piezoelectric accelerometer to estimate sleep and wake episodes (Lichstein et al., 2006; Sadeh, 2011). Wrist actigraphs provide high sensitivity (i.e., accuracy in detecting sleep) particularly in normal sleepers (de Souza et al., 2003; Kripke et al., 2010). Yet some studies have raised concerns about the specificity (i.e., accuracy in detecting wakefulness) of actigraphy in some populations or devices (de Souza et al., 2003; Paquet, Kawinska, & Carrier, 2007). Even in 21 healthy volunteers, de Souza and colleagues reported high sensitivity but relative low specificity, in which actigraphy (zero crossing mode) systematically overestimated total sleep time (TST; sleep period minus sum of minutes to fall asleep and minutes awake during sleep period), sleep latency (SL; minutes to fall asleep) and underestimated wake after sleep onset (WASO; sum of minutes awake during the sleep period) compared to PSG using the Bland and Altman technique (de Souza et al., 2003). In contrast, both actigraphy and diaries underestimated TST compared to PSG in 17 middle aged adults with primary insomnia, yet similarly overestimated WASO (Vallieres & Morin, 2003). These contradictory findings suggest sufficiently powered studies are needed in adults with insomnia compared to healthy controls using similar devices and scoring modes.

Also in healthy adults, Paquet and colleagues found that increased wakefulness during sleep diminished actigraphy to PSG concordance, suggesting the capacity of actigraphy to detect quiet wakefulness impacts its validity in clinical populations with restless sleep (Paquet et al., 2007). Moreover, actigraphy measures wrist movement rather than encephalographic sleep and may not detect small movements associated with wakefulness (Sadeh, 2011).Some discrepancy between actigraphy and PSG, therefore, is expected (Chae et al., 2009).

Recent methodological reviews suggest the need for uniform scoring algorithms, consistency of information in published studies involving actigraphy, and standard devices and modes for interpreting the data to advance sleep science (Berger et al., 2008; Sadeh,

2011). For example, there are currently three modes for interpreting actigraphy: *zero crossing mode* which counts the number of times the accelerometer waveform crosses 0 for each time period, *proportional integral mode* (PIM) measures the area under the curve, and *time above threshold* (TAT) uses a set threshold, which measures the time above that threshold. A threshold of 20 seconds is considered low, 40 is medium, and 80 is high. TAT provided the best results in a large scale study comparing TST from actigraphy and PSG in adolescents with and without sleep-disturbed breathing (Johnson et al., 2007). In contrast, Blackwell et al. found that PIM mode correlated the best with PSG in older women (Blackwell et al., 2008) and older men (Blackwell, Ancoli-Israel, Redline, & Stone, 2011). These studies underscore the difficulties in comparing studies with divergent scoring algorithms. (Ancoli-Israel et al., 2003; Berger et al., 2008). Despite noted methodological concerns, actigraphy allows objective data collection in large samples in which participant burden is a concern, and when demonstration of chronic behavior is needed (Blackwell et al., 2008). It has been suggested that validity of actigraphy exceeds common medical and psychological tests, and methodological discrepancies are not random, thus, are correctable (Tryon, 2004).

Daily Sleep Diary

The sleep diary, a daily self-report of sleep and wake patterns, provides important information about behavioral dimensions of sleep. Accessibility, conceptualization of the experience of sleep as subjective, and concordance with PSG support the use of diaries in sleep research (Sadeh, 2011). Although diary self-reporting is a relatively simple and inexpensive, self-reporting of sleep has drawbacks. Daily monitoring of sleep may focus attention on sleep habits, and data may be affected by selective reporting, social desirability bias, and misperceptions (Edinger et al., 2000).

Evidence suggests sleep misperceptions may be common in insomnia sufferers (Edinger et al., 2000; Harvey & Tang, 2012). For instance, when self-reported estimates of TST, SL, and WASO were compared to PSG, individuals with chronic insomnia showed a greater propensity than normal sleepers to underestimate TST and overestimate WASO (Means, Edinger, Glenn, & Fins, 2003). Sleep misperceptions can be distributed across a broad continuum such that they may range from small to large overestimated or underestimated sleep parameters relative to PSG. This suggests the accuracy of actigraphy and sleep diary scores relative to PSG and the nature of misperceptions may be related to the underlying sleep problem (Edinger & Fins, 1995). Relatively little is known about sleep misperceptions in medically comorbid insomnia populations, particularly over a period of weeks. Understanding the perception or even misperception of insomnia is essential to its effective management.

Comparisons of Actigraphy and Sleep Diary

Across healthy adolescent and adult populations TST measured by actigraphy is typically shorter than by diary (Lauderdale et al., 2006; Short, Gradisar, Lack, Wright, & Carskadon, 2012; Van Den Berg et al., 2008), but longer in other studies (Kawada, 2008; Vallieres & Morin, 2003). For example, Short et al. compared eight days of actigraphy and diaries in adolescents (N = 385) aged 13-18, and found that actigraphy recorded lower TST and higher

WASO than diaries (Short et al., 2012). Likewise, Landerdale and colleagues reported mean actigraphically measured TST was nearly an hour shorter than diary among early-middle aged adults (N = 699) (Landerdale et al., 2006). In community-dwelling older adults (N = 969), TST by actigraphy was shorter than diary (Van Den Berg et al., 2008).

In contrast, Kawada calculated agreement rates for waking and sleeping in actigraphy versus diary in 76 healthy adults. Mean TST by actigraphy (629.6 minutes) was higher than diary (482.3 minutes), but percentages of agreement were 86.1% (SD = 6.2) for sleep and 77.5% for wake time (SD = 10.2) (Kawada, 2008). Vallieres and Morin found actigraphy recorded more TST and less WASO compared to diaries in 17 adults with insomnia, consistent with the notion that insomnia sufferers underestimate TST. Compared to PSG, however, both actigraphy and diaries underestimated TST and overestimated WASO (Vallieres & Morin, 2003). Similarly, in community dwelling older adults (N = 39), actigraphy overestimated TST and WASO compared to diary, and actigraphy was correlated more highly with PSG. (Friedman et al., 2000).

It is worth mentioning that the three studies in which TST measured by actigraphy was shorter than by diary, one study used actigraphy with a zero crossing mode (Short et al., 2012), one used TAT with a threshold of 20 (Van Den Berg et al., 2008), and one did not specify the mode (Lauderdale et al., 2006). In those studies in which actigraphically measured TST was longer than diaries, one study used TAT with a threshold of 40 (Kawada, 2008) and the other two did not specify the mode (Friedman et al., 2000; Vallieres & Morin, 2003).

Relatively few insomnia studies have compared actigraphy and diaries in populations with medical conditions (Sadeh, 2011), and even fewer in cancer patients. Thus, it is unclear if individuals with insomnia and medical conditions underestimate sleep duration similar to those with primary insomnia. Dean and colleagues reported shorter TST, and longer SL by actigraphy compared to diary assessments in patients with lung cancer (N = 35) (Dean et al., 2013). Using actigraphy and electronic diary to evaluate sleep in patients with fibromyalgia, Okifuji and Hare found a 73 absolute minute difference in methods per night per patient. TST was longer for electronic diary compared to actigraphy about half the nights (Okifuji & Hare, 2011). Sinclair and colleagues (2014) reported poor agreement between actigraphy and diary assessments of WASO in patients with traumatic brain injury (TBI; n = 21) and healthy controls (n = 21). TBI patients showed weaker agreement between methods compared to controls (Sinclair, Ponsford, & Rajaratnam, 2014). Similarly, Wang et al., reported that sleep diary measurement of TST was 1.25 times higher, and WASO was lower than actigraphy measure in stable heart failure patients (Wang, Hung, Tsai, 2011). Limitations of these studies include brief measurement periods, small sample sizes, and limited analysis of participant level discrepancies.

Study Overview

The purpose of this study was to compare actigraphically and sleep diary measured TST and WASO in breast cancer survivors with at least subthreshold insomnia as determined by Insomnia Severity Index (ISI) scores of 8 or higher. The specific aims of these analyses were: (1) examine the overall average difference between sleep diary and actigraphy

measures by examining the average discrepancy between all pairs of actigraphy and sleep diary measurements collected over time, (2) capitalize on the longitudinal nature of the data to examine discrepancies at the level of the participant, and (3) determine if any observed differences between actigraphy and sleep diary could be accounted for by demographic, medical, or individual difference variables. Discrepancies at the level of the participant were examined by calculating the average discrepancy between the two methods by each participant and by calculating the variability in discrepancy in each participant over time.

Method

This is a secondary analysis of a randomized clinical trial whose primary endpoint was sleep improvement and quality of life; full details of the research protocol are described elsewhere (Matthews et al., 2014). Briefly, once randomized to cognitive behavioral therapy for insomnia (CBTI) or control (behavioral placebo therapy), 43 breast cancer survivors with insomnia wore actigraphy and completed daily diaries over five weeks.

Participants and Setting

Forty-three women were recruited from local outpatient oncology clinics, breast cancer support groups, and Western U.S. communities. This research was approved by the Institutional Review Board, and all women provided signed, informed consent. The CBTI and control groups were combined for these analyses and group differences were analyzed.

Women were eligible if they completed treatment for stage I-III breast cancer and met criteria for chronic insomnia (i.e., self-reported SL or WASO >30 min on ≥3 nights/ week for ≥1 month and a score of at least 8 on the Insomnia Severity Index (Morin, 1993). Other inclusion criteria were: age between 21-65, 1-36 months after completion of cancer treatment, insomnia that started or worsened at diagnosis as determined by a clinical interview, and ability to speak and write English. Exclusion criteria were unstable/untreated psychiatric or serious non-cancer medical condition, sleep disorder other than insomnia, unstable doses of medications that affect sleep, and night-shift employment. The presence of other sleep disorders was determined by a screening interview. Women reporting a diagnosed sleep disorder (e.g., sleep apnea, restless legs syndrome) were excluded from the study.

Measurements

Actigraphy—Activity/sleep periods were measured by the *Actiwatch 2* (Philips Respironics, Bend, OR) worn on the non-dominant wrist. To minimize researcher scoring bias, sleep parameters were calculated using the Respironics Actiware 5.59.0015 software, based on the previously validated algorithm (Sadeh, Aster, Urbach, & Lavie, 1989).

Actigraphy output was digitally downloaded and automatically scored using the manufacturer's *Actiware* software. Each 1-minute epoch was scored as either sleep or wake by comparing activity counts for the epoch and those surrounding it (two minutes in either direction) to a pre-set sensitivity threshold. If the number of activity counts exceeds the threshold, the epoch is scored as wake and if the number of counts is equal to or below the threshold, the epoch is scored as sleep. The TAT mode with a medium wake threshold (40

seconds) was used in this study. We used the Actiware default setting of 10 consecutive immobile minutes with at least some activity (but < one minute) to define sleep onset and 10 consecutive minutes of activity to define sleep end. Two independent scorers, trained by the manufacturer, inspected the data for periods of prolonged inactivity indicating watch removal. Periods of inactivity greater than 60 minutes (most often at the beginning and end of the measurement period) were not included in the analysis.

The Morin Sleep Diary (Morin, 1993) measures common sleep behaviors including (1) time getting into bed; (2) time of lights out and intended to fall asleep; (3) sleep onset latency; (4) number of awakenings; (5) duration of awakenings; (6) time of final awakening; (7) rise time; and (8) perceived sleep quality (via Likert scale). Diaries were completed each morning and were discussed at study visits. If diary entries were unclear or incomplete, clarification of responses was sought and an agreed upon response was entered.

On average, the women had 27.60 repeated actigraphy measurements (SD: 11.07, Range: 1-50) and 33.60 repeated diary measurements (SD: 5.16, Range: 7-38); 1187 usable pairs of actigraphy and diary measurements from 43 women were collected during the study. Although both sleep diary and actigraphy measurements were collected for five weeks, individual subjects may have had fewer matched pairs of measurements due to missing diary data or malfunction or misuse of the actigraph. Forty-one women had 7 or more matched pairs of data.

Baseline demographic/medical characteristics included demographic (age, marital status, employment) and medical information (months since diagnosis, cancer stage and sleep aid use).

Insomnia Severity Index (ISI), a 7-item questionnaire validated in cancer patients, measures the global severity and impact of insomnia (Savard, Savard, Simard, & Ivers, 2005). Items were rated from 0 (none) to 4 (very severe); scores range from 0-28; higher scores indicated greater impairment ($M = 18.24$, $SD = 3.68$). In this study reliability was high ($a = .90$).

Hospital Anxiety and Depression Scale (HADS) measures self-reported anxiety and depression (Zigmond & Snaith, 1983). This 14-item, 0-3 scale, is divided into two 7-item anxiety and depression subscales. The maximum score for each subscale is 21; higher scores reflect greater anxiety or depression ($M = 7.93$, $SD = 4.29$ for anxiety; $M = 4.79$, $SD = 3.28$ for depression). Cronbach's a was .85 for the anxiety subscale and .83 for the depression subscale.

Piper Fatigue Scale (PFS) (Piper et al., 1998), is a 22-item scale which measures four dimensions of fatigue: behavioral/severity, affective meaning, sensory, and cognitive/mood. The total fatigue score on this instrument ranges from 0 (no fatigue) to 10 (extreme fatigue) ($M = 5.71$, $SD = 1.59$), and the internal consistency reliability was high ($a = .97$).

Hot Flash Severity was assessed with the first item of the Menopause Rating Scale (Schneider, Heinemann, Rosemeier, Potthoff, & Behre, 2000). This item asks participants to rate the severity of hot flashes and sweating on a 5-point scale from 0 (none) to 4 (very severe). Scores on this item encompassed the full range of the scale, with only 4.88%

participants reporting no hot flashs and 17.07% reporting very severe symptoms (*M* = 2.19, *SD* = 1.1).

Procedures

Women completed demographic/medical information, Insomnia Severity Index (ISI), Hospital Anxiety and Depression Scale (HADS), Piper Fatigue Scale (PFS), and hot flash severity rating at baseline. The five weeks of actigraphy and diary data were collected during the consecutive intervention sessions, and were used to estimate the parameters of interest (TST and WASO) in this analysis, (Matthews et al., 2014).

Data Analysis

Data were double-entered and inspected for artifacts, missing or out of range values, and non-normality. All analyses were performed in SAS version 9.3. Three methods of discrepancy in TST and WASO were evaluated: 1) the average discrepancy between all pairs of actigraphy and diary ("Mean Discrepancy, All Measurements") (Aim 1), 2) the average discrepancy between the two methods by each participant ("Mean Discrepancy, Participant Averages") (Aim 2), and 3) the variability in discrepancy in each participant over time ("Mean SD of Discrepancy") (Aim 2). The overall discrepancy assesses average disagreement between actigraphy and diary. The average discrepancy and variability in discrepancy calculated at the participant level determines if disagreement varied by participant. It also allowed for analyses testing whether we could identify subgroups of women with better or worse agreement.

Bland-Altman plots (Bland & Altman, 1999) were used to evaluate the agreement between actigraphy and diary TST and WASO (Aims 1 and 2). These graphs plot the difference between actigraphy and diary versus the mean of the two measurements and provide assessment of the overall agreement between measures and variability in agreement. Plots were developed for the overall dataset (n=43) and to illustrate discrepancies in individual participants. For the n=41 women with 7 or more repeated measures, box and whiskers plots illustrate if and how the agreement between the measures varied across participants (Aim 2). The 2 women with fewer than 7 repeated measures were excluded from this individual level analysis because there was insufficient data to adequately characterize their within subject variability in discrepancy scores. Pearson product-moment correlation coefficients were calculated between women's mean discrepancies and demographic (age, educational level), medical (cancer stage, sleep aid use), and individual difference variables (insomnia severity, mood, fatigue, hot flash severity) to determine whether lack of agreement between the measures was related to the participant characteristics (Aim 3). An independent t-test examined whether there were differences in discrepancies between intervention groups to ensure that sleep improvements in the CBTI group were not accounting for any observed discrepancies between actigraphy and diaries.

Results

A full analysis of demographic and baseline symptoms was performed and reported elsewhere (Matthews et al., 2014). Of the 43 women in the present analyses, the typical

participant was middle aged (average age of 51.86, SD = 7.76), Caucasian (93.02%), married (53.49%), and well-educated (65.12% college educated), and just under half were employed full-time (46.51%). Cancer stage was distributed approximately equally between stages 1-3 and most had received radiation and/or chemotherapy.

Aim 1: Overall Actigraphy/Diary Discrepancy Averaged Across All Participants

Averaged over 5-weeks, actigraphy measured 55.75 (SD = 112.42) less TST minutes and 85.19 (SD = 81.36) more WASO minutes than diaries (Mean Discrepancy, All Measurements, Table 1). Bland-Altman plots revealed relatively poor agreement between actigraphy and diary TST and WASO (Figures 1-2). The y-axis for these figures represents actigraphy minus diary. Positive scores represent higher actigraphy than diary scores; negative scores represent higher diary than actigraphy. Both higher and lower TST were measured by actigraphy compared to diary data, as illustrated by points scattered above and below the 0 discrepancy line (Figure 1). Discrepancies ranged from -402.95 to 441.0 minutes. WASO measured by actigraphy tended to be higher than diary, with more points above the 0 discrepancy line, and a range of -246.0 to 407.2. The hard line along the left edge of the Bland-Altman plot in Figure 2 indicates that actigraphy measured some WASO even if women reported no WASO in the diary.

Aim 2: Discrepancy between Actigraphy and Diary: Participant Level

For TST, the average participant had a mean discrepancy (actigraphy - diary) of -62.3 minutes (Range: -267.4 to 74.2) across repeated measures with an average standard deviation (i.e., Mean SD of Discrepancy) of 80.97 (Range: 39.2 to 161.3). If we assume that discrepancies are normally distributed, only 42.3% of the average participant's actigraphy and diary TST measurements would be expected to agree within 1 hour. For WASO, the average participant had a mean discrepancy of 86.8 minutes (Range: 10.8 to 240.6) with a Mean SD of Discrepancy score of 60.7 (Range: 19.4 to 138.7) (Table 1). Again, assuming a normal distribution, only 32.2% of the average participant's actigraphy and diary WASO measurements would be expected to agree within one hour.

Box and whiskers plots (Figures 3-4) are presented for the participants with 7 or more repeated measures where a separate box and whiskers is shown for each participant to further demonstrate the variability in agreement between measures and to demonstrate individual level patterns in agreement/disagreement. For example, a narrow box plot, centered around 0, indicates a participant had relatively good agreement between actigraphy and sleep diary measures. A narrow boxplot centered above (or below) 0 indicates that a participant's actigraphy measurements were consistently higher (or lower) than sleep diary measurements. A wide boxplot indicates high variability in agreement within a participant. As the figures show, there was a range of agreement patterns where some individuals showed relatively good agreement, despite the average disagreement across participants.

Participant-level Bland-Altman plots were also created to display the patterns of measurement agreement across time within participants. Examples from three participants are shown in Figures 5 and 6 for illustration purposes. While some women showed good agreement in TST and WASO with similar changes over time (e.g., Participant 3), others

were more variable. Participant 1 illustrates the case where TST and WASO were both over-and underestimated by actigraphy relative to diary, while Participant 2 illustrates TST measured by actigraphy was consistently lower than TST measured by diary, and WASO was substantially greater than zero by actigraphy, but consistently close to zero by diary.

Aim 3: Participant Characteristics and Actigraphy/Diary Discrepancy

There were no significant correlations between TST or WASO discrepancy and participants' age, cancer stage, educational level, sleep medication use, insomnia severity/impact, anxiety or depression, fatigue, or hot flash severity (Table 2). This indicates that discrepancies between actigraphy and sleep diaries cannot be explained by any of these measured variables.

There were also no statistically significant differences in the mean TST and WASO discrepancies (actigraphy - diary) between women in the CBTI and control groups. For TST, those in the CBTI group had a mean discrepancy of −52.71 (SD = 65.16), while those in the control group had a mean discrepancy of −73.40 minutes (SD = 94.82) (p=0.43, Cohen's d=0.26). For WASO, those in the CBTI group had a mean discrepancy of 74.78 minutes (SD = 40.12), while those in the control group had a mean discrepancy of 100.7 minutes (SD = 68.42) (p=0.16, Cohen's d=0.47). It is plausible that sleep diary entries of women in the CBTI group might reflect shorter TST in keeping with sleep restriction instructions, particularly in the beginning of CBTI treatment. The lack of significant group differences and the small effect sizes associated with the tests of possible differences suggest that women receiving CBTI did not complete the diaries based on sleep restriction expectations, thus, treatment and differential sleep improvement did not significantly impact the TST and WASO measurement discrepancies.

Discussion

To our knowledge, research related to discrepancies between actigraphy and diary reports in a breast cancer population has not been previously reported. The current analysis examined agreement between measurements of TST and WASO using three methods (mean discrepancy, all measurements; mean discrepancy, participant averages; mean SD of discrepancy). Across all three methods, we observed relatively poor agreement between actigraphy and diary measurements of TST and WASO. Accurate metrics of sleep improvement are needed to determine treatment efficacy, particularly in understudied medical populations with insomnia.

Our findings, in which actigraphy measured less TST and more WASO than sleep diaries at an overall level and at the participant level are consistent with other studies of insomnia in patients with a variety of medical conditions including cancer. Lack of agreement between actigraphy and diary TST and WASO has been reported in adults with lung cancer (Dean et al., 2013), fibromyalgia (Okifuji & Hare, 2011), and heart failure (Wang, Hung, & Tsai, 2011).

Our results in women with insomnia comorbid with breast cancer suggest that misperceptions of sleep recorded in a diary are fairly common, the accuracy of TST

perception varies widely, and breast cancer survivors often have fragmented sleep. Consistent with variability of TST and WASO measurement across some women in the present study, patients with fibromyalgia syndrome (FMS) (N = 75; 97% female) had a near equal distribution of underestimation and overestimation of TST by actigraphy and sleep diary assessment (Okifuji & Hare, 2011). The authors concluded that FMS patients were most likely to misperceive TST when sleep was restless. Apart from misperceptions, actigraphy overestimation of wake during restless sleep, and underestimation during motionless awakenings may be related to a high sensitivity of the actigraphy to wrist movement, or may simply indicate agitated or fragmented sleep (Cole, Kripke, Gruen, Mullaney, & Gillin, 1992; Friedman et al., 2000; Hauri & Wisbey, 1992; Okifuji & Hare, 2011).

In the present study, variability in agreement could not be explained by participant characteristics such as age, cancer stage, education, insomnia, mood, fatigue, or hot flash severity. Although null findings provide important information, it remains unknown which variables do predict discrepancies or may be associated with subgroups or phenotypes of discrepancies. Contrary to our findings, it has been suggested that mood can affect a person's ability to estimate TST, and depression and anxiety may be more predictive of subjective sleep complaints then objective disturbances (Edinger et al., 2000). Another potential factor in accurate estimation of sleep is an "anchoring effect." When a person tries to estimate sleep in the previous night, he/she may anchor the answer to how he/she felt at the time (Okifuji & Hare, 2011). If a person is tired, sleepy or non-restored when completing a sleep diary, he/she may deduce that sleep was poor and this may be reflected in TST and WASO estimates.

Although nocturnal hot flash severity would be expected to reflect poor sleep diary ratings (e.g., less TST and more WASO minutes) (Carpenter, Johnson, Wagner, & Andrykowski, 2002; Savard, Savard, Trudel-Fitzgerald, Ivers, & Quesnel, 2011), hot flash severity did not explain measurement discrepancy in the present study. This may be due to the single item measure and the importance of assessing other characteristics of hot flashes. Savard at al., (2013) found that hot flashes frequency was not associated with more sleep disturbances among breast cancer survivors. Instead, the time to reach peak conductance was most consistently associated with sleep impairments (Savard, Savard, Caplette-Gingras, Ivers, & Bastien, 2013). Slow-developing hot flashes may lead to more enduring symptoms such as sweating and increase the propensity for sleep awakenings (Savard et al., 2013). It is unclear if slow peak hot flashes are perceived by women as more severe or if slow peak hot flashes affect the concordance between subjective and objective assessments of sleep.

Depending on factors that influence subjective reporting and the nature of the sleep disturbance, it is possible that breast cancer survivors fall into three phenotypes of actigraphy–diary agreement (Figures 5-6). The first phenotype consists of women exhibiting unpredictable agreement with actigraphy (*both* over or underestimation of sleep and wake in any given night). A second phenotype includes women with poor agreement in a consistent direction such as overestimation of TST and underestimation of WASO by diary. The last phenotype may include women with diary reports that are consistent with actigraphy.

Many clinicians and researchers evaluate sleep using actigraphy, diaries or both; actigraphy monitoring can be expected to increase in future research (Berger et al., 2008). Thus, concordance between subjective and objective sleep measures has important clinical and research implications. The perception or even misperception of symptom frequency, severity, and impact is essential to identification and management, and sleep disturbance is no exception.

An interesting line of research with implications for treatment is the importance of sleep perception in medical populations with concurrent insomnia. Evidence suggests that underestimation of TST does not appear to be a universal characteristic of all insomnia sufferers (Edinger & Fins, 1995). Many questions about sleep time perception remain to be answered, including the influence of sleep environment, personality, and constitutional factors on sleep time perceptions (Means et al., 2003). Further research efforts to shed light on the characteristics that influence sleep diary accuracy may include measures of social desirable responding, personality, and perceived accuracy of diary measures. Sleep inertia and other neurophysiological attributes, which may impair ability to recognize wakefulness and approximate time awake may be helpful for improved prediction of discrepancies between measures (Harvey & Tang, 2012).

Sleep experts recommend using both sleep diaries and actigraphy (Carney, Lajos, & Waters, 2004; Kushida et al., 2001; Vallieres & Morin, 2003) and conducting sensitivity analyses to determine if similar treatment effects can be measured with actigraphy and diaries over the course of longitudinal sleep studies. If similar treatment effects are observed for actigraphy and diary despite disagreement between the methods, this can increase confidence in the likelihood that a true treatment effect exists. Conversely, treatment effects observed for one method but not the other should be interpreted with caution. Accuracy may be improved by helping study participants remember to fill out the diaries electronically and by providing phone and other electronic reminders. It may be worthwhile to delineate the most appropriate measure(s) based on the sleep phenomenon of interest, apart from the merits and drawbacks of individual devices and questionnaires unlike established scoring guidelines for PSG, actigraphs lack standard algorithms and software for calculating sleep parameters. If established, actigraphy standards for epoch length, data cleaning, and scoring could lead to the development of uniform empirical evidence across insomnia populations.(Edinger, Means, Stechuchak, & Olsen, 2004).

Most studies have compared actigraphy and diary across seven or fewer nights. Longitudinal actigraphy and diary data strengthened this study and allowed us to examine trends in sleep parameters over longer periods of time and determine if the two measurements show agreement within and across participants. Limitations of the present study, however, must be considered. First, PSG was not used. It was not possible to determine, therefore, if the lack of agreement between actigraphy and sleep diary measures of TST and WASO was due to inaccuracies in sleep diaries, in actigraphy, or both. Second, the lack of association between individual factors and measurement agreement may be related to the fact that variable characteristics such as mood and fatigue were measured at baseline, but sleep diary and actigraphy were measured subsequently in weeks 2-6. Associations may have been found if all variables were measured concurrently. These participant level characteristics (e.g., mood,

fatigue) were also assessed post-treatment (week 7); because neither baseline nor post-treatment measures corresponded perfectly with the sleep diary and actigraphy measures, we chose to use the baseline assessments because the posttest assessments may have been affected by treatment condition. Finally, the sample demographic suggests homogeneity. Women were predominately middle aged, Caucasian and well educated. This sample represents women with breast cancer in our region and those who volunteer for clinical studies, yet breast cancer affects people with a range of racial/ethnic backgrounds and educational attainment, which requires a diverse sample.

In summary, we suggest that investigators conducting longitudinal sleep studies use both actigraphy and sleep diaries, take steps to minimize discrepancies (e.g., reminders, clear instructions for diary/device use), and provide sufficient details about actigraphy (e.g., algorithm, sensitivity level, placement of device) and diary use to allow comparisons between studies.

Acknowledgements

Reference List

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(3):342–392. [PubMed: 12749557]

Bender CM, Sereika SM, Ryan CM, Brufsky AM, Puhalla S, Berga SL. Does lifetime exposure to hormones predict pretreatment cognitive function in women before adjuvant therapy for breast cancer? Menopause. 2013; 20(9):922–929. [PubMed: 23481123]

Berger AM, Wielgus KK, Young-McCaughan S, Fischer P, Farr L, Lee KA. Methodological challenges when using actigraphy in research. J Pain Symptom.Manage. 2008; 36(2):191–199. [PubMed: 18400460]

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. [PubMed: 21897772]

Blackwell T, Redline S, Ancoli-Israel S, Schneider JL, Surovec S, Johnson NL, et al. Comparison of sleep parameters from actigraphy and polysomnography in older women: the SOF study. Sleep. 2008; 31(2):283–291. [PubMed: 18274276]

Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat.Methods Med Res. 1999; 8(2):135–160. [PubMed: 10501650]

Carney CE, Lajos LE, Waters WF. Wrist actigraph versus self-report in normal sleepers: sleep schedule adherence and self-report validity. Behav.Sleep Med. 2004; 2(3):134–143. [PubMed: 15600229]

Carpenter JS, Johnson D, Wagner L, Andrykowski M. Hot flashes and related outcomes in breast cancer survivors and matched comparison women. Oncol Nurs.Forum. 2002; 29(3):E16–E25. [PubMed: 11979290]

Chae KY, Kripke DF, Poceta JS, Shadan F, Jamil SM, Cronin JW, et al. Evaluation of immobility time for sleep latency in actigraphy. Sleep Med. 2009; 10(6):621–625. [PubMed: 19103508]

Cole RJ, Kripke DF, Gruen W, Mullaney DJ, Gillin JC. Automatic sleep/wake identification from wrist activity. Sleep. 1992; 15(5):461–469. [PubMed: 1455130]

de Souza L, Benedito-Silva AA, Pires ML, Poyares D, Tufik S, Calil HM. Further validation of actigraphy for sleep studies. Sleep. 2003; 26(1):81–85. [PubMed: 12627737]

Dean GE, Redeker NS, Wang YJ, Rogers AE, Dickerson SS, Steinbrenner LM, et al. Sleep, mood, and quality of life in patients receiving treatment for lung cancer. Oncol Nurs.Forum. 2013; 40(5):441–451. [PubMed: 23989018]

Dhruva A, Paul SM, Cooper BA, Lee K, West C, Aouizerat BE, et al. A longitudinal study of measures of objective and subjective sleep disturbance in patients with breast cancer before, during, and after radiation therapy. J Pain Symptom.Manage. 2012; 44(2):215–228. [PubMed: 22795049]

Edinger JD, Fins AI. The distribution and clinical significance of sleep time misperceptions among insomniacs. Sleep. 1995; 18(4):232–239. [PubMed: 7618020]

Edinger JD, Fins AI, Glenn DM, Sullivan RJ Jr. Bastian LA, Marsh GR, et al. Insomnia and the eye of the beholder: are there clinical markers of objective sleep disturbances among adults with and without insomnia complaints? J Consult Clin Psychol. 2000; 68(4):586–593. [PubMed: 10965634]

Edinger JD, Means MK, Stechuchak KM, Olsen MK. A pilot study of inexpensive sleep-assessment devices. Behav.Sleep Med. 2004; 2(1):41–49. [PubMed: 15600223]

Friedman L, Benson K, Noda A, Zarcone V, Wicks DA, O'Connell K, et al. An actigraphic comparison of sleep restriction and sleep hygiene treatments for insomnia in older adults. J Geriatr.Psychiatry Neurol. 2000; 13(1):17–27. [PubMed: 10753003]

Harvey AG, Tang NK. (Mis)perception of sleep in insomnia: a puzzle and a resolution. Psychol.Bull. 2012; 138(1):77–101. [PubMed: 21967449]

Hauri PJ, Wisbey J. Wrist actigraphy in insomnia. Sleep. 1992; 15(4):293–301. [PubMed: 1519002]

Johnson NL, Kirchner HL, Rosen CL, Storfer-Isser A, Cartar LN, ncoli-Israel S, et al. Sleep estimation using wrist actigraphy in adolescents with and without sleep disordered breathing: a comparison of three data modes. Sleep. 2007; 30(7):899–905. [PubMed: 17682661]

Kawada T. Agreement rates for sleep/wake judgments obtained via accelerometer and sleep diary: a comparison. Behav.Res Methods. 2008; 40(4):1026–1029. [PubMed: 19001393]

Koppelmans V, Breteler MM, Boogerd W, Seynaeve C, Schagen SB. Late effects of adjuvant chemotherapy for adult onset non-CNS cancer; cognitive impairment, brain structure and risk of dementia. Crit Rev.Oncol Hematol. 2013; 88(1):87–101. [PubMed: 23768778]

Kripke DF, Hahn EK, Grizas AP, Wadiak KH, Loving RT, Poceta JS, et al. Wrist actigraphic scoring for sleep laboratory patients: algorithm development. J Sleep Res. 2010; 19(4):612–619. [PubMed: 20408923]

Kushida CA, Chang A, Gadkary C, Guilleminault C, Carrillo O, Dement WC. Comparison of actigraphic, polysomnographic, and subjective assessment of sleep parameters in sleep-disordered patients. Sleep Med. 2001; 2(5):389–396. [PubMed: 14592388]

Lauderdale DS, Knutson KL, Yan LL, Rathouz PJ, Hulley SB, Sidney S, et al. Objectively measured sleep characteristics among early-middle-aged adults: the CARDIA study. Am J Epidemiol. 2006; 164(1):5–16. [PubMed: 16740591]

Lichstein KL, Stone KC, Donaldson J, Nau SD, Soeffing JP, Murray D, et al. Actigraphy validation with insomnia. Sleep. 2006; 29(2):232–239. [PubMed: 16494091]

Martin JL, Hakim AD. Wrist actigraphy. Chest. 2011; 139(6):1514–1527. [PubMed: 21652563]

Matthews EE, Berger AM, Schmiege SJ, Cook PF, McCarthy MS, Moore CM, et al. Cognitive Behavioral Therapy for Insomnia Outcomes in Women After Primary Breast Cancer Treatment: A Randomized Controlled Trial. Oncol Nurs.Forum. 2014

Means MK, Edinger JD, Glenn DM, Fins AI. Accuracy of sleep perceptions among insomnia sufferers and normal sleepers. Sleep Med. 2003; 4(4):285–296. [PubMed: 14592301]

Minton O, Stone PC. A comparison of cognitive function, sleep and activity levels in disease-free breast cancer patients with or without cancer-related fatigue syndrome. BMJ Support.Palliat.Care. 2013; 2(3):231–238.

Morin, CM. Insomnia: Psychological Assessment and Management. Guilford Press; New York: 1993.

Okifuji A, Hare BD. Nightly analyses of subjective and objective (actigraphy) measures of sleep in fibromyalgia syndrome: what accounts for the discrepancy? Clin J Pain. 2011; 27(4):289–296. [PubMed: 21178589]

Paquet J, Kawinska A, Carrier J. Wake detection capacity of actigraphy during sleep. Sleep. 2007; 30(10):1362–1369. [PubMed: 17969470]

Piper BF, Dibble SL, Dodd MJ, Weiss MC, Slaughter RE, Paul SM. The revised Piper Fatigue Scale: psychometric evaluation in women with breast cancer. Oncol Nurs.Forum. 1998; 25(4):677–684. [PubMed: 9599351]

Sadeh A. The role and validity of actigraphy in sleep medicine: an update. Sleep Med Rev. 2011; 15(4):259–267. [PubMed: 21237680]

Sadeh A, Aster J, Urbach D, Lavie P. Actigraphically based automatic bedtime sleep-wake scoring: validity and clinical application. J Ambulat Monitoring. 1989; 2:209–216.

Sanchez-Ortuno MM, Edinger JD, Means MK, Almirall D. Home is where sleep is: an ecological approach to test the validity of actigraphy for the assessment of insomnia. J Clin Sleep Med. 2010; 6(1):21–29. [PubMed: 20191934]

Savard J, Ivers H, Villa J, Caplette-Gingras A, Morin CM. Natural course of insomnia comorbid with cancer: an 18-month longitudinal study. J Clin Oncol. 2011; 29(26):3580–3586. [PubMed: 21825267]

Savard MH, Savard J, Caplette-Gingras A, Ivers H, Bastien C. Relationship between objectively recorded hot flashes and sleep disturbances among breast cancer patients: investigating hot flash characteristics other than frequency. Menopause. 2013; 20(10):997–1005. [PubMed: 23632657]

Savard MH, Savard J, Simard S, Ivers H. Empirical validation of the Insomnia Severity Index in cancer patients. Psychooncology. 2005; 14(6):429–441. [PubMed: 15376284]

Savard MH, Savard J, Trudel-Fitzgerald C, Ivers H, Quesnel C. Changes in self-reported hot flashes and their association with concurrent changes in insomnia symptoms among women with breast cancer. Menopause. 2011; 18(9):985–993. [PubMed: 21552166]

Schneider HP, Heinemann LA, Rosemeier HP, Potthoff P, Behre HM. The Menopause Rating Scale (MRS): reliability of scores of menopausal complaints 3. Climacteric. 2000; 3(1):59–64. Retrieved from PM:11910611. [PubMed: 11910611]

Short MA, Gradisar M, Lack LC, Wright H, Carskadon MA. The discrepancy between actigraphic and sleep diary measures of sleep in adolescents. Sleep Med. 2012; 13(4):378–384. [PubMed: 22437142]

Sinclair KL, Ponsford J, Rajaratnam SM. Actigraphic Assessment of Sleep Disturbances following Traumatic Brain Injury. Behav.Sleep Med. 2014; 12(1):13–27. [PubMed: 23394102]

Tilmanne J, Urbain J, Kothare MV, Wouwer AV, Kothare SV. Algorithms for sleep-wake identification using actigraphy: a comparative study and new results. J Sleep Res. 2009; 18(1):85–98. [PubMed: 19250177]

Tryon WW. Issues of validity in actigraphic sleep assessment. Sleep. 2004; 27(1):158–165. [PubMed: 14998254]

Vallieres A, Morin CM. Actigraphy in the assessment of insomnia. Sleep. 2003; 26(7):902–906. [PubMed: 14655927]

Van Den Berg JF, Van Rooij FJ, Vos H, Tulen JH, Hofman A, Miedema HM, et al. Disagreement between subjective and actigraphic measures of sleep duration in a population-based study of elderly persons. J Sleep Res. 2008; 17(3):295–302. [PubMed: 18321246]

Van OC, Paul SM, Lee K, Dunn L, Aouizerat BE, West C, et al. Trajectories of sleep disturbance and daytime sleepiness in women before and after surgery for breast cancer. J Pain Symptom.Manage. 2013; 45(2):244–260. [PubMed: 22921179]

Wang MY, Hung HL, Tsai PS. The sleep log and actigraphy: congruency of measurement results for heart failure patients. J Nurs.Res. 2011; 19(3):173–180. [PubMed: 21857324]

Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr.Scand. 1983; 67(6):361–370. [PubMed: 6880820]
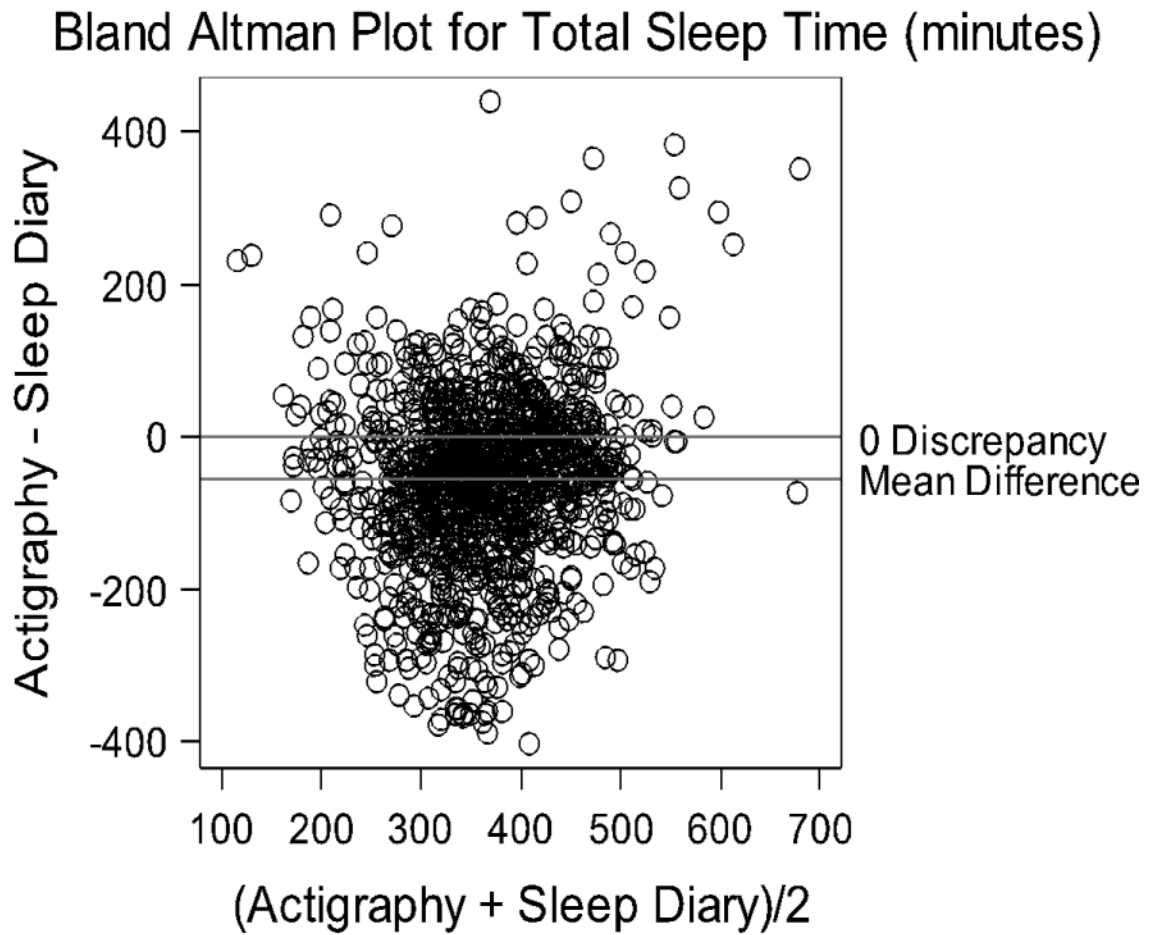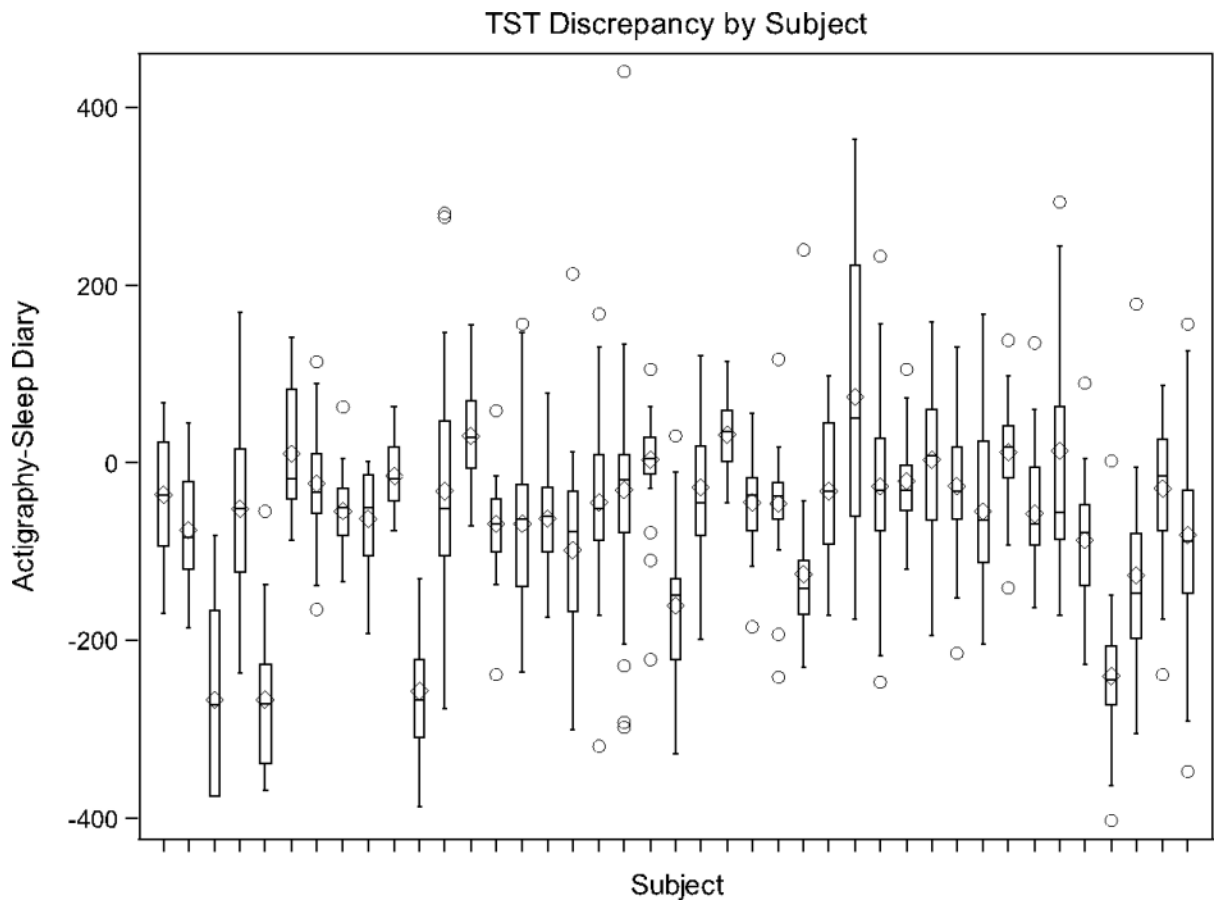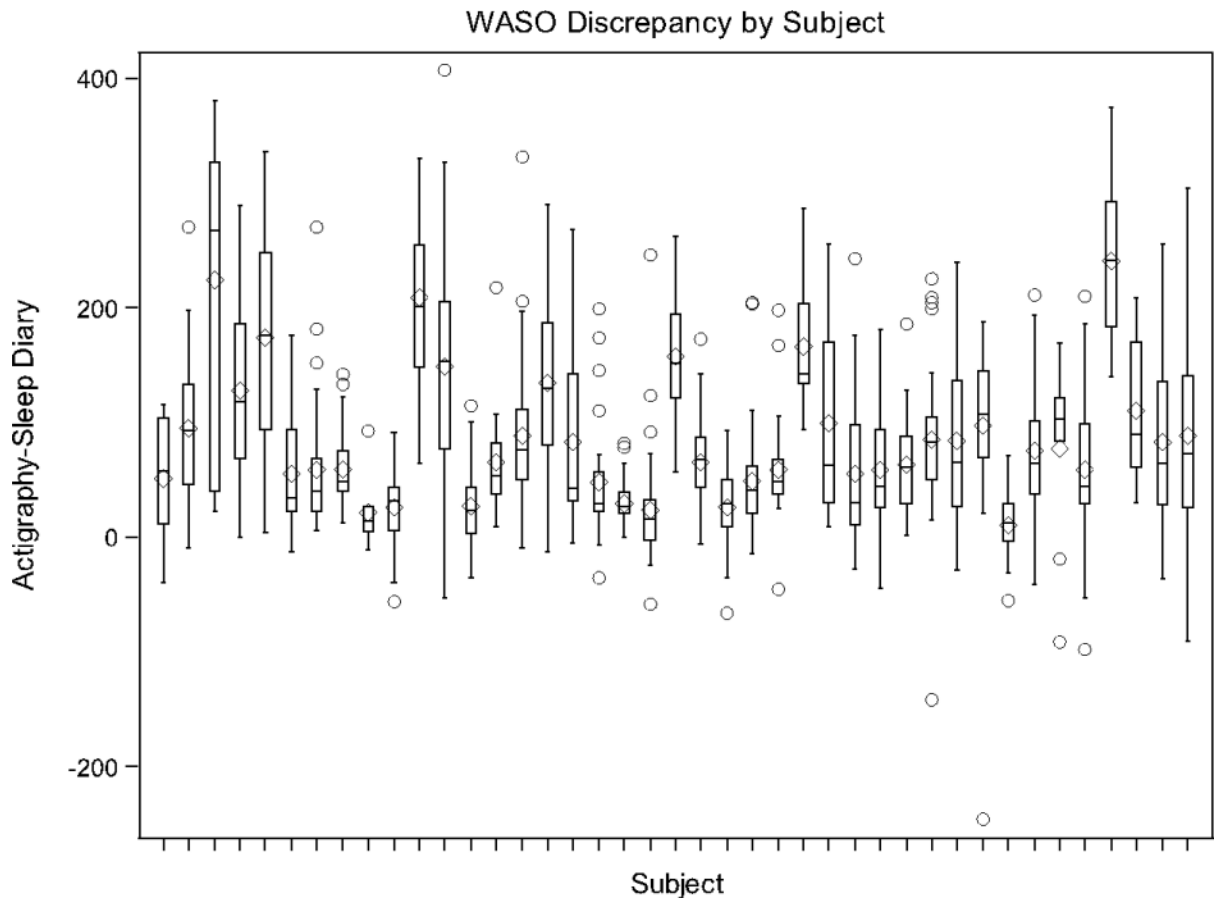
Figure 1.
Bland Altman Plot of overall TST agreement between actigraphy and sleep diary.

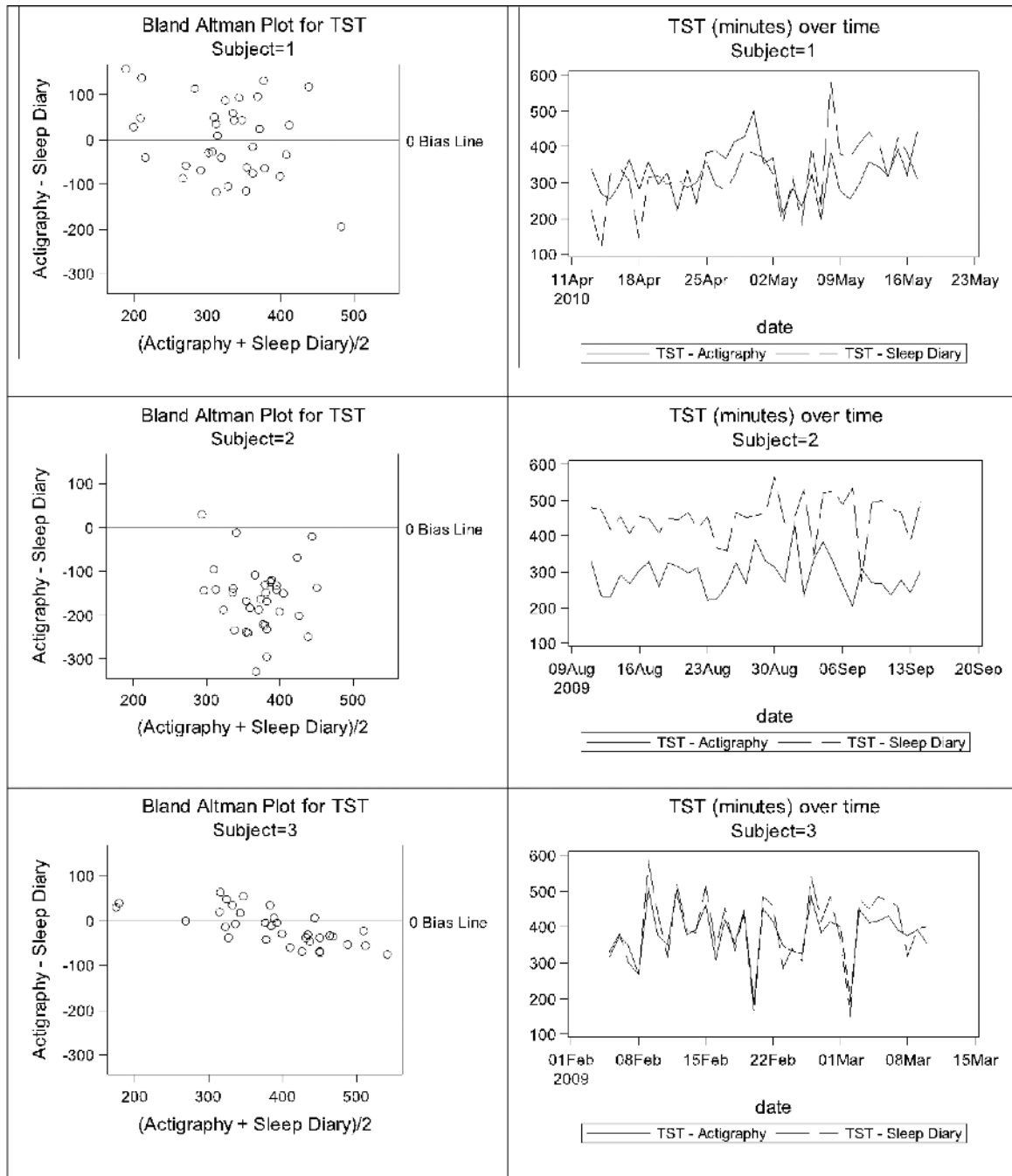# Bland Altman Plot for Wake After Sleep Onset (minutes)
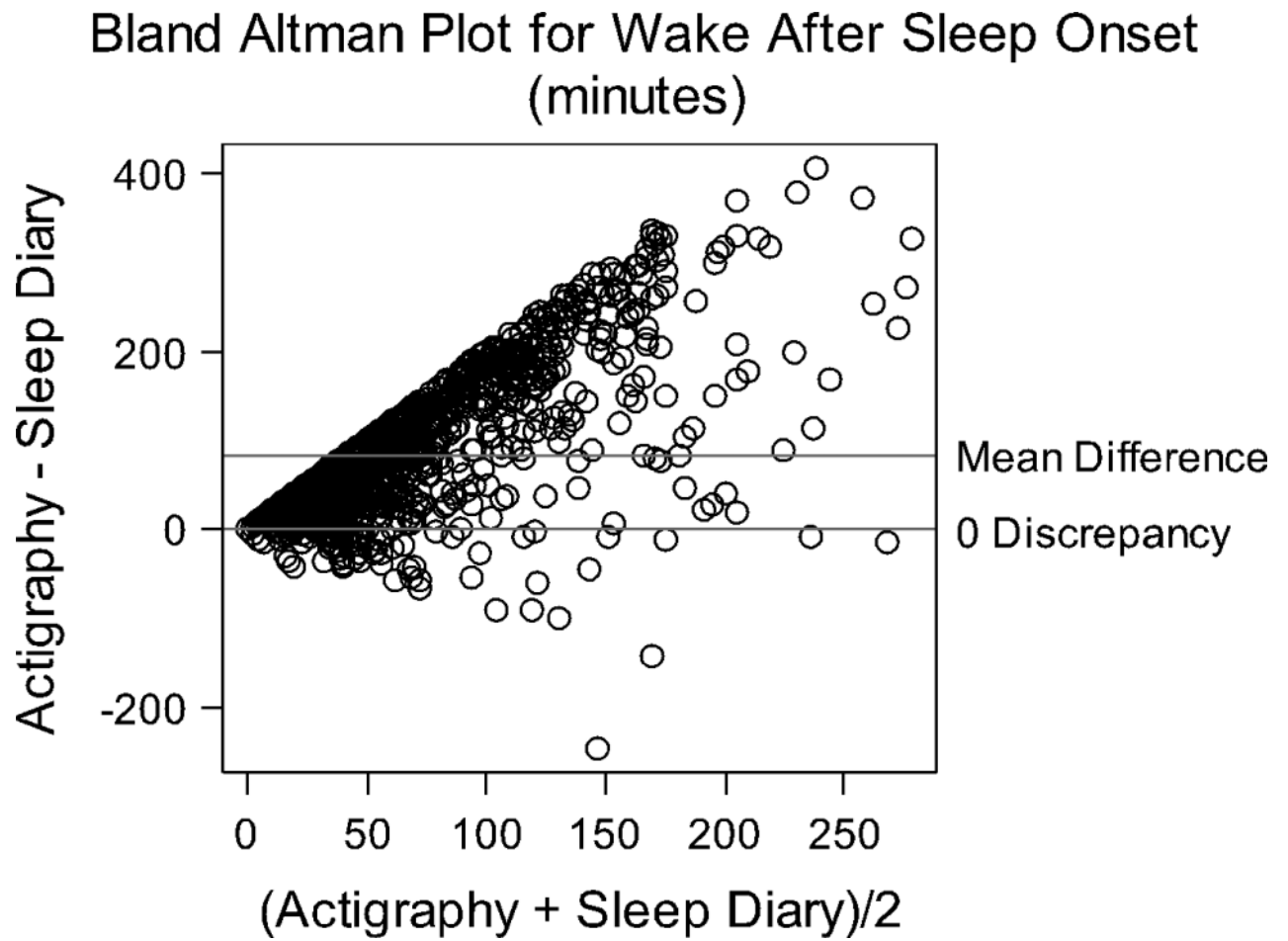
Figure 2.
Bland Altman Plot of overall WASO agreement between actigraphy and sleep diary.
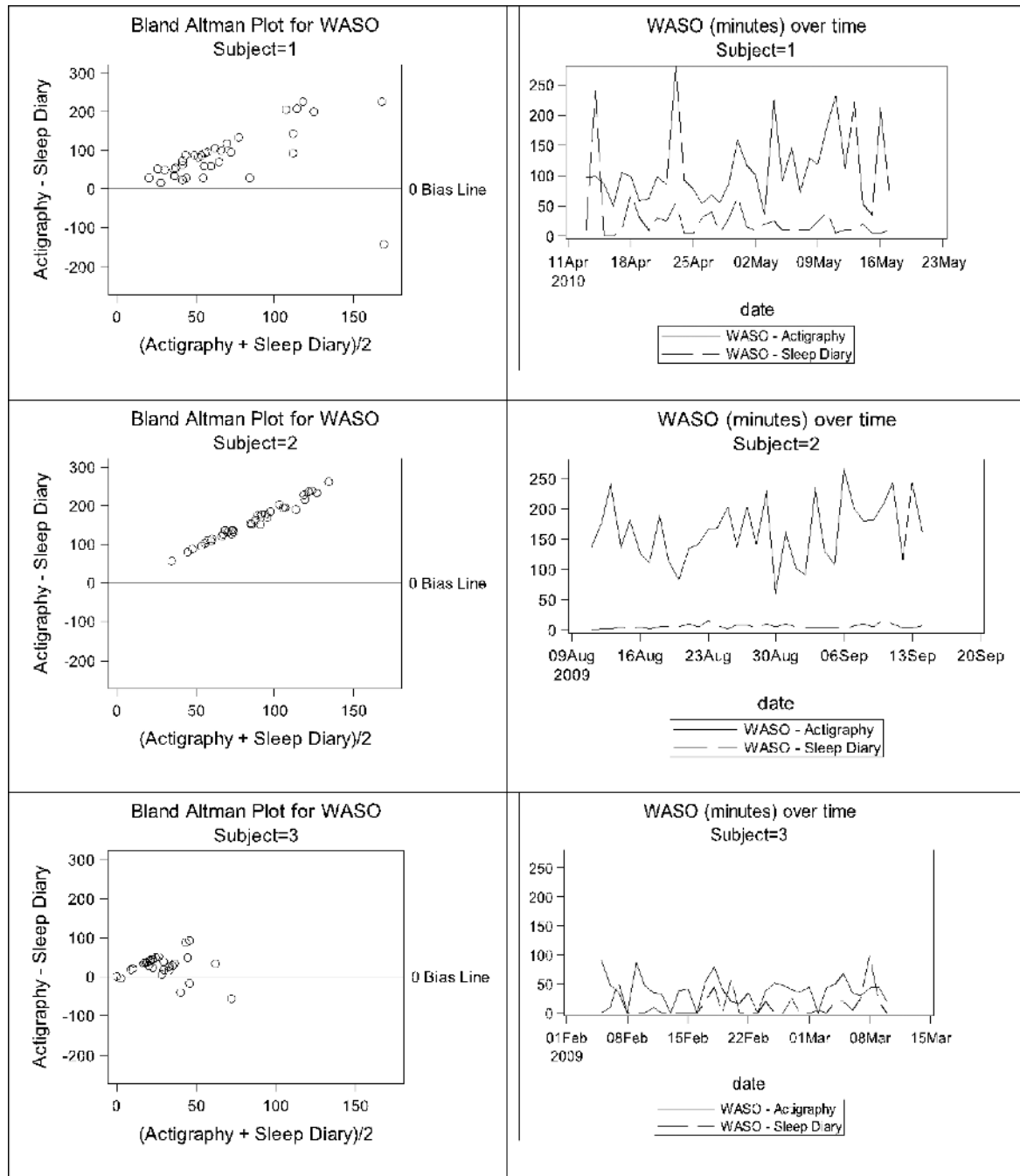
Figure 3.
Box-whisker plot of TST discrepancies by participant.

Figure 4.
Box-whisker plot of WASO discrepancies by participant.

Figure 5.
Illustration of TST discrepancies for three example participants.

Figure 6.
Illustration of WASO discrepancies for three example participants.

Table 1

Mean Discrepancies (Actigraphy minus Sleep Diary) in TST and WASO

Variable	Mean Discrepancy (SD), all measurements (n=1187)	Min	Max	Mean Discrepancy (SD), participant averages (n=41)	Min	Max	Mean SD of Discrepancy (participant standard deviations, n=41)	Min	Max
TST	−55.75 (112.42)	−402.95	441.00	−62.30 (79.91)	−267.37	74.24	80.97 (28.89)	39.22	161.27
WASO	85.19 (81.36)	−246.00	407.23	86.79 (55.88)	10.79	240.64	60.73 (22.68)	19.44	138.67

Table 2

Correlations between Mean Participant Discrepancies and Demographic/Individual Difference Characteristics

	TST Discrepancy	WASO Discrepancy
Variable	Pearson Correlation (P-Value)	Pearson Correlation (P-Value)
Age	0.12 (0.44)	−0.25 (0.12)
Cancer Stage	−0.06 (0.71)	−0.05 (0.75)
Educational Level	−0.03 (0.86)	0.03 (0.85)
Sleep Medication Use	−0.05 (0.73)	0.09 (0.58)
Insomnia Severity Index	0.08 (0.64)	0.10 (0.54)
HADS Depression Score	−0.09 (0.56)	0.17 (0.30)
HADS Anxiety Score	0.17 (0.30)	−0.15 (0.35)
Piper Fatigue Score	−0.17 (0.30)	0.28 (0.07)
Hot Flash Severity	0.04 (0.79)	0.15 (0.34)