Title

Correlates of poor sleep based upon wrist actigraphy data in bipolar disorder.

Authors

Kaufmann, Christopher N
Lee, Ellen E
Wing, David
et al.

Publication Date

2021-09-01

DOI

10.1016/j.jpsychires.2021.06.038

Peer reviewed

HHS Public Access

Author manuscript

J Psychiatr Res. Author manuscript; available in PMC 2022 September 01.

Published in final edited form as:

J Psychiatr Res. 2021 September ; 141: 385–389. doi:10.1016/j.jpsychires.2021.06.038.

Correlates of Poor Sleep Based Upon Wrist Actigraphy Data in Bipolar Disorder

Christopher N. Kaufmann, PhD, MHS[a,*], **Ellen E. Lee, MD**[b,c,d], **David Wing, MS**[e], **Ashley N. Sutherland, MA**[b], **Celestine Christensen, BS**[f], **Sonia Ancoli-Israel, PhD**[b,c], **Colin A. Depp, PhD**[b,c,d], **Ho-Kyoung Yoon, MD, PhD**[g], **Benchawanna Soontornniyomkij, PhD**[b], **Lisa T. Eyler, PhD**[b,c]

[a]Division of Epidemiology and Data Science in Gerontology, Department of Aging and Geriatric Research, University of Florida College of Medicine, Gainesville, FL

[b]Department of Psychiatry, University of California San Diego, La Jolla, CA, USA

[c]Stein Institute for Research on Aging, University of California San Diego, La Jolla, CA, USA

[d]San Diego VA Healthcare System, La Jolla, CA, USA

[e]Center for Wireless and Population Health Systems, University of California San Diego, La Jolla, CA, USA

[f]State University of New York Downstate College of Medicine, Brooklyn, NY, USA

[g]Department of Psychiatry, Korea University College of Medicine, Seoul, South Korea

Abstract

BACKGROUND: Wrist-worn actigraphy can objectively measure sleep, and has advantages over self-report, particularly for people with Bipolar Disorder (BD) for whom self-reports might be influenced by affect. Clinically useful data reduction approaches are needed to explore these complex data.

METHODS: We created a composite score of sleep metrics in BD based on 51 BD and 80 healthy comparison (HC) participants. Subjects wore an actigraph for up to 14 consecutive 24-hour periods, and we assessed total sleep time (TST), wake after sleep onset (WASO), percent sleep (PS), and number of awakenings (NA). We focused on participants who had at least 5 nights of actigraphy data. We computed z-scores for within-person means of sleep measures for BD

[*]**Corresponding Author:** Christopher N. Kaufmann, PhD, MHS, Assistant Professor, Division of Epidemiology and Data Science in Gerontology, Department of Aging and Geriatric Research, University of Florida, 2004 Mowry Rd, Gainesville, FL 32603, P: 323-528-2786, ckaufmann@ufl.edu.

CONFLICT OF INTEREST

All authors report no conflicts of interest.

AUTHOR STATEMENT

CNK conceived of project, completed data analyses, and wrote the manuscript. LTE oversaw analyses, and wrote the manuscript. EEL, DW, ANS, CC, SAI, CAD, HKY and BS provided feedback on the manuscript.

subjects versus HCs, which were averaged to create a composite measure. We correlated this composite with participant characteristics, and used LASSO regression to identify sleep measures best explaining variability in identified correlates.

RESULTS: Sleep measures and the composite did not differ between BDs and HCs; however, there was considerable variability in z-scores among those with BD. In BDs, the composite score was higher in women ($t_{(49)} = 2.28$, $p = 0.027$) and those who were employed ($t_{(34)} = 2.34$, $p = 0.025$), and positively correlated with medication load ($r = 0.41$, $p = 0.003$) while negatively correlated with Young Mania Rating Scale (YMRS; $r = -0.35$, $p = 0.030$). In Lasso regression, TST and NA best explained medication load while PS best explained employment and YMRS.

CONCLUSION: While a composite score of sleep metrics may provide useful information about sleep quality globally, our findings suggest that selection of theory-driven sleep measures may be more clinically meaningful.

Keywords

sleep; bipolar disorder; actigraphy; data reduction

INTRODUCTION

In recent years, there has been an explosion in the availability of consumer wearables (e.g., Fitbit, Apple watch) (Voets, 2013) which are able, through accelerometry, heart rate, and “real time” physiologic signal, to provide insights on lifestyle, including sleep. The ubiquity of these devices means that a wealth of data is now available for understanding sleep in diverse populations, including those with serious mental illness such as bipolar disorder (BD). Sleep disturbance is highly prevalent in BD and associated with psychopathology (Harvey et al., 2009; Sylvia et al., 2018; Sylvia et al., 2012). Using objective forms of sleep measurement (e.g., polysomnography, actigraphy) in BD may be more accurate than clinical interviews, as self-report of sleep may be influenced not only by recall, but by current affective states inherent in BD (Baillet et al., 2016). While polysomnography is the most validated means of measuring sleep, wearables can generate important data about sleep quality that has been validated against polysomnography (Ancoli-Israel et al., 2015; Cole et al., 1992; Sadeh et al., 1994). Indeed, past studies show altered sleep patterns as measured by actigraphy in BD vs. healthy subjects (see De Crescenzo et al. (2017) as example).

Accelerometry can estimate a number of sleep parameters, including total sleep time (TST; time a subject was asleep in minutes), wake after sleep onset (WASO; time awake while in bed in minutes after sleep onset), percent sleep (PS; percent of time spent asleep at night), number of awakenings (NA; number of times awaken at night) amongst others (Ancoli-Israel et al., 2003). While these individual parameters, by themselves, are clinically relevant and provide important data to identify mechanisms for outcomes, there are some statistical challenges, including Type 1 error which can be problematic when looking at multiple outcomes (e.g., different affective states, psychiatric symptoms, inflammatory biomarkers) as are commonly measured in studies of BD (Ketter, 2010).

Distilling sleep parameters into one joint construct may help identify those with poor sleep and in need of sleep intervention. Approaches such as machine learning can cluster people based upon patterns of sleep, and past studies in sleep using these methods have shown promising results (Sathyanarayana et al., 2016; Willetts et al., 2018; Wolz et al., 2017). However, these approaches have limitations in their applicability and generalizability to individual patients (Graham et al., 2019), and are often hard to communicate to those unfamiliar with these methods, limiting their clinical utility. Other investigators have sought to identify actigraphic measures most important in distinguishing sleep quality across subjects. For example, Natale et al. (2009) used linear discriminant function analysis to find that TST, sleep onset latency, and NA were the best combination of actigraphy statistics differentiating those with insomnia from those with healthy sleep patterns. In neuropsychological assessment, cognitive test scores are commonly standardized as compared to a normative population, and can then be merged into a composite. Using a similar approach, but with actigraphic sleep measurements, may yield a clinically meaningful way to summarize sleep quality across measures among those with BD.

Sleep quality is associated with a number of demographic and clinical characteristics. For example, in community-based samples, poor sleep is associated with lower overall health quality (Medic et al., 2017), increased risk for depression (Zhai et al., 2015), higher body mass index (BMI) (Rahe et al., 2015), lower cognition (Lo et al., 2016), higher levels of inflammation (Irwin et al., 2016), among other negative health outcomes. In BD, poor sleep is associated with these same characteristics but at a more pronounced level; and poor sleep is also associated with worsening BD symptoms (e.g., depression and mania) (Barbini et al., 1996; Fava and Kellner, 1991; Gruber et al., 2011; Jackson et al., 2003; Perlman et al., 2006), poorer overall mood (Cretu et al., 2016), impaired cognition (Bradley et al., 2020), and greater inflammation (Dolsen et al., 2018), among others. To have construct validity and be clinically meaningful, a composite sleep measure in BD should be correlated with demographic and clinical characteristics consistent with past studies of sleep quality.

In this study we created a composite score for sleep across accelerometer-derived sleep variables in a sample of those with BD. The main aims of this study were to identify the clinical utility of this composite measure by examining demographic, clinical, and biological correlates within subjects, and to identify the sleep variables most contributing to associations between the composite measure and these associated variables. We hypothesized that better sleep quality as seen in our composite measure would be associated with fewer depressive and mania symptoms, greater medication load, and lower cognition and inflammation. We hypothesized that any observed correlations would be driven by multiple sleep indices included in the overall composite score.

MATERIAL AND METHODS

Study design and participants

Data came from a longitudinal study of cognition and inflammation in BD. We recruited those with BD and HCs from outpatient clinics, community settings, and other research studies at UC San Diego. BD was defined as a diagnosis of Bipolar I or II DSM-IV Disorder receiving outpatient care. Exclusion criteria included: acute illness or pregnancy, a recent

vaccination, history of various health conditions (dementia, seizures, Parkinson's, stroke, or head trauma), cancer treatment in past, diabetes/hypertension that is not controlled, among others. Among HCs, we also excluded individuals who had a history of DSM-IV Axis I disorders, previous use of psychotropic medications, as well as having a first degree relative with history of depression, BD, or schizophrenia. Each year of the study, participants were asked to complete a 2-week burst of assessments consisting of three in-person visits and completion of up to 14 24-hour periods of wrist-worn actigraphy. We focused on a subset of 51 persons with BD and a comparable group of healthy controls [HCs] (n=80) who had valid wrist actigraphy data which we defined as having actigraphy data recorded for at least five nights. Of note, some individuals had more than one period of actigraphy assessment (e.g., in years 1 and 2 of the study) or did not have valid actigraphy data for year 1 but valid data for year 2. Consequently, we analyzed the first valid actigraphy assessment period available for each participant. This study was approved by the UCSD Institutional Review Board and was carried out in accordance with the Declaration of Helsinki. All participants completed informed consent prior to study involvement.

Actigraphy data and processing

Data were gathered from wrist-worn Actisleep-BT device (Actigraph, Pensacola, FL) which measures raw acceleration data in gravitational units (g's) using a tri-axial accelerometer sampling at 30Hz. Over continuous 24-hour wear periods, this device can be used to monitor movement allowing for estimation of sleep/wake patterns similar to methods described by Ancoli-Israel et al. (2003). In-bed and out-of-bed times were set based upon both actigraphy data and information from morning surveys on a cell phone survey. If participant sleep records were missing entirely, sleep onset and awake time were manually determined by a specially trained research assistant using the detection methods outlined by Full, et al. (Full et al., 2018). TST, WASO, PS, and NA were computed based upon these in/out bed intervals. To construct our composite score, in the absence of publicly-available norms, we chose to normalize BD subjects' sleep measures based upon the HCs as a normative sample. We first computed means and standard deviations of actigraphy sleep measures of TST, WASO, PS, and NA in the HC group. Based upon this, we created standardized z-scores in the BD group by subtracting each BD individual's sleep measure (for TST, WASO, PS, and NA) from the means of measures in HCs, and dividing by the SDs in the HC group. We multiplied WASO and NA by −1 to keep variables in the same direction (higher scores = better sleep). We then computed the mean of z-scores across all measures to create our composite.

Measures

We also assessed age (years), education (years), gender (female, male), minority status (Caucasian, non-Caucasian), marital status (single/divorced/separated/widowed, married/cohabitating), employment (which we categorized as employed full- or part-time, unemployed including on disability), depression (using the Patient Health Questionaire-9, PHQ-9) (Kroenke et al., 2001), cognitive functioning (the Measurement and Treatment Research to Improve Cognition Schizophrenia [MATRICS] Consensus Cognitive Battery (Nuechterlein and Green, 2006) global cognitive and working memory [T-scores]), body mass index (BMI), the Perceived Stress Scale total score (Cohen et al., 1983), sleep medication use as reported on item 7 (component # 6) of the Pittsburg Sleep Quality Index

(PSQI) (Buysse et al., 1989), age of onset, duration of illness, Young Mania Rating Scale (YMRS) (Young et al., 1978), Hamilton Depression Rating Scale (HAM-D) (Hamilton, 1960), Brief Psychiatric Rating Scale (BPRS) (Overall and Gorham, 2016), and global assessment of functioning (GAF), as well as a computed medication load variable that combines information about the number, and relative dose, of all psychotropic medications based on methods previously discussed elsewhere (Versace et al., 2008). Additionally, subjects were asked to answer daily surveys sent to a mobile device regarding use of alcohol and cannabis. We computed the percentage of surveys answered in which the subject indicated they used alcohol or cannabis that day. Participants also gave three fasting blood samples over a two-week period, and biomarkers of inflammation were assayed using a Meso Scale Discovery (MSD) MULTI-SPOT® Assay System and processed on a SECTOR Imager 2400 instrument (Rockville, MD, USA). We focused on the Interleukin-6 (IL-6), tumor necrosis factor (TNF)-alpha, and C-reactive protein (CRP) inflammatory markers as these biomarkers have been shown to be associated with sleep quality (Irwin et al., 2016). We log-transformed inflammatory values and averaged the results across the three blood draws for analyses.

Analyses

After computing the individual sleep statistics and composite z-scores, we sought to determine whether the sleep was of overall worse quality in BDs as compared to HCs by testing if the z-scores were different from 0 (no difference between groups). We then assessed the correlation of the composite with demographic and clinical variables using Pearson correlations (continuous measures) and t-tests (binary/categorical variables). Finally, for variables significantly correlated with the composite sleep z-score, we conducted LASSO regression (Ahrens et al., 2018, 2019) to identify the individual sleep statistics most contributing to this correlation. The demographic and clinical variables served as outcomes and predictors were the z-scores for TST, WASO (which was inverse coded), PS, and NA (also inverse coded). LASSO identified the sleep metrics or combination that best predicted demographic and clinical variables. All analyses were completed in Stata SE version 15 (Stata Corp, College Station, TX).

RESULTS

The mean age of our BD sample was 47.9 (SD = 9.18) years and mean education was 14.4 (SD = 2.10) years. Over half were female (61%), 73% were Caucasian, and 28% were married or cohabitating. Compared to HCs, the BD group had fewer years of education and were less likely to be employed and married or cohabitating (all p's ≤ 0.038). Groups did not differ by age, gender, and minority status.

Figure 1 shows the distribution of computed z-scores of all BD subjects for TST, WASO, PS, and NA, as well as the composite (the mean of TST, WASO, PS, and NA z-scores) based upon respective means from the HC group. While the NA z-score differed statistically significantly from 0 (mean = 0.35, SD = 1.18, p = 0.039) suggesting more awakenings in the BD group, the composite and most individual z-scores did not (Composite: mean = 0.13, SD = 0.98, p = 0.340; TST: mean = 0.05, SD = 1.44, p = 0.817; WASO: mean = 0.09, SD =

1.08, $p = 0.570$; PS: mean = 0.04, SD = 1.25, $p = 0.806$) indicating comparability between the BD and HC groups. In the BD group, there was substantial inter-individual variability in z-scores by about 1 SD.

Table 1 presents correlations between the composite z-score and demographic and clinical variables. Higher scores on the composite z-score was associated with gender (women had a higher composite; $t_{(49)} = 2.28$, $p = 0.027$), employment (those employed had a higher composite; $t_{(34)} = 2.34$, $p = 0.025$), higher medication load (r = 0.41, $p = 0.003$), and lower YMRS score (r = −0.35, $p = 0.030$). Although more than one individual sleep metric was significantly related to each of these factors, results from LASSO regression which included all of them individually in the same predictive model identified TST as most contributing to correlations with medication load (r = 0.51, $p < 0.001$); PS contributing most to correlations for employment ($t_{(34)} = 2.62$, $p = 0.013$), and YMRS (r = −0.35, $p = 0.029$); and NA most contributing to medication load (r= 0.26, p=0.070; Supplemental Table 1) but this was not significant.

DISCUSSION

Past research on sleep in BD has focused on individual actigraphy measures (e.g., TST, WASO), and there is a need for a global measure of sleep derived from actigraphy. In this study, we explored whether a composite score would relate to expected clinical, cognitive, and biological factors and whether including multiple sleep variables was important to the observed relationships. While the composite score was related to variables in BD including gender, employment, medication load, and mania symptoms, these associations were for the most part driven by only one of the individual sleep metrics with the exception of medication load for which TST and NA jointly contributed. Our results suggest a composite score does not yield gains in predictive power over individual sleep metrics. In studies of BD, it may be more appropriate to choose metrics to examine based on theory rather than summarizing multiple sleep metrics together.

Our finding that a composite score is not as informative as examining individual sleep metrics alone may be explained by a number of factors. We focused on the averaged selected sleep indices measured over a two-week period, and it may be that examining changes in patterns of sleep over time, rather than means of sleep indices, could better capture more variability and thus compute a more nuanced composite score. Our goal, however, was to identify an intuitive approach for summarizing poor sleep in clinical practice, and thus we only focused on commonly used and often readily available sleep metrics. It is possible, however, that incorporating measures of circadian patterns would better support a composite measure approach. Additionally, population-based norms of actigraphic sleep indices do not currently exist, and thus we chose to use our healthy control sample as the norms for construction of z-scores. It is important that more research be conducted to generate norms for the purpose of computing composite measures as we sought to do in this study.

While a sleep composite may not provide additional utility beyond that of individual measures themselves, there was substantial variability in z-scores indicating there are persons with BD who may have worse or better sleep as compared to HCs by as much

as one standard deviation. This highlights the need for future research to identify cut-points based on z-scores identifying the worse sleepers. Similarly, it may be important to examine how daily fluctuations in sleep could be incorporated into a composite score. In BD, sleep often varies night-to-night (Kaufmann et al., 2016) and it could be that these nightly fluctuations are representative of poor sleep rather than simply average sleep measures across nights.

While clinical relationships to the composite z-score were generally driven by only one sleep measure, we found correlations with clinical variables pertinent to BD confirming previous literature. A higher score on our composite measure (e.g., better sleep) was associated with being female, being unemployed, having greater medication load and lower mania symptomatology. Similarly, studies in BD show that poor sleep (as defined by self-report or *individual* actigraphy statistics) is associated with worsening BD symptoms (e.g., depression and mania) (Barbini et al., 1996; Fava and Kellner, 1991; Gruber et al., 2011; Jackson et al., 2003; Perlman et al., 2006), among other correlates. Based upon LASSO regression, TST and NA most contributed to the correlation with medication load which may be reflective of the sedating properties of many psychotropic medications (both increasing sleep duration and lowering number of awakenings). PS contributed most to the correlations with employment, and mania symptoms, which may relate to sleep fragmentation and variability that has previously been shown to be associated with these variables (Gold and Sylvia, 2016).

Because wrist actigraphy is easily administered and is less invasive compared to an inlab sleep evaluation and is easier to get longitudinal measures over the span of weeks, it is important for future research to identify other approaches that could reduce these voluminous data to actionable insights, especially for treating patients with BD where management of sleep is paramount. Clinicians could utilize a composite measure to identify patients with poor sleep overall and triage these patients to appropriate sleep treatment options based upon their individual sleep metrics. In the future, as such accelerometry data becomes available in many different populations, it may soon be possible to identify when poor sleep begins to emerge with the possibility of predicting a mood episode to offer just-in-time clinical interventions. More research is needed to develop tools using these data for future prediction of events for clinical monitoring.

Our study has some limitations. First, our analysis was cross-sectional and retrospective. Research should explore ways that changes in sleep quality longitudinally may be incorporated in this composite measure. Second, we focused on the means of sleep parameters as our main purpose was to examine an intuitive composite score for poor sleep. Future studies may want to identify whether night-to-night variability in sleep or circadian patterns can improve a composite measure (Ancoli-Israel et al., 2003). Third, in the absence of published norms for actigraphic sleep measures for healthy individuals of comparable age to our BD sample, we used our own HC sample as the normative group. To the extent that our HC sample was relatively small (n=80) and participants were not selected on the basis of having no reported sleep abnormalities, this may have introduced some bias into the composite scores. It may be important for future studies to compare this method to other approaches which do not use a normative sample (e.g., linear discriminant function).

Fourth, our sample size was small and given the number of correlates assessed with the composite score, there is a possibility of significant findings due to chance alone. However, our study evaluates a potential way to combine actigraphic measurements, and future studies with larger samples may help examine this further. Finally, we did not have measures of sleep apnea which may contribute to disturbed sleep (especially sleep fragmentation as characterized by PS, WASO, and NA). However, sleep apnea is often undiagnosed (Simpson et al., 2013), and in a clinical setting, clinicians may need to base their assessment of sleep on wrist actigraphy alone.

In conclusion, we found that while a sleep composite measure based upon actigraphy measures was correlated with patient characteristics similar to that in other studies, it does not add more information beyond individual sleep metrics alone and future research might benefit from selecting individual sleep metrics based on theory rather than use a composite measure approach. While our approach may have limited utility in BD, it may be important for research to examine this in other clinical groups, including those with other serious mental illnesses. As sleep becomes more frequently measured by actigraphy, efforts to improve the use and applicability of these unique data will be important for understanding the dynamics of sleep in those with BD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We thank Rebecca Daly for data management assistance, Sheena Dev, Jessica Carrasco, Sydney Willis, Laura Narayan Hein, Nana Kori, and Rudolph Benitez for their help with data collection and Michael Higgins for his assistance with sleep data aggregation.

This study was supported by funding from the National Institutes of Health (NIMH: R01MH103318, T32MH019934; K23MH119375; NIA: K01AG061239). Ms. Christensen was supported by the National Institute on Aging T35 grant AG26757 (PI: Dilip V. Jeste, MD), and the Stein Institute for Research on Aging and the Center for Healthy Aging at the University of California, San Diego.

REFERENCES

Ahrens A, Hansen CB, Schaffer ME, 2018. LASSOPACK: Stata module for lasso, square-root lasso, elastic net, ridge, adaptive lasso estimation and cross-validation.

Ahrens A, Hansen CB, Schaffer ME, 2019. lassopack: Model selection and prediction with regularized regression in Stata.

Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP, 2003. The role of actigraphy in the study of sleep and circadian rhythms. Sleep 26(3), 342–392. [PubMed: 12749557]

Ancoli-Israel S, Martin JL, Blackwell T, Buenaver L, Liu L, Meltzer LJ, Sadeh A, Spira AP, Taylor DJ, 2015. The SBSM Guide to Actigraphy Monitoring: Clinical and Research Applications. Behav Sleep Med 13 Suppl 1, S4–S38. [PubMed: 26273913]

Baillet M, Cosin C, Schweitzer P, Peres K, Catheline G, Swendsen J, Mayo W, 2016. Mood Influences the Concordance of Subjective and Objective Measures of Sleep Duration in Older Adults. Front Aging Neurosci 8, 181. [PubMed: 27507944]

Barbini B, Bertelli S, Colombo C, Smeraldi E, 1996. Sleep loss, a possible factor in augmenting manic episode. Psychiatry Res 65(2), 121–125. [PubMed: 9122286]

Bradley AJ, Anderson KN, Gallagher P, McAllister-Williams RH, 2020. The association between sleep and cognitive abnormalities in bipolar disorder. Psychol Med 50(1), 125–132. [PubMed: 30646966]

Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ, 1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 28(2), 193–213. [PubMed: 2748771]

Cohen S, Kamarck T, Mermelstein R, 1983. A global measure of perceived stress. J Health Soc Behav 24(4), 385–396. [PubMed: 6668417]

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

Cretu JB, Culver JL, Goffin KC, Shah S, Ketter TA, 2016. Sleep, residual mood symptoms, and time to relapse in recovered patients with bipolar disorder. J Affect Disord 190, 162–166. [PubMed: 26519636]

De Crescenzo F, Economou A, Sharpley AL, Gormez A, Quested DJ, 2017. Actigraphic features of bipolar disorder: A systematic review and meta-analysis. Sleep Med Rev 33, 58–69. [PubMed: 28185811]

Dolsen MR, Soehner AM, Harvey AG, 2018. Proinflammatory Cytokines, Mood, and Sleep in Interepisode Bipolar Disorder and Insomnia: A Pilot Study With Implications for Psychosocial Interventions. Psychosom Med 80(1), 87–94. [PubMed: 28914726]

Fava GA, Kellner R, 1991. Prodromal symptoms in affective disorders. Am J Psychiatry 148(7), 823–830. [PubMed: 2053620]

Full KM, Kerr J, Grandner MA, Malhotra A, Moran K, Godoble S, Natarajan L, Soler X, 2018. Validation of a physical activity accelerometer device worn on the hip and wrist against polysomnography. Sleep Health 4(2), 209–216. [PubMed: 29555136]

Gold AK, Sylvia LG, 2016. The role of sleep in bipolar disorder. Nat Sci Sleep 8, 207–214. [PubMed: 27418862]

Graham S, Depp C, Lee EE, Nebeker C, Tu X, Kim HC, Jeste DV, 2019. Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. Curr Psychiatry Rep 21(11), 116. [PubMed: 31701320]

Gruber J, Miklowitz DJ, Harvey AG, Frank E, Kupfer D, Thase ME, Sachs GS, Ketter TA, 2011. Sleep matters: sleep functioning and course of illness in bipolar disorder. J Affect Disord 134(1–3), 416–420. [PubMed: 21683450]

Hamilton M, 1960. A rating scale for depression. J Neurol Neurosurg Psychiatry 23, 56–62. [PubMed: 14399272]

Harvey AG, Talbot LS, Gershon A, 2009. Sleep Disturbance in Bipolar Disorder Across the Lifespan. Clin Psychol (New York) 16(2), 256–277. [PubMed: 22493520]

Irwin MR, Olmstead R, Carroll JE, 2016. Sleep Disturbance, Sleep Duration, and Inflammation: A Systematic Review and Meta-Analysis of Cohort Studies and Experimental Sleep Deprivation. Biol Psychiatry 80(1), 40–52. [PubMed: 26140821]

Jackson A, Cavanagh J, Scott J, 2003. A systematic review of manic and depressive prodromes. J Affect Disord 74(3), 209–217. [PubMed: 12738039]

Kaufmann CN, Gershon A, Eyler LT, Depp CA, 2016. Clinical significance of mobile health assessed sleep duration and variability in bipolar disorder. J Psychiatr Res 81, 152–159. [PubMed: 27451108]

Ketter TA, 2010. Strategies for monitoring outcomes in patients with bipolar disorder. Prim Care Companion J Clin Psychiatry 12(Suppl 1), 10–16.

Kroenke K, Spitzer RL, Williams JB, 2001. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 16(9), 606–613. [PubMed: 11556941]

Lo JC, Groeger JA, Cheng GH, Dijk DJ, Chee MW, 2016. Self-reported sleep duration and cognitive performance in older adults: a systematic review and meta-analysis. Sleep Med 17, 87–98. [PubMed: 26847980]

Medic G, Wille M, Hemels ME, 2017. Short- and long-term health consequences of sleep disruption. Nat Sci Sleep 9, 151–161. [PubMed: 28579842]

Natale V, Plazzi G, Martoni M, 2009. Actigraphy in the assessment of insomnia: a quantitative approach. Sleep 32(6), 767–771. [PubMed: 19544753]

Nuechterlein KH, Green MF, 2006. MATRICS Consensus Cognitive Battery Manual. MATRICS Assessment, Inc, Los Angeles, CA.

Overall JE, Gorham DR, 2016. The Brief Psychiatric Rating Scale. Psychological Reports 10(3), 799–812.

Perlman CA, Johnson SL, Mellman TA, 2006. The prospective impact of sleep duration on depression and mania. Bipolar Disord 8(3), 271–274. [PubMed: 16696829]

Rahe C, Czira ME, Teismann H, Berger K, 2015. Associations between poor sleep quality and different measures of obesity. Sleep Med 16(10), 1225–1228. [PubMed: 26429750]

Sadeh A, Sharkey KM, Carskadon MA, 1994. Activity-based sleep-wake identification: an empirical test of methodological issues. Sleep 17(3), 201–207. [PubMed: 7939118]

Sathyanarayana A, Joty S, Fernandez-Luque L, Ofli F, Srivastava J, Elmagarmid A, Arora T, Taheri S, 2016. Sleep Quality Prediction From Wearable Data Using Deep Learning. JMIR Mhealth Uhealth 4(4), e125. [PubMed: 27815231]

Simpson L, Hillman DR, Cooper MN, Ward KL, Hunter M, Cullen S, James A, Palmer LJ, Mukherjee S, Eastwood P, 2013. High prevalence of undiagnosed obstructive sleep apnoea in the general population and methods for screening for representative controls. Sleep Breath 17(3), 967–973. [PubMed: 23161476]

Sylvia LG, Chang WC, Kamali M, Tohen M, Kinrys G, Deckersbach T, Calabrese JR, Thase ME, Reilly-Harrington N, Bobo WV, Kocsis JH, McInnis MG, Bowden CL, Ketter TA, Friedman ES, Shelton RC, McElroy SL, Gao K, Rabideau DJ, Nierenberg AA, 2018. Sleep disturbance may impact treatment outcome in bipolar disorder: A preliminary investigation in the context of a large comparative effectiveness trial. J Affect Disord 225, 563–568. [PubMed: 28881294]

Sylvia LG, Dupuy JM, Ostacher MJ, Cowperthwait CM, Hay AC, Sachs GS, Nierenberg AA, Perlis RH, 2012. Sleep disturbance in euthymic bipolar patients. J Psychopharmacol 26(8), 1108–1112. [PubMed: 21965189]

Versace A, Almeida JR, Hassel S, Walsh ND, Novelli M, Klein CR, Kupfer DJ, Phillips ML, 2008. Elevated left and reduced right orbitomedial prefrontal fractional anisotropy in adults with bipolar disorder revealed by tract-based spatial statistics. Arch Gen Psychiatry 65(9), 1041–1052. [PubMed: 18762590]

Voets J, 2013. Study Shows that Over 252 Million Activity Trackers Will Ship Globally in the Coming Five Years. http://numrush.com/2013/05/17/study-shows-that-over-252-million-activity-trackers-will-ship-globally-in-the-coming-five-years/. (Accessed April 2, 2019.

Willetts M, Hollowell S, Aslett L, Holmes C, Doherty A, 2018. Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. Sci Rep 8(1), 7961. [PubMed: 29784928]

Wolz R, Munro J, Guerrero R, Hill DL, Dauvilliers Y, 2017. Predicting sleep/wake patterns from 3-axis accelerometry using deep learning. Alzheimer's & Dementia: The Journal of the Alzheimer's Association 13(7), P1012.

Young RC, Biggs JT, Ziegler VE, Meyer DA, 1978. A rating scale for mania: reliability, validity and sensitivity. Br J Psychiatry 133, 429–435. [PubMed: 728692]

Zhai L, Zhang H, Zhang D, 2015. Sleep Duration and Depression among Adults: A Meta-Analysis of Prospective Studies. Depress Anxiety 32(9), 664–670. [PubMed: 26047492]

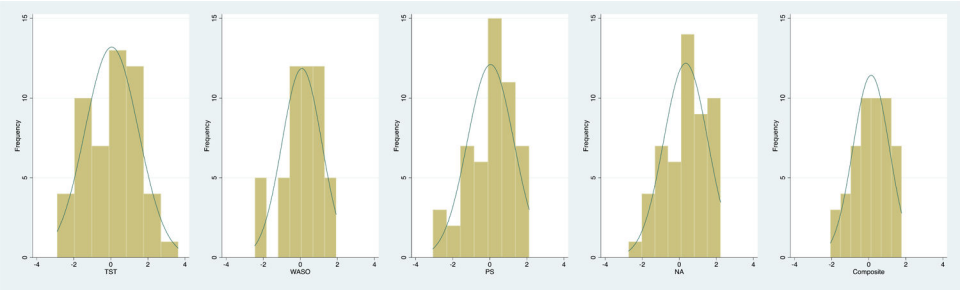

Figure 1:

Distribution of z-scores for both individual sleep statistics as well as composite in subjects with Bipolar Disorder (n=51)

Note: TST = total sleep time, WASO = wake after sleep onset, PS = percent sleep, NA = number of awakenings. WASO and NA are inverse coded.

Table 1:

Relationships between z-scores for composite sleep measure and demographic and clinical variables in subjects with Bipolar Disorder (n=51)

	N	Composite r or t-score
Demographics		
Age	50	r=−0.09
Education	50	r=0.01
Gender		
Female	31	$t_{(49)}=2.27^{*}$
Male	20	
Race/Ethnicity		
Caucasian	37	$t_{(49)}=-0.22$
Non-Caucasian	14	
Marital Status		
Single/Divorced/Separated/Widowed	36	$t_{(48)}=1.12$
Married/Cohabitating	14	
Employment		
Employed full- or part-time	15	$t_{(34)}=2.34^{*}$
Unemployed	21	
Clinical Characteristics		
PHQ-9 Depression Scale	38	r=−0.19
Global Cognitive T-Score	36	r=−0.10
Working Memory T-Score	37	r=0.04
Body Mass Index	51	r=−0.10
Perceived Stress Scale Total Score	38	r=−0.14
Use of Alcohol (mean % of surveys)	51	r=−0.22
Use of Cannabis (mean % of surveys)	51	r=−0.06
PSQI Component 6: Sleep Medication	50	r=−0.11
Medication Load	49	$r=0.41^{*}$
Age of Onset	41	r=0.21
Duration of Illness	41	r=−0.03
Young Mania Rating Scale	39	$r=-0.35^{*}$
Hamilton Depression Rating Scale	39	r=0.10
Brief Psychiatric Rating Scale	39	r=−0.22
Current Global Assessment of Functioning Score	43	r=−0.08
Inflammatory Biomarkers (log transformed)		
IL-6	51	r=−0.27
TNF-alpha	51	r=−0.18
CRP	51	r=−0.17

Notes:

* significant at $p<0.05$. Inflammatory biomarkers are log transformed. Pearson correlations are presented for continuous variables and t-scores for categorical variables. Positive t-scores indicate better sleep quality in the first category.