

Original Article

Evaluating sleep in bipolar disorder: comparison between actigraphy, polysomnography, and sleep diary

Kaplan KA, Talbot LS, Gruber J, Harvey AG. Evaluating sleep in bipolar disorder: comparison between actigraphy, polysomnography, and sleep diary. *Bipolar Disord* 2012; 14: 870–879. © 2012 John Wiley & Sons A/S. Published by Blackwell Publishing Ltd.

Objectives: Bipolar disorder is an illness characterized by sleep and circadian disturbance, and monitoring sleep in this population may signal an impending mood change. Actigraphy is an important clinical and research tool for examining sleep, but has not yet been systematically compared to polysomnography or sleep diary in bipolar disorder. The present study compares actigraphy, polysomnography, and sleep diary estimates of five standard sleep parameters in individuals with bipolar disorder and matched controls across two nights of assessment.

Methods: Twenty-seven individuals who met diagnostic criteria for bipolar disorder type I or II and were currently between mood episodes, along with 27 matched controls with no history of psychopathology or sleep disturbance, underwent two nights of research laboratory monitoring. Sleep was estimated via polysomnography, actigraphy, and sleep diary.

Results: Over the 108 nights available for comparison, sleep parameter estimates from actigraphy and polysomnography were highly correlated and did not differ between the two groups or across the two nights for sleep onset latency, wake after sleep onset, number of awakenings, total sleep time, or sleep efficiency percentage. The medium wake threshold algorithm in the actigraphy software was the most concordant with polysomnography and diaries across the five sleep parameters. Concordance between actigraphy, polysomnography, and sleep diary was largely independent of insomnia presence and medication use.

Conclusions: Actigraphy is a valid tool for estimating sleep length and fragmentation in bipolar disorder.

Katherine A Kaplan^a, Lisa S Talbot^a, June Gruber^b and Allison G Harvey^a

^aDepartment of Psychology, University of California, Berkeley, Berkeley, CA, ^bDepartment of Psychology, Yale University, New Haven, CT, USA

doi: 10.1111/bdi.12021

Key words: actigraphy – bipolar disorder – polysomnography – sleep – sleep diary

Received 13 May 2011, revised and accepted for publication 5 September 2012

Corresponding author:
Allison G. Harvey, Ph.D.
Sleep and Psychological Disorders Laboratory
Psychology Department
University of California, Berkeley
3210 Tolman Hall #1650
Berkeley, CA 94720-1650
USA
Fax: + 1-510-642-5293
E-mail: aharvey@berkeley.edu

Bipolar disorder is a severe, recurrent mental illness affecting 1–4% of the population (1). Characterized by episodes of mania, hypomania, and often depression, bipolar disorder frequently includes both sleep disturbance (2, 3) and circadian disruption (4–6). Reduced need for sleep and insomnia/hypersomnia are common symptoms of the manic and depressive phases of bipolar disorder, respectively. Furthermore, sleep disturbance persists in the period between episodes (the *inter-episode* period) and may be a mechanism contributing to illness relapse. For example, one

study found that 70% of an inter-episode bipolar disorder group had a clinically significant sleep problem, and 55% met diagnostic criteria for insomnia while inter-episode (7). A second study revealed that sleep disturbance is the most common prodrome of mania and the sixth most common prodrome of bipolar depression (8). Furthermore, a prospective study utilizing daily monitoring reported that changes in self-reported sleep duration (change in total sleep time > three hours) was strongly suggestive of an impending mood episode (9). Though it is unclear as yet whether sleep

disturbance is a symptom of current episodes, prodromal to emerging episodes, or a mechanism of future illness relapse, taken together these results underscore the importance of monitoring sleep in bipolar disorder.

Actigraphy has proved to be a useful tool for estimating sleep and daily activity patterns (10), but has not been systematically evaluated in bipolar disorder. Actigraphy is a small wrist-watch-like device containing an accelerometer that records and stores motion, from which periods of sleep and wakefulness may be inferred. Though polysomnography (PSG) is a widely accepted objective measure of sleep, actigraphy has several advantages over PSG in estimating sleep: it is relatively inexpensive, minimally disruptive, and can be worn continuously in ecological environments for periods of one month or more. When compared to daily sleep diaries, actigraphy is again preferable insofar as it is less burdensome to participants and not subject to the recall biases of self-report estimates (11).

Validation work with healthy, non-disordered adults has established an encouraging 90% concordance between actigraphy and PSG (12, 13). However, there are at least two reasons to suspect that actigraphy estimates may be less accurate in bipolar disorder. First, actigraphy appears to be less accurate in populations with fragmented sleep (14) and in periods of quiet wakefulness, such as the sleep onset period (15). Numerous studies have documented that actigraphy has a tendency to overestimate total sleep time and underestimate wake during sleep in insomnia (15–18). As over half of individuals with bipolar disorder meet diagnostic criteria for insomnia, and many more demonstrate other types of sleep disturbance or subclinical insomnia, it is unclear to what extent actigraphy may similarly exhibit underestimation of wakefulness in a bipolar population. Second, numerous medications for bipolar disorder have sedative properties which may lead actigraphy to overestimate sleep (based on reduced mobility) in this population. Hypnotic use has been shown to yield overestimates of actigraphically determined sleep in patients with chronic insomnia (15), adding support to this potential limitation.

A handful of studies have successfully used actigraphy to document sleep and circadian disturbance in individuals with bipolar disorder (3, 7, 19, 20) and in those at elevated risk for bipolar episodes (21). However, without systematic comparison to PSG or sleep diaries, the interpretability of these actigraphy findings is unclear. Moreover, as sleep diaries are often a standard assessment tool by which sleep disturbance is evaluated (22), it is

important to compare actigraphy estimates against both PSG (a similarly *objective* measure of sleep) and sleep diaries (a *subjective* measure of sleep) to determine its utility in bipolar sleep monitoring.

The aims of the present research were twofold. First, concordance rates between PSG, actigraphy, and sleep diary were compared between diagnostic groups. Based on prior insomnia validation work, Aim 1 tested the hypothesis that concordance between actigraphy, PSG, and diaries would be lower in individuals with bipolar disorder than in a matched control group free of lifetime psychiatric illness. Aim 2, included on an exploratory basis, was to systematically evaluate the effects of insomnia presence and psychotropic medications on actigraphy concordance, to evaluate the utility of actigraphy in bipolar sleep monitoring. The present research focused on the inter-episode period given the potential for early detection of illness episodes via changes in sleep.

Materials and methods

Participants

Adult participants between the ages of 18 and 65 years were recruited from advertisements, online bulletins, and psychiatric referrals. For the bipolar disorder group, a diagnosis of bipolar disorder type I (85%) or II (15%) was determined at the baseline visit using the Structured Clinical Interview for the DSM-IV, Non-Patient version (SCID-NP) (23) and inter-episode status was defined using established cutoff scores of ≤ 11 on the Inventory of Depressive Symptomatology–Clinician version (IDS-C) (24) and ≤ 7 on the Young Mania Rating Scale (YMRS) (25). Participants with bipolar disorder were excluded if they met criteria for current substance abuse and/or dependence, given their myriad effects on sleep, if they endorsed active suicidality, or if they were not under the care of a psychiatrist (requirements of the ethics committee). Participants with bipolar disorder were not excluded on the basis of psychiatric comorbidities or pharmacological treatments, given that both are common features of bipolar disorder. Participants with bipolar disorder were also not excluded for insomnia on the basis that sleep disturbance is a hallmark feature of the illness even while inter-episode (7).

Control participants were excluded if they met criteria for any current or lifetime Axis I disorder from the SCID-NP, or major sleep disorders using the Duke Structured Interview for Sleep Disorders (DSISD) based on research diagnostic criteria for sleep disorders (26). All participants were excluded for unstable major illnesses (e.g.,

HIV/AIDS) and severe neurological injuries (e.g., head trauma).

Semi-structured interviews were completed by trained doctoral students and post-doctoral fellows. To assess diagnostic inter-rater reliability, independent interviewers blind to diagnostic group evaluated a randomly selected sample of SCID interviews ($n = 13$). Primary diagnoses matched those made by the original interviewer in all cases ($k = 1.00$).

Five individuals, two from the bipolar disorder group and three from the control group, were excluded from the present analyses after their first laboratory overnight revealed obstructive sleep apnea (Apnea Hypopnea Index > 5) and/or periodic limb movements (Periodic Limb Movement Arousal Index > 15). An additional two individuals were excluded based on PSG or actigraphy malfunction.

The final sample included 27 individuals with bipolar disorder type I ($n = 23$) or type II ($n = 4$), along with 27 age- and sex-matched control individuals without a history of psychiatric or sleep disorder.

Instrumentation

Actigraphy. All participants were equipped with wristwatch actigraphy for each laboratory night [Actiwatch AW-64 (Mini Mitter, Philips Respironics Inc., Bend, OR, USA)]. This device features a sensitivity of 0.05 g and a bandwidth between 3 and 11 Hz, with a sampling frequency of 32 Hz. Actigraphy data were stored in 30-sec epochs to mirror PSG analyses. Following the work of other actigraphy validation papers (15, 18), the following sleep parameters were extracted from actigraphic output: sleep onset latency (SOL), wakefulness after sleep onset (WASO), number of awakenings (nWAK), total sleep time (TST), and sleep efficiency (SE), which was calculated from dividing TST by time in bed. Following previous validation studies (14), sleep onset data were analyzed using the Immobile Minutes algorithm in Actiware 5.57 (Mini Mitter Philips Respironics Inc.), with start and endpoints of rest intervals mirroring lights-out and lights-on times of PSG data. Though the accuracy of the Immobile Minutes algorithm has been called into question (15), it remains a widely used parameter in actigraphy validation research (e.g., 14, 18). Three settings are available within the actigraphy software to detect wake threshold: low [20 activity counts (Act-Low)], medium [40 counts (Act-Med)], and high wake threshold [80 counts (Act-High)], with lower wake thresholds corresponding to higher sensitivity. As these thresholds

have varied considerably in actigraphy research, we had no *a priori* assumptions about the sensitivity threshold that would yield the highest concordance between actigraphy and PSG. Instead, following previous validation studies (14, 15, 18), we analyzed our five primary outcome variables on each of the three wake threshold settings (Act-Low, Act-Med, and Act-High). All actigraphy data were scored by trained researchers blind to diagnostic group.

PSG. All participants underwent two nights of PSG in one of two dedicated laboratory-based PSG setups (Compumedics, Siesta802 Wireless system, Charlotte, NC, USA). The recording montage was comprised of two electrooculograms (EOGs) referenced to a supramedial electrode (1.5 cm above the nasion), and four electroencephalograms (EEGs) referenced to linked mastoids (C3, C4, O1, and O2), and two submental electromyogram (EMG) sensors. On the first night, participants' heart rate and blood oxygen (via oximetry), nasal and oral air flow, thoracic and abdominal effort (respiratory bands), and leg motion (right and left tibial EMGs) were also monitored. The data were acquired and stored on a Dell Optiplex GX280, Intel Pentium, running windows XP at a sampling rate of 512 Hz. PSG data were visually staged off-line using Compumedics Profusion PSG2+. Sleep onset was defined as three consecutive epochs of stage 1 or one epoch of stage 2, 3, 4, or REM. Awakenings were scored according to parameters set forth by Rechtschaffen and Kales (27), as data collection and sleep scoring began prior to the introduction of revised guidelines set forth in late 2007 (28); even so, research has shown these two sets of scoring guidelines applied to non-elderly subjects do not yield significant differences on sleep/wake parameters of interest in the present research (29).

Carefully trained researchers, blind to diagnostic group, scored the PSG data using standard criteria for sleep staging (27). Following the procedures of Perlis et al. (30), five PSG files were checked for inter-rater reliability. At least 90% of epochs were scored identically to the original file.

Sleep diaries. Standard sleep diaries were completed within 10 min of awakening in the morning after laboratory overnights. The sleep diary has been shown to be a reliable measure of these estimates across participants and study days (31, 32). All five variables extracted from both actigraphy and polysomnography (SOL, WASO, nWAK, TST, and SE) were similarly extracted from participants' sleep diaries.

Procedures

All procedures were approved by the University of California, Berkeley Committee for the Protection of Human Subjects. At the baseline visit, which took place in the daytime, doctoral student or postdoctoral interviewers assessed the diagnostic status and symptom severity of participants by administering the SCID, the DSISD, the YMRS, and the IDS-C. Eligible participants then visited the lab on a separate occasion for their first overnight visit (*night one*) that included assessment for the presence of sleep disorders (i.e., sleep apnea and periodic limb movement disorder). Eligible participants were invited back approximately one month later (45.1 ± 33.1 days between visits) for their second overnight (*night two*). As part of a separate study (33), participants underwent a neutral mood induction (34) prior to sleep. For both visits, participants arrived at the laboratory approximately two hours prior to their typical bedtime. The YMRS and IDS-C were administered each night to confirm inter-episode status, which was followed by attaching the PSG equipment. After the PSG hook-up or neutral mood induction, the participant was escorted to the laboratory bedroom and lights were turned out. Participants were awakened at a time of their request, with the constraint that they could not stay in the lab past 9:00 a.m. due to staffing and resource limitations. Upon awakening in the morning, participants completed the sleep diary.

Statistical analysis

All statistical analyses were performed using SPSS version 16 (SPSS Inc., Chicago, IL, USA). The parameters SOL and WASO were positively skewed, and so a square root transformation was applied. Two-way repeated measures analyses of variance (ANOVAs) with Instrument (PSG, Act-Low, Act-Med, Act-High, diary) and Night (first, second) as within-subject factors, and Group (bipolar, control) as the between-subject factor were performed on each of the five sleep parameters. Main effects were evaluated via post-hoc Tukey honestly significant difference (HSD) tests and interactions were evaluated with Bonferroni-adjusted *t*-tests. Greenhouse-Geisser corrections were applied to correct for violations of sphericity, with original degrees of freedom reported below.

Pearson correlations were used to evaluate agreement between each pair of instruments for each sleep parameter. Additionally, because correlations may be artificially influenced by sample characteristics, concordance between PSG and

actigraphy for each sleep parameter was further examined using the Bland–Altman method (35). The Bland–Altman method involves plotting the average of two instruments on the x-axis against the difference between the two instruments on the y-axis for each subject and parameter. After evaluating each plot for systematic bias, mean difference scores, along with standard deviations of the differences, were calculated for each plot. Positive difference scores indicate actigraphy overestimation compared to PSG, whereas negative differences indicate underestimation of actigraphy compared to PSG. Both methods have been used in prior actigraphy validation research.

Results

Participant characteristics

Participant characteristics at the baseline visit are presented in Table 1. There were no significant differences between the groups on age, gender, race/ethnicity, or annual income level. The bipolar disorder group, though established as inter-episode prior to each laboratory night, reported a greater number of both manic (YMRS) and depressed (IDS-C) symptoms than control participants at the baseline visit. Given that all participants were confirmed inter-episode at each laboratory overnight,

Table 1. Participant characteristics

Demographic variables	Bipolar disorder (n = 27)	Controls (n = 27)	χ^2 or <i>t</i>
Age, years, mean (SD)	33.1 (10.3)	38.1 (13.0)	1.6
Gender, female, %	85.2	70.4	1.7
Race/ethnicity, n			1.3
Caucasian	19	15	
Non-Caucasian	8	12	
Annual income, n			1.3
< \$50000	12	17	
> \$50000	11	8	
IDS-C total score, mean (SD)	7.0 (3.8)	2.8 (2.4)	4.5 ^a
YMRS total score, mean (SD)	2.9 (2.1)	0.9 (1.4)	3.9 ^a
Psychotropic medication use, n			
None	1	27	50.1 ^a
Monotherapy	3	0	3.2
Mood stabilizers	19	0	29.3 ^a
Antidepressants	22	0	37.1 ^a
Antipsychotics	13	0	17.1 ^a
Anxiolytics	6	0	6.8 ^b
Hypnotics/sleep agents	1	0	1.0

IDS-C = Inventory of Depressive Symptomatology–Clinician version; SD = standard deviation; YMRS = Young Mania Rating Scale.

^a*p* < 0.001.

^b*p* < 0.01.

we did not control for symptom variation in subsequent analyses.

Instrument concordance across groups and nights

Table 2 summarizes sleep data across the two nights as estimated by each instrument. Five repeated measures ANOVAs were performed for each sleep parameter, with Instrument (PSG, Act-Low, Act-Med, Act-High, diary) and Night (first, second) as within-subject factors, and Group (bipolar, control) as the between-subject factor. Main effects were evaluated using post-hoc Tukey HSD, and interactions were decomposed with independent samples *t*-tests using a Bonferroni-adjusted $\alpha = 0.01$.

The ANOVA performed on SOL revealed a main effect of Instrument ($F_{4,36} = 16.6$, $p < 0.001$, $\eta_p^2 = 0.31$). A post-hoc Tukey HSD test indicated that diary estimates of SOL were significantly higher than PSG and all three actigraphy estimates ($p \leq 0.001$ for all). There were no significant main effects for Group or Night ($F_{4,36} \leq 1.8$, $p \geq 0.19$, $\eta_p^2 \leq 0.05$), nor were interactions present ($F_{4,36} \leq 0.52$, $p \geq 0.60$, $\eta_p^2 \leq 0.01$).

For WASO, there was a main effect for Instrument ($F_{4,31} = 21.1$, $p < 0.001$, $\eta_p^2 = 0.38$). Follow-up tests indicated that the Act-Low algorithm yielded higher estimates of WASO compared to all other instruments ($p \leq 0.001$ for all), that the three actigraphy algorithms differed from one another ($p \leq 0.001$), and that the Act-Low and Act-Med algorithms yielded higher estimates of WASO when compared to diary ($p \leq 0.001$). No significant main effects were observed for Group or Night ($F_{4,31} \leq 2.4$, $p \geq 0.19$, $\eta_p^2 \leq 0.05$), nor were interactions present ($F_{4,31} \leq 0.52$, $p \geq 0.13$, $\eta_p^2 \leq 0.07$).

Analysis of nWAK again showed a main effect for Instrument ($F_{4,33} = 143.1$, $p < 0.001$, $\eta_p^2 = 0.80$). Tukey HSD indicated that diary reports of nWAK were significantly lower than the other four instruments ($p < 0.001$ for all), and Act-Low and Act-Med yielded higher estimates of nWAK when compared to PSG ($p < 0.001$ for both); only the Act-High algorithm did not differ from PSG. There were no significant main effects for Group or Night ($F_{4,33} \leq 0.2$, $p \geq 0.66$, $\eta_p^2 \leq 0.01$), nor were interactions present ($F_{4,33} \leq 1.3$, $p \geq 0.27$, $\eta_p^2 \leq 0.04$).

For TST, a main effect for Instrument was detected ($F_{4,28} = 5.4$, $p = 0.01$, $\eta_p^2 = 0.15$). Follow-up tests revealed that the three actigraphy algorithms differed from one another in accordance with their wake threshold (i.e., Act-Low < Act-Med < Act-High; $p < 0.001$ for all). There were no main effects for Group or Night ($F_{4,28} \leq 1.9$, $p \geq 0.18$, $\eta_p^2 \leq 0.06$), though a

Table 2. Sleep parameters for each recording night

Sleep parameter	Night one					Night two				
	PSG	Act-Low	Act-Med	Act-High	Diary	PSG	Act-Low	Act-Med	Act-High	Diary
Sleep onset latency (min)^a										
Bipolar disorder	12.6 (13.7)	11.8 (14.4)	11.8 (14.4)	11.8 (14.4)	27.5 (19.7)	15.9 (17.9)	12.8 (21.0)	12.8 (21.0)	12.8 (21.0)	27.7 (27.4)
Controls	10.8 (11.7)	18.2 (28.0)	18.2 (28.0)	18.2 (28.0)	24.4 (20.5)	11.6 (13.4)	10.0 (15.9)	10.0 (15.9)	10.0 (15.9)	19.4 (18.4)
Wakefulness after sleep onset (min)										
Bipolar disorder	23.0 (18.0)	46.8 (20.8)	32.3 (17.1)	23.2 (14.9)	27.4 (25.8)	32.3 (29.0)	49.8 (25.3)	33.3 (16.9)	30.7 (11.6)	30.1 (30.7)
Controls	34.6 (26.6)	50.3 (15.4)	34.8 (13.1)	23.3 (11.0)	21.9 (20.6)	25.7 (19.0)	48.5 (15.4)	33.2 (15.7)	22.4 (11.8)	23.1 (19.0)
No. of awakenings										
Bipolar disorder	16.8 (7.6)	29.9 (13.4)	25.0 (11.3)	17.0 (7.4)	2.6 (1.6)	18.8 (8.7)	30.7 (12.2)	27.7 (12.8)	22.0 (11.0)	2.5 (1.6)
Controls	18.0 (7.6)	31.1 (12.1)	27.2 (11.3)	19.5 (8.6)	2.9 (1.4)	20.7 (8.4)	30.4 (11.7)	27.1 (10.8)	20.7 (8.8)	2.8 (1.4)
Total sleep time (min)										
Bipolar disorder	411.9 (85.2)	412.0 (74.5)	426.3 (76.4)	435.4 (77.4)	446.6 (84.3)	426.8 (84.4)	412.5 (87.8)	429.1 (89.4)	438.3 (95.1)	409.8 (88.2)
Controls	370.3 (74.7)	373.3 (70.3)	388.2 (71.9)	400.2 (71.8)	420.6 (69.6)	380.5 (81.2)	371.5 (77.9)	387.7 (79.4)	398.9 (82.0)	408.4 (70.4)
Sleep efficiency (%)										
Bipolar disorder	92.4 (4.7)	87.4 (5.1)	90.4 (4.6)	92.4 (4.3)	88.9 (9.0)	89.2 (8.7)	86.1 (7.0)	89.7 (5.9)	91.9 (5.2)	86.5 (11.8)
Controls	89.4 (7.2)	84.3 (7.2)	87.7 (7.1)	90.5 (6.8)	89.7 (8.2)	90.9 (5.9)	86.2 (6.2)	90.1 (5.5)	92.7 (4.9)	89.9 (6.9)

Values are reported as mean (standard deviation). Act-High = Actiware high threshold algorithm; Act-Low = Actiware low threshold algorithm; Act-Med = Actiware medium threshold algorithm; PSG = polysomnography.

^aSleep onset latency is calculated by a separate immobile minutes algorithm in Actiware and is unaffected by sensitivity thresholds.

significant interaction between Instrument and Night was observed ($F_{4,28} = 5.9$, $p < 0.01$, $\eta_p^2 = 0.16$). Follow-up comparative tests across recording nights revealed that sleep diary estimates of TST decreased from night one to night two [$t(43) = -3.42$, $p < 0.01$], while all other instrument estimates showed no significant changes across recording nights.

The final ANOVA for SE detected a main effect for Instrument ($F_{4,29} = 8.4$, $p < 0.001$, $\eta_p^2 = 0.21$). Follow-up Tukey HSD tests determined that Act-Low yielded a lower estimate of SE compared to PSG and Act-High ($p \leq 0.001$ for both). No main effects or interactions were detected for Group or Night ($F_{4,29} \leq 3.2$, $p > 0.05$, $\eta_p^2 \leq 0.09$).

There was no main effect of recording night observed for any sleep parameter, suggesting that the equipment used to detect sleep disorders on night one and the neutral mood induction introduced on night two did not differentially impact sleep. Data were subsequently collapsed across nights to reflect 108 total comparison points. To evaluate concordance between instruments, mean difference (bias) and standard deviation of bias were plotted for each instrument and each sleep parameter. Figure 1 illustrates these plots, with positive values indicating overestimation of actigraphy compared to PSG and diary. As can be seen, actigraphy estimates were fairly concordant with PSG estimates across the five parameters. When compared with sleep diaries, all three actigraphy algorithms underestimated SOL (bias range 8.1–15.0) and overestimated nWAK (bias range 17.1–28.0) in both the bipolar and control groups. To evaluate the best fitting actigraphy algorithm, mean biases for each parameter were subsequently standardized (z -score) and collapsed to form a single composite representing the bias of each actigraphy algorithm. The Act-Med algorithm showed the best concordance and was used in subsequent analyses.

As a final measure of evaluating concordance between instruments, Pearson correlations were calculated for each sleep parameter based on the Act-Med algorithm. These correlations are presented in Table 3. Given the number of correlations computed ($n = 30$) we adjusted our p -value to 0.01 following previous actigraphy validation work (15). As illustrated in Table 3, significant correlations were observed between actigraphy and PSG for nWAK, TST, and SE across both groups. SOL correlations reached our significance criterion in the control group only, and WASO correlations reached significance in the bipolar disorder group only. Correlations between actigraphy and diary

were low, with a significant correlation observed for TST (both groups) and SOL (bipolar disorder group only). Finally, the correlations between PSG and sleep diary revealed significant correlations for TST (both groups), SE (both groups), and SOL (control group only).

Insomnia and medication use

Discrepancy scores were computed between pairs of instruments (PSG, Act-Med, and diary) for all participants. Independent sample t -tests or Pearson correlations with appropriate Bonferroni corrections were used to compare discrepancy scores based on insomnia presence/absence and medication use.

Insomnia. Insomnia diagnoses were made using the DSISD based on research diagnostic criteria for insomnia (26). Six individuals in the bipolar disorder group met diagnostic criteria for insomnia. There were no significant differences between discrepancy scores based on presence or absence of insomnia, and effect size estimates for all comparisons were small to moderate (Cohen's $d = 0.02$ – 0.70). Furthermore, when considered dimensionally, insomnia severity [as measured by the Insomnia Severity Index (36)] was not associated with discrepancy scores on any parameter after Bonferroni-adjusted correlations were computed ($r \leq 0.17$, $p \geq 0.24$).

Medication use. Only one of 27 participants with bipolar disorder was not taking psychotropic medications; by contrast, three individuals reported taking one medication and 23 reported more than one medication. Of the medications reported, 19 individuals reported taking mood stabilizers, 22 took antidepressants, 13 took antipsychotics, six took anxiolytics, and one individual reported taking hypnotics. There were no significant differences in discrepancy between instruments based on individual classes of medications taken, nor were significant differences observed based on medication administration (monotherapy versus polytherapy). Given the small number of participants in each subgroup, effect sizes (Cohen's d) were simultaneously calculated for each. Effect sizes were in the small-to-moderate range (0.01 to 0.59), with two exceptions; anxiolytic users were more likely to overestimate diary reports of SOL and underestimate SE when compared with actigraphy ($d = 0.77$), and individuals using mood stabilizers had actigraphy estimates that overestimated nWAK compared with PSG ($d = 0.70$).

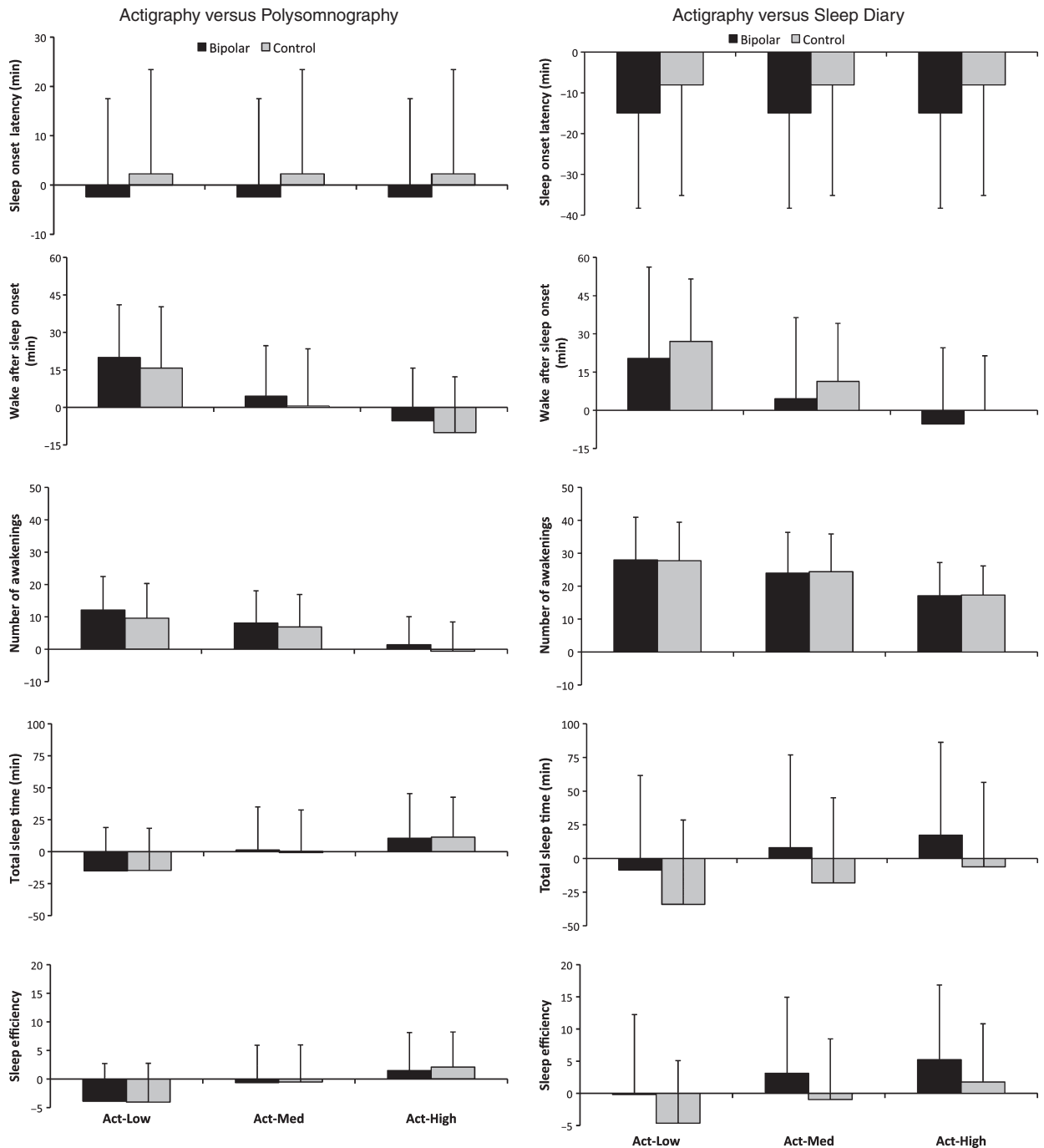


Fig. 1. Bland–Altman estimates of bias and standard deviation of mean bias for each sleep parameter. Bias estimates are between each of three actigraphy threshold algorithms: Actiware low threshold algorithm (Act-Low), Actiware medium threshold algorithm (Act-Med), and Actiware high threshold algorithm (Act-High) and polysomnography and sleep diary.

Discussion

The present research evaluated concordance between actigraphy, PSG, and sleep diary in individuals diagnosed with bipolar disorder and matched controls. Aim 1 evaluated group differences in instrument concordance. Based on previous studies documenting reduced concordance between actigraphy

and PSG in insomnia, coupled with high rates of sleep disturbance in bipolar disorder, we predicted that instrument concordance would be reduced in a bipolar sample. Contrary to this hypothesis, we found no group differences in actigraphic estimates of sleep parameters when compared to PSG and sleep diaries, and effect sizes were trivial. Actigraphy concordance was equally

Table 3. Pearson correlations between instruments

Instrument comparison	Sleep onset latency	Wakefulness after sleep onset	No. of awakenings	Total sleep time	Sleep efficiency
ACT versus PSG					
Bipolar disorder	0.33 ^a	0.59 ^c	0.58 ^c	0.92 ^c	0.49 ^b
Controls	0.41 ^b	0.35 ^a	0.39 ^b	0.91 ^c	0.51 ^c
ACT versus Diary					
Bipolar disorder	0.43 ^b	0.07	-0.10	0.68 ^c	0.24
Controls	0.22	0.17	-0.10	0.57 ^c	0.13
PSG versus Diary					
Bipolar disorder	0.17	0.23	0.02	0.71 ^c	0.48 ^b
Controls	0.62 ^c	0.28 ^a	-0.03	0.74 ^c	0.53 ^c

Actigraphy (ACT) values here are calculated using the medium threshold algorithm in Actiware™. PSG = polysomnography.

^a $p < 0.05$.

^b $p < 0.01$.

^c $p < 0.001$.

accurate for individuals with bipolar disorder and control individuals. These results held even when excluding the four individuals diagnosed with bipolar disorder type II from analyses.

Results demonstrated the utility of actigraphy in estimating sleep parameters, particularly when compared to PSG. Associations were observed between actigraphy and PSG in estimating WASO, nWAK, TST, and SE (though it should be noted that, after omitting bipolar II disorder participants from analyses, WASO no longer met our significance threshold). Actigraphic estimates of SOL in the bipolar disorder group, though not significantly different from PSG estimates, did not reach our Bonferroni-adjusted criteria for a strong correlation ($p < 0.01$), replicating previous validation findings (15, 16) and adding to previous recommendations (15) for a re-evaluation of the sleep onset algorithm.

Actigraphic concordance was more variable when compared with sleep diaries. Significant differences were observed between actigraphy and diary estimates of SOL and nWAK, whereby actigraphy underestimated SOL and overestimated nWAK when compared to diaries. Participants reported longer SOL on diaries when compared to actigraphy and PSG, a common finding in previous research (15, 17, 18, 37, 38) perhaps explained by differences in the sleep stages in which objective and subjective sleep onset are determined (typically stage 1 for objective estimates and stage 2 for subjective estimates) (39). The tendency for individuals to underestimate nocturnal awakenings on sleep diaries compared to actigraphy and PSG has also been observed in both healthy (38, 40) and in sleep-disordered samples (15, 18), perhaps explained by the temporal difference between a recalled nocturnal arousal and a 30-sec epoch scored as wake (15). Such findings call into

question the comparability between subjective and objective estimates of nocturnal awakenings.

The Bland–Altman method of plotting bias estimates illustrated differences between actigraphy algorithms. Specifically, the low-wake threshold algorithm tended to yield larger biases compared to diary and PSG, and the medium-wake threshold algorithm was determined to be the best fit when collapsing across the five sleep parameters. Previous research has tended to utilize both the low and medium threshold algorithms; based on our findings, we recommend the medium threshold setting within MiniMitter Actiware software for bipolar sleep monitoring. McCall and McCall (41), in a recent comparison of actigraphy, PSG, and sleep diary in depressed insomniacs, used the Bland–Altman method to identify systematic discrepancies between instruments. These researchers reported that differences between instruments increased as SOL and WASO increased, suggesting that sleep measurement is less concordant as values grow more extreme. In our own data, we found evidence for this pattern only when comparing sleep diaries to actigraphy or PSG; as noted above, actigraphy and PSG tended to be fairly concordant, even as SOL and WASO increased in duration.

The second aim of the present study, included on an exploratory basis, concerned the influence of insomnia and of medication use on concordance estimates. Given that insomnia is prevalent in bipolar disorder, we evaluated its presence as a possible contributor to the relationship between actigraphy and PSG/diary estimates. Though the number of individuals meeting criteria for insomnia was small, results suggest that concordance was independent of insomnia presence and pointed to the utility of actigraphy in bipolar sleep monitoring.

Medications present a challenge for research in bipolar disorder. We conducted analyses by medication class and administration type in the bipolar disorder group. We did not find evidence that medication classes influenced sleep parameters, though given the small sample size of these subdivided groups this finding should be interpreted with caution. However, effect size estimates suggested that anxiolytic users were more likely to overestimate SOL on sleep diaries when compared to non-anxiolytic users, perhaps explained by the relationship between anxiety disorders and subjective sleep disturbance (42). Notably, anxiolytic use did not influence objective estimates of sleep latency; instead, and perhaps paradoxically, only self-report estimates of sleep onset latency were prolonged in anxiolytic users.

For prior actigraphy validation paradigms sufficiently powered to detect medication effects (e.g., 15), hypnotics have been shown to reduce actigraphic accuracy compared to PSG, presumably because of their sedative properties. Recall that actigraphy infers sleep based on movement, and any agent that reduces nocturnal mobility is likely to overscore sleep. There are multiple medications commonly used to treat bipolar disorder with sedative properties [e.g., atypical antipsychotics (43)]. However, medication effects are not straightforward; several medications commonly prescribed for bipolar disorder can be associated with either sedating or alerting side effects [e.g., aripiprazole, venlafaxine, sertraline, and zonisamide (44, 45)]. Furthermore, validation of actigraphy in medication-free bipolar samples is unrepresentative and lacks generalizability.

Several limitations should be noted. First, though our objective was to describe concordance between sleep parameters that provide a summary of the night's sleep (e.g., WASO, TST, and SE), we encourage future research to conduct epoch-by-epoch analyses to examine sensitivity and specificity (sleep and wake accuracy, respectively). Second, our design evaluated instruments across two nights of recording and cannot address the significant night-to-night variability that exists in populations with disturbed sleep; we encourage future validation work to compare actigraphy and sleep diary estimates over a longer recording period. Though it is possible that participants' sleep was affected by the unfamiliar laboratory environment, we compared diary reports of SOL, WASO, and TST in the laboratory to a week of sleep diaries (not reported in this paper) kept at home by the same participants; we found no evidence for differences between self-reported sleep in the lab and in the home environment. Third, though our final eligible sam-

ple included 27 inter-episode individuals with bipolar disorder, a larger sample is needed to verify medication subgroup effects, and our results on the effects of medications should be interpreted with caution. We note that our sample included individuals diagnosed with both bipolar I and II disorder and, though main findings did not change when excluding individuals with bipolar II from analyses, it is as yet unclear what features of sleep may differ between these two subtypes. Finally, it should be emphasized that our results are specific to the brand and algorithm of actigraphy monitoring. Here, we used MiniMitter Actiwatches with data analysis conducted in Actiware 5.57 on each of three threshold settings. It should be underscored that our favorable concordance findings are specific to these parameters, and will not necessarily generalize to other actigraphy makes or algorithms.

The present research suggests actigraphy is a valid instrument for estimating sleep and wakefulness in a bipolar population, despite sedating medication use and the presence of clinical sleep disturbance (insomnia). Though not considered here, actigraphy may be of further interest to both researchers and clinicians because it allows for observation of circadian rhythmicity and circadian disruption across the 24-hour sleep/wake cycle.

Acknowledgements

This project was supported by a National Institute of Mental Health Grant No. R34 MH080958 (AGH), a National Science Foundation Graduate Research Fellowship Grant (KAK), and Ruth L. Kirschstein National Research Service Awards (NRSA) Institutional Research Training Grant T32 in Affective Science (LST and JG).

Disclosures

The authors of this paper do not have any commercial associations that might pose a conflict of interest in connection with this manuscript.

References

1. Merikangas KR, Akiskal HS, Angst J et al. Lifetime and 12-month prevalence of bipolar spectrum disorder in the National Comorbidity Survey replication. *Arch Gen Psychiatry* 2007; 64: 543–552.
2. Gruber J, Harvey AG, Wang PW et al. Sleep functioning in relation to mood, function, and quality of life at entry to the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *J Affect Disord* 2009; 114: 41–49.
3. Millar A, Espie C, Scott J. The sleep of remitted bipolar outpatients: a controlled naturalistic study using actigraphy. *J Affect Disord* 2004; 80: 145–153.
4. Ehlers CL, Frank E, Kupfer DJ. Social zeitgebers and biological rhythms: a unified approach to understanding

- the etiology of depression. *Arch Gen Psychiatry* 1988; 45: 948–952.
5. Frank E. Interpersonal and social rhythm therapy: managing the chaos of bipolar disorder. *Biol Psychiatry* 2000; 48: 593–604.
 6. Mansour HA, Monk TH, Nimgaonkar VL. Circadian genes and bipolar disorder. *Ann Med* 2005; 37: 196–205.
 7. Harvey AG, Schmidt DA, Scarna A, Semler CN, Goodwin GM. Sleep-related functioning in euthymic patients with bipolar disorder, patients with insomnia, and subjects without sleep problems. *Am J Psychiatry* 2005; 162: 50–57.
 8. Jackson A. A systematic review of manic and depressive prodromes. *J Affect Disord* 2003; 74: 209–217.
 9. Bauer M, Grof P, Rasgon N, Bschor T, Glenn T, Whybrow PC. Temporal relation between sleep and mood in patients with bipolar disorder. *Bipolar Disord* 2006; 8: 160–167.
 10. Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak C. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep* 2003; 26: 342–392.
 11. Mercer J, Bootzin R, Lack L. Insomniacs perception of wake instead of sleep. *Sleep* 2002; 25: 564–571.
 12. Mullaney D, Kripke D, Messin S. Wrist-actigraphic estimation of sleep time. *Sleep* 1980; 3: 83.
 13. Jean-Louis G, Kripke DF, Cole RJ, Assmus JD, Langer RD. Sleep detection with an accelerometer actigraph: comparisons with polysomnography. *Physiol Behav* 2001; 72: 21–28.
 14. Paquet J, Kawinska A, Carrier J. Wake detection capacity of actigraphy during sleep. *Sleep* 2007; 30: 1362.
 15. Lichstein KL, Stone KC, Donaldson J et al. Actigraphy validation with insomnia. *Sleep* 2006; 29: 232–239.
 16. Vallieres A, Morin CM. Actigraphy in the assessment of insomnia. *Sleep* 2003; 26: 902–906.
 17. Verbeek I, Arends J, Declerck G, Beecher L. Wrist actigraphy in comparison with polysomnography and subjective evaluation in insomnia. *Sleep-Wake Research in The Netherlands* 1994; 5: 163–169.
 18. Kushida CA, Chang A, Gadkary C, Guillemainault C, Carrillo O, Dement WC. Comparison of actigraphic, polysomnographic, and subjective assessment of sleep parameters in sleep-disordered patients. *Sleep Med* 2001; 2: 389–396.
 19. Jones SH, Hare DJ, Evershed K. Actigraphic assessment of circadian activity and sleep patterns in bipolar disorder. *Bipolar Disord* 2005; 7: 176–186.
 20. Benedetti F, Dallaspesza S, Fulgosi MC, Barbini B, Colombo C, Smeraldi E. Phase advance is an actimetric correlate of antidepressant response to sleep deprivation and light therapy in bipolar depression. *Chronobiol Int* 2007; 24: 921–937.
 21. Jones SH, Tai S, Evershed K, Knowles R, Bentall R. Early detection of bipolar disorder: a pilot familial high-risk study of parents with bipolar disorder and their adolescent children. *Bipolar Disord* 2006; 8: 362–372.
 22. Buysse DJ, Ancoli-Israel S, Edinger JD, Lichstein KL, Morin CM. Recommendations for a standard research assessment of insomnia. *Sleep* 2006; 29: 1155–1173.
 23. First MB, Spitzer RL, Gibbon M, Williams JBW. Structured Clinical Interview for DSM-IV Non-patient Edition (SCID-NP). New York: New York State Psychiatric Institute, 1995.
 24. Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med* 1996; 26: 477.
 25. Young R, Biggs J, Ziegler V, Meyer D. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry* 1978; 133: 429–435.
 26. Edinger JD, Bonnet MH, Bootzin RR et al. Derivation of research diagnostic criteria for insomnia: report of an American Academy of Sleep Medicine Work Group. *Sleep* 2004; 27: 1567–1596.
 27. Rechtschaffen A, Kales A. A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. Los Angeles: Brain Information Service/Brain Research Institute, 1968.
 28. Iber C, Ancoli-Israel S, Chesson A, Quan S. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specification, 1st edn. Westchester: American Academy of Sleep Medicine, 2007.
 29. Moser D, Anderer P, Gruber G et al. Sleep classification according to AASM and Rechtschaffen & Kales: effects on sleep scoring parameters. *Sleep* 2009; 32: 139–149.
 30. Perlis ML, Smith MT, Orff HJ, Andrews PJ, Gillin JC, Giles DE. The effects of an orally administered cholinergic agonist on REM sleep in major depression. *Biol Psychiatry* 2002; 51: 457–462.
 31. Morin CM, Espie CA. *Insomnia: A Clinical Guide to Assessment and Treatment*. New York: Kluwer Academic/Plenum Publishing, 2003.
 32. Bootzin R, Engle-Friedman M. The assessment of insomnia. *Behav Assess* 1981; 3: 3107–3126.
 33. Talbot LS, Hairston IS, Eidelman P, Gruber J, Harvey AG. The effect of mood on sleep onset latency and REM sleep in interepisode bipolar disorder. *J Abnorm Psychol* 2009; 118: 448–458.
 34. Eich E, Macaulay D, Ryan L. Mood dependent memory for events of the personal past. *J Exp Psychol* 1994; 123: 201–215.
 35. Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 327: 307–310.
 36. Bastien C. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med* 2001; 2: 297–307.
 37. Van Den Berg JF, Van Rooij FJA, Vos H et al. Disagreement between subjective and actigraphic measures of sleep duration in a population-based study of elderly persons. *J Sleep Res* 2008; 17: 295–302.
 38. Lockley SW, Skene DJ, Arendt J. Comparison between subjective and actigraphic measurement of sleep and sleep rhythms. *J Sleep Res* 1999; 8: 175–183.
 39. Tryon WW. Issues of validity in actigraphic sleep assessment. *Sleep* 2004; 27: 158–165.
 40. Baekeland F, Hoy P. Reported vs. recorded sleep characteristics. *Arch Gen Psychiatry* 1971; 24: 548–551.
 41. McCall C, McCall WV. Comparison of actigraphy with polysomnography and sleep logs in depressed insomniacs. *J Sleep Res* 2012; 21: 122–127.
 42. Ramsawh HJ, Stein MB, Belik SL, Jacobi F, Sareen J. Relationship of anxiety disorders, sleep quality, and functional impairment in a community sample. *J Psychiatr Res* 2009; 43: 926–933.
 43. Kane J, Sharif Z. Atypical antipsychotics: sedation versus efficacy. *J Clin Psychiatry* 2008; 69: 18–31.
 44. GlaxoSmithKline. *Psychotropic Prescribing Guide*, 8th edn. New York: Thomson PDR, 2005.
 45. PDR Staff. *Physicians Desk Reference 2008: Hospital/Library Version*, 62nd edn. Montvale, NJ: Physician's Desk Reference, 2007.