

## Original Research

**Cite this article:** Schneider J, Bakštein E, Kolenič M, Vostatek P, Correll CU, Novák D, and Španiel F (2022). Motor activity patterns can distinguish between interepisode bipolar disorder patients and healthy controls. *CNS Spectrums* 27(1), 82–92.

<https://doi.org/10.1017/S1092852920001777>

Received: 17 June 2020

Accepted: 04 August 2020

**Key words:**


Bipolar disorder; classification; sleep; circadian rhythm; actigraphy

**Author for correspondence:**

\*Jakub Schneider,

Email: [schnejak@fel.cvut.cz](mailto:schnejak@fel.cvut.cz)

# Motor activity patterns can distinguish between interepisode bipolar disorder patients and healthy controls

Jakub Schneider<sup>1,2\*</sup> , Eduard Bakštein<sup>1,2</sup>, Marian Kolenič<sup>2</sup>, Pavel Vostatek<sup>3</sup>, Christoph U. Correll<sup>4,5,6,7</sup>, Daniel Novák<sup>1</sup> and Filip Španiel<sup>2</sup>

<sup>1</sup>Department of Cybernetics, Czech Technical University in Prague, Prague, Czech Republic, <sup>2</sup>Applied Neuroscience and Neuroimaging, National Institute of Mental Health, Klecany, Czech Republic, <sup>3</sup>MINDPAX, Prague, Czech Republic, <sup>4</sup>Department of Psychiatry, The Zucker Hillside Hospital, Northwell Health, Glen Oaks, New York, USA, <sup>5</sup>Zucker School of Medicine at Hofstra/Northwell, Department of Psychiatry and Molecular Medicine, Hempstead, New York, USA, <sup>6</sup>The Feinstein Institute for Medical Research, Center for Psychiatric Neuroscience, Manhasset, New York, USA, and <sup>7</sup>Department of Child and Adolescent Psychiatry, Charité Universitätsmedizin Berlin, Berlin, Germany

**Abstract**

**Background.** Bipolar disorder (BD) is linked to circadian rhythm disruptions resulting in aberrant motor activity patterns. We aimed to explore whether motor activity alone, as assessed by longitudinal actigraphy, can be used to classify accurately BD patients and healthy controls (HCs) into their respective groups.

**Methods.** Ninety-day actigraphy records from 25 interepisode BD patients (ie, Montgomery–Asberg Depression Rating Scale (MADRS) and Young Mania Rating Scale (YMRS) < 15) and 25 sex- and age-matched HCs were used in order to identify latent actigraphic biomarkers capable of discriminating between BD patients and HCs. Mean values and time variations of a set of standard actigraphy features were analyzed and further validated using the random forest classifier.

**Results.** Using all actigraphy features, this method correctly assigned 88% (sensitivity = 85%, specificity = 91%) of BD patients and HCs to their respective group. The classification success may be confounded by differences in employment between BD patients and HCs. When motor activity features resistant to the employment status were used (the strongest feature being time variation of intradaily variability, Cohen's  $d = 1.33$ ), 79% of the subjects (sensitivity = 76%, specificity = 81%) were correctly classified.

**Conclusion.** A machine-learning actigraphy-based model was capable of distinguishing between interepisode BD patients and HCs solely on the basis of motor activity. The classification remained valid even when features influenced by employment status were omitted. The findings suggest that temporal variability of actigraphic parameters may provide discriminative power for differentiating between BD patients and HCs while being less affected by employment status.

**Introduction****Bipolar disorder**

The factors contributing to relapse in bipolar disorder (BD) are not yet clearly understood, but it has been suggested that there is an association with dysregulation of circadian rhythm<sup>1,2</sup> and sleep,<sup>3–7</sup> where BD has been reported to be associated with a later chronotype.<sup>8,9</sup> Circadian rhythm dysregulation appears both in acute episodes and in euthymic periods of BD. Therefore, measurements of circadian rhythm via motor activity profiles may provide a valid trait marker of BD,<sup>10</sup> and a deeper understanding of this dysregulation may contribute to improved management of the disease.<sup>11,12</sup>

**Actigraphy studies in BD patients**

Actigraphy is a convenient way to study motor activity patterns. Existing findings from actigraphy studies suggest that circadian rhythm and sleep are disrupted in patients with BD, even in the remitted state. Current evidence, including reviews (meta-analyses),<sup>1,4,12,13</sup> documents lower overall activity<sup>7,12,14–16</sup> and longer and more disrupted sleep in remitted BD patients than in healthy controls (HCs).<sup>5,6,17–19</sup> Similar observations have been found also in unaffected child and adolescent offspring of parents with BD.<sup>20</sup> Although previous studies have improved the understanding of motor activity in BD patients, most existing studies are based on a limited period of actigraphy monitoring and therefore miss the opportunity to assess and account for

intraindividual temporal variations in actigraphy parameters. Variability in sleep and in circadian parameters, obtained from actigraphy, suggests lower levels of synchronization of BD patients with the day and night rhythm<sup>5,13,16,17,21</sup> and may be closely connected with the symptomatic periods.<sup>12,22</sup> The short duration of the studies (mostly < 14 days, the longest being 50 days—see Supplementary Material) is a limitation for variability assessment.<sup>6,17,23</sup> In order to overcome these issues, we increased the observation period in the actigraphy study presented here to 90 days, aiming to focus on intraindividual long-term temporal variability (LTTV) in circadian rhythm and sleep parameters.

Contrary to statistical evaluation, the machine-learning techniques provide a means to quantify between-group differences by evaluation classification power of a set of features (biomarkers), considering complex nonlinear relationships and correction among features. There are at least two recent actigraphy studies employing this approach for actigraphy-based BD–HC classification. The first was done by Faedda et al.<sup>24</sup> who reached 83% accuracy with 64% sensitivity and 92% specificity, using 3 to 5 days of actigraphy and diary data from children (5–18 years old). There was no medication used, and all data were recorded during a similar regime (school days). The second recent study by Krane-Gartiser et al.<sup>25</sup> applied classification algorithms to a set of 61 HCs and 61 remitted BD patients, with stable medication, resulting in 78% accuracy (75% sensitivity and 80% specificity) using selected actigraphy features and MADRS scores, resulting in 70% accuracy using actigraphy alone. The main advantages were the use of matched groups (including employment status) and strict remission criteria (MADRS and YMRS  $\leq 8$  for  $\geq 3$  months).

### Literature-based differences between BD patients and HCs

Following the available literature, we expected lower overall motor activity<sup>7,14–16,26</sup> and also lower peak activity<sup>27</sup> in BD patients vs HCs. Based on diminished adaptability to changes in circadian rhythm, lower rhythm robustness was expected.<sup>28</sup> Additionally, due to greater mood instability, higher fragmentation of activity profiles within a day and instability between days was expected, including higher variability in most of the actigraphy parameters, both motor activity based or time based.<sup>9</sup>

Reduction in sleep quality has been reported in BD patients<sup>13,29</sup>; therefore, higher motor activity and longer awake or mobile periods were expected during night sleep. Further, since BD is associated with longer sleep<sup>1,4,6,18,30</sup> (though some reports did not confirm this finding<sup>7,14,31</sup>), we expected sleep time to be longer and more variable. Moreover, since longer sleep latency is associated with BD,<sup>5,6,17,18</sup> we expected lower activity before sleep onset and greater activity (restless sleep [RSL]) after sleep onset, with higher variability in both sleep latency and RSL. Finally, BD is associated with later chronotype,<sup>1,8,9</sup> represented as a later activity peak and a later sleep midtime.

### Variability measurements and primary objectives

This work is focused on motor activity and intraindividual temporal changes in motor activity during waking hours and during sleep. Motor activity was measured using a wrist-worn actigraphy device, an instrument specifically tailored for use in psychiatry (MINDPAX). Temporal variability is connected to changes in daily routine and in circadian rhythm synchronization. Temporal variability may, therefore, be a more straightforward way to measure the assumed triggers/predictors of BD symptoms<sup>1,10</sup> than a

standard comparison of average activity levels. We also expected the variability measurements to be comparatively insensitive to basic differences in daily routine between BD patients and HCs.

Aims of the study were: (a) to evaluate the motor activity profiles of interepisode BD patients vs HCs, (b) to use machine learning to distinguish between BD patients and HCs using actigraphy-derived features with a focus on variability measurements, and (c) to evaluate the effect of employment status on the results (post hoc).

## Data and Method

### Participants

Actigraphy data were recorded for more than 90 days in 35 BD patients mainly with BD type I diagnosis, recruited from the outpatient BD clinic at the National Institute of Mental Health (NIMH), in Klecany, Czech Republic, and in 26 HCs, matched for age and sex, who were recruited by advertisement in the community. All BD patients underwent a baseline psychiatric examination by a trained institutional psychiatrist, confirming euthymic state or low levels of depressive/manic symptoms, using the Montgomery–Asberg Depression Rating Scale (MADRS)<sup>32</sup> and the Young Mania Rating Scale (YMRS).<sup>33</sup> Inclusion criteria: all BD patients were diagnosed according to Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5) criteria.<sup>34</sup> At the study entry, all patients had to be euthymic or be in a remitted state<sup>35</sup> (ie, YMRS  $\leq 12$  and MADRS  $\leq 9$ , see Table 1) with no reported mood episodes for  $\geq 60$  days prior to study entry.

Exclusion criteria for BD patients were the presence of an acute depressive episode, dysthymia, suicidal thoughts, a (hypo)manic episode or diagnosis of schizoaffective disorder at enrolment.

HC exclusion criteria were: past or acute presence of a moderate depressive or (hypo)manic episode or suicidal thoughts, diagnosed neurological, sleep or mood disorders, or a first-degree family history of mood or psychotic disorder among their first-order relatives.

The study was approved by the ethical committee of the NIMH, Czech Republic, and all BD patients and HCs signed written informed consent.

### Procedure

On enrolment into the study, all participants answered a demographic questionnaire. The HC pool was contacted through an email screening questionnaire that asked for basic information (age, sex, and employment status) and family disease history (neurological: epilepsy, Parkinson's disease, etc., sleep disorders: insomnia, sleep apnea, narcolepsy, etc.). Subjects who fulfilled the screening criteria were further evaluated using the M.I.N.I. structured questionnaire<sup>36</sup> for neuropsychiatric disorders.

All participants were equipped with a wrist-worn actigraphic monitoring device (MINDPAX) and were instructed to remove it only when necessary.

During follow-up (the period when the data were recorded), BD participants were assessed monthly by their treating psychiatrist via in-person visits and/or a telephone interview to identify their current psychiatric state. We allowed for some minor increase of symptoms during follow-up (ie, YMRS < 15 and MADRS < 15). The criteria for clinical episodes included psychiatric hospitalizations, work incapacity, MADRS  $\geq 15$ , YMRS  $\geq 15$ ,<sup>37</sup> suicidal ideation, and/or substantial deterioration of the patient's clinical state.

**Table 1.** Demographic, Health and Activity Characteristics in Patients and Controls Groups

Metadata	BD	HCS	P value
Participants count final (original) [BDI/BDII]	25 (35) [16 BDI, 9 BDII]	25 (26)	–
Age	39.72 (SD 12.85, range 22-63)	39.68 (SD 11.19, range 25-63)	$P = .6549^a$
Sex	60% female (N = 15)	68% female (N = 17)	$P = .7688^b$
Days in study (recorded/valid)	134 (SD 39) [range 61-179]/ 86 (SD 21) [range 50-124]	97 (SD 12) [range 72-126]/86 (SD 13) [range 62-108]	(for recorded) $P = .0018^{c*}$ (for valid) $P = .2977^d$
Working status (full-time/part-time/no work)	6/12/7	22/3/0	$P = .0125^{c*}$
Total days in the study (recorded/valid)	3341/2158	2426/2113	–
MADRS	At admission to the study 2.3 (SD 3.9) Through the study—with relapses 8.8 (SD 7.1) —without relapses 6.6 (SD 3.8)	–	–
YMRS	At admission to the study 0.4 (SD 0.81) Through the study—with relapses 1.8 (SD 3.3) —without relapses 2.1 (SD 3.5)	–	–
Mood episodes	In study: two patients with episode/s (2 dep., 1 hypomania) Excluded: 5 patients with episode/s (7 dep., 1 mania - psychosis)	–	–
<b>Lithium</b>	<b>7 (28%)</b>	–	–
<b>Antipsychotics used:</b>	<b>12 (48%)</b>	–	–
Quetiapine	10 (40%)	–	–
Olanzapine	4 (16%)	–	–
<b>Antidepressants used:</b>	<b>6 (24%)</b>	–	–
Bupropion	5 (20%)	–	–
Sertraline	2 (8%)	–	–
<b>Antiepileptic used:</b>	<b>9 (36%)</b>	–	–
Lamotrigine	6 (24%)	–	–
Valproate	3 (12%)	–	–
Carbamazepine	1 (4%)	–	–
<b>Mixed/single treatment</b>	19 patients/6 patients	–	–
<b>Average features<sup>d</sup></b>	605 (SD 110)	778 (SD 92)	
ADA (average daily activity)	8.98 (SD 1.22) h	7.40 (SD 0.51) h	
Sleep duration	0.78 (SD 0.12)	0.66 (0.07)	
Circadian quotient	14.7 (SD 1.3) o'clock	14.9 (SD 1.3) o'clock	For statistical comparison see Table 2
M10 time	2.6 (SD 0.9) %	2.1 (SD 0.6) %	
RSL (restless sleep)	3.7 (SD 1.0) o'clock	3.6 (SD 1.1) o'clock	
MSFcS (chronotype)			
<b>LTTV in features<sup>d</sup></b>	103 (SD 32)	94 (SD 25)	
ADA	1.7 (SD 0.6) h	1.3 (SD 0.3) h	
Sleep duration	0.07 (SD 0.02)	0.05 (SD 0.01)	
IV (intradaily variability)	0.06 (SD 0.02)	0.06 (SD 0.02)	
IS (interdaily stability)	166 (SD 50)	148 (SD 42)	For statistical comparison see Table 2
M10	2.1 (SD 0.8) h	2.6 (SD 0.5) h	
M10 time	31 (SD 23)	38 (SD 23)	
L5	1.8 (SD 0.6) h	1.4 (SD 0.5) h	
L5 time	1.8 (SD 0.8)%	1.5 (SD 1.0)%	
RSL			

Abbreviations: ADA, average daily activity; BD, bipolar disorder; IS, interdaily stability; IV, intradaily variability; LTTV, long-term temporal variability; MADRS, Montgomery–Asberg Depression Rating Scale (MADRS); RSL, restless sleep; SD, standard deviation; YMRS, Young Mania Rating Scale.

<sup>a</sup>Mann–Whitney test.

<sup>b</sup>Fisher exact test.

<sup>c</sup> $\chi^2$  test (Chi-square = 8.77).

<sup>d</sup>Selected features, for all see Supplementary Material.

\* $P < .05$ .

\*\* $P < .01$ .

\*\*\* $P < .001$ .

The relapse episodes were omitted from the data, and, thus, we used only patient data from interepisode periods, therefore, the results were not affected by motor activity changes due to clinical episodes. The rating scales were not used in any of the models. In order to reduce possible seasonal effects, the data were collected in both groups during a similar time period from December 2016 to May 2017.

### Actigraphy data and feature extraction

The actigraphy wearable was set to capture motor activity aggregated over 30-second epochs. The data were transferred wirelessly through a base station at the subject's home to a server for storage and offline evaluation. Missing data, detected wearable lay-off periods, and segments starting 1 week before and ending 1 week after an episode of mania or depression were excluded from the trait-focused analysis. Data from mood episodes (mania/depression) were excluded in order to minimize contamination of the data, which might lead to increased differences between BD patients and HCs. For feature estimation, 80% of valid samples within a given time frame were required; otherwise, they were marked as a missing value. All calculations were performed in local time, using weekends and national holidays as free days.

In order to perform the statistical and machine-learning analyses, we derived a set of standard features from the raw actigraphy recordings, based on cosinor, nonparametric and sleep analysis. All actigraphy features were calculated for each day, while a 7-day sliding time window was used for some features (see below). Chronotype and social jetlag (SJL) were estimated using the whole actigraphy recording.<sup>38</sup>

### Circadian analysis

Cosinor analysis<sup>28,39</sup> estimated regular activity patterns by fitting a cosine function with a fixed 24-hour period. The resulting features are the acrophase (the time of the activity peak), the mesor (the offset of the fitted cosine function), the amplitude, the circadian quotient (CQ) (the amplitude and mesor ratio, that is, the robustness of the daily activity rhythm), and the goodness of fit (GOF) (the sum of absolute residuals from fitted cosine function).

Nonparametric analysis<sup>14,27,40</sup> estimated activity patterns for each day, without assuming an underlying analytical function. The estimated features are M10 (average activity during the most active 10 hours), L5 (average activity during the least active 5 hours), timings of M10 and L5 (midwindow daytime), and RA (relative amplitude:  $(M10 - L5)/(M10 + L5)$ ). Additional features estimated for each day include ADA (average daily activity) and AQA1–4 (average activity in quarters of a day).

The nonparametric features, based on the 7-day sliding window, were IV (intradaily variability, describing rhythm fragmentation) and IS (interdaily stability). For an estimate of IV and IS, the signal was aggregated into 20-minute segments, according to previously reported more favorable properties.<sup>27</sup>

### Sleep analysis

Sleep epochs were detected using the MINDPAX algorithm.<sup>41</sup> To reduce overestimation of sleep duration (SDur) based on algorithmic analysis, as observed by Boudebesse et al.,<sup>42</sup> the results were manually checked and corrected. The distinction between sleep and wake time using actigraphy is generally considered reliable.<sup>31,43</sup> The main SDur, that is, the longest SDur on a given day, with minor interruptions, was used for the analyses.

In order to cope with issues connected with sleep parameters requiring a sleep diary (eg, efficiency, latency), we decided to use actigraphy features only. The extracted features were SDur, sleep onset, sleep offset, midsleep (MS), RSL (the percentage of minutes within the main sleep, above preset inactivity threshold), and immobile parts of sleep (ISL: the percentage of minutes within the main sleep, below preset activity threshold), sleep instability (RMSSD: Root mean square of successive differences—based on raw actigraphy during detected sleep), and average activity 2 hours prior to sleep onset (APSO), and 2 hours after sleep onset (AASO).

### Chronotype and social jetlag

Circadian chronotype and SJL are commonly evaluated by the Munich Chronotype Questionnaire (MCTQ),<sup>38</sup> estimated as the mid-sleep on free days corrected for sleep debt on work days (MSFsc). In the analysis presented here, we estimated both MSFscs and SJL using actigraphy-derived sleep starts and sleep ends<sup>44</sup> and calendar free days, using the whole follow-up period. The non-parametric M10 and L5 times were used for the LTTV evaluation.

### Statistical analysis

The LTTV and average values were calculated from all available daily values, using the standard deviation and the mean, respectively. Between-group statistical comparison was performed on an a priori selected subset of features (Table 2), based on the available literature.

The features were checked for normality using Q–Q plots and were normalized on the basis of skewness and kurtosis (for details see Supplementary Material). When normality was not disproved in the transformed values (Jarque–Bera test,  $\alpha = 5\%$ ), a student *t* test was used; otherwise, the Wilcoxon rank-sum test was used for non-normally distributed data. One-sided tests were used, based on an a priori hypothesis from the existing literature (“Literature-based differences between BD patients and HC” section). See the supplement for details on existing studies and feature normality. See the data processing scheme in Figure 1.

The results were corrected for multiple comparisons using the Holm<sup>45</sup> procedure ( $n = 25$ ). The corrected results are marked as “corr” after each result in the “Results” section. The effect size was calculated as the standardized mean difference (SMD). The area under the receiver operating characteristic (AUC) was computed to measure the classification power of individual actigraphy features.

All data processing and statistical analyses were conducted with Matlab 2015b, The MathWorks, Inc.

### Classification

In order to illustrate the discriminatory power of the entire feature set combined (as opposed to the statistical analysis, which was aimed at individual features/biomarkers), we designed a set of classifiers for discriminating between BD patients and HCs. In total, we used three models differing by the features that were employed: (a) a model with all the features presented above, (b) a model based only on temporal variabilities, and (c) a model using only features with low dependency on employment status (see details below).

The models were trained using a random forest (RF) classifier,<sup>46</sup> commonly used for heterogeneous biomedical data including actigraphy,<sup>24</sup> and the out-of-sample performance was estimated using fivefold cross-validation. In each fold, data from 20 BD patients and 20 HC participants were used for training the

**Table 2.** Group Differences Between Patients and Controls

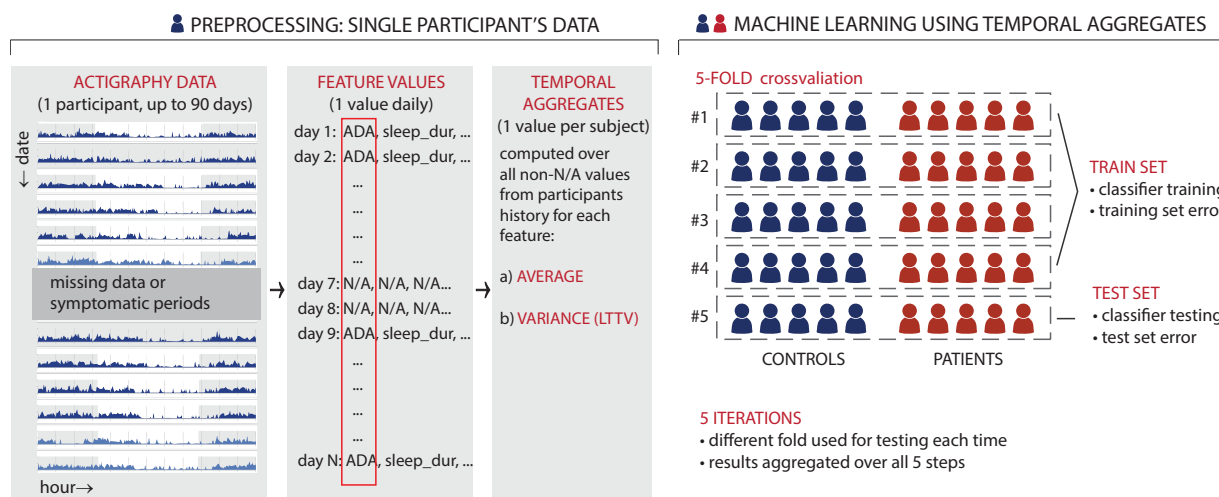
Temporal Variability				Average Values				Rationale
Hypothesis—LTTV in Feature is Higher/Lower in BD Patients	<i>P</i> value	AUC	SMD (nonparam)	Hypothesis—Average Value in Feature is Higher/Lower in BD Patients	<i>P</i> value	AUC	SMD (nonparam)	
Var. in IV is higher in BD patients	<b>&lt;.001***</b>	0.8544	1.33 (0.97)	IV is higher in BD patients	.131	0.5872	0.32 (0.26)	Fragmentation of activity within a 24-h cycle. <sup>1</sup>
Var. of M10 time is higher in BD patients <sup>†</sup>	<b>&lt;.001*</b>	0.7680	−0.74 (−0.71)	M10 time is later in BD patients	.707	0.5168	−0.15 (−0.24)	Finding daily activity extremes on a daily basis. <sup>1</sup>
Var. of L5 time is higher in BD patients	<b>.003*</b>	0.7456	0.71 (0.65)	L5 time is later in BD patients <sup>†</sup>	.197	0.5712	0.04 (0.35)	See M10 time
Var. in sleep duration is higher in BD patients	<b>.004</b>	0.7168	0.79 (0.56)	Sleep duration is higher in BD patients <sup>†</sup>	<b>&lt;.001***</b>	0.8592	1.68 (2.03)	A basic sleep-describing feature.
Var. in AASO is higher in BD patients	<b>.039</b>	0.6448	0.47 (0.40)	AASO is higher in BD patients	<b>&lt;.001**</b>	0.76	0.99 (0.88)	Used as an approximation of sleep latency.
Var. of CQ is higher in BD patients	<b>.041</b>	0.6352	0.5 (0.27)	CQ is higher in BD patients <sup>†</sup>	<b>&lt;.001***</b>	0.8512	1.23 (1.23)	An estimate of how well-circumscribed periods of activity are during a day; a proxy for rhythm robustness <sup>23</sup>
Var. of M10 is higher in BD patients	.074	0.6112	0.41 (0.23)	M10 is lower in BD patients	<b>&lt;.001**</b>	0.792	−1.16 (−0.53)	Approximates the amplitude of peak daily activity (M10) and sleep quality (L5) for each day. It is related to motor capability <sup>21</sup>
Var. in RSL is higher in BD patients	.081	0.6304	0.26 (0.41)	RSL is higher in BD patients	<b>.012</b>	0.6624	0.67 (0.31)	Feature describing sleep quality. Represents sleep inefficiency based on actigraphy
Var. in IS is higher in BD	.110	0.6272	0.29 (0.31)	IS is lower in BD	.417	0.5328	−0.11 (−0.14)	Synchronization to the light–dark cycle and stability of daily rhythm.
Var. in ADA is higher in BD	.151	0.5904	0.31 (0.34)	ADA is lower in BD	<b>&lt;.001***</b>	0.8992	−1.71 (−1.05)	Describes how active a person is throughout the day.
Var. in APSO is lower in BD	.160	0.5808	−0.27 (−0.24)	APSO is lower in BD	<b>&lt;.001**</b>	0.7712	−1.14 (−0.45)	See AASO
Var. in L5 is higher in BD	.952	0.3424	−0.32 (−0.25)	L5 is higher in BD	.932	0.3696	−0.38 (−0.38)	See M10
				BD patients are generally a later chronotype	.249	0.5568	0.08 (0.15)	Objectively assess the values from MCTQ <sup>36</sup> questionnaire.

Significance after Holm's<sup>45</sup> ( $n = 25$ ) \* $P < .05$ , \*\* $P < .01$ , \*\*\* $P < .001$ .

Abbreviations: AASO, activity 2 hours after to sleep onset; ADA, average daily activity; APSO, activity 2 hours prior to sleep onset; AUC, area under the curve; BD, bipolar disorder; CQ, circadian quotient; IS, interdaily stability; IV, intradaily variability; LTTV, long-term temporal variability; MCTQ, Munich Chronotype Questionnaire; RSL, restless sleep; SD, standard deviation; SMD, standardized mean difference.

<sup>†</sup>tested using Wilcoxon rank-sum test (non-normally distributed data).





**Figure 1.** Preprocessing and machine-learning classification scheme. The left side shows the estimation of individual values (average and long-term temporal variability [LTTV]) from features, based on all valid days for each patient. In the right part, the machine-learning cross-validation process is shown, where the same number of bipolar disorder (BD) patients and healthy controls (HCs) is repetitively left out from the learning dataset for results.

classifier, the rest were used for evaluating the classification performance. In subsequent folds, the data from 5 + 5 different subjects were used for validation, until all patients were iterated. The entire fivefold procedure was repeated 100 times to estimate the uncertainty of the results, caused by the random division of the patients into folds and random feature selection in RF. See the supplement for more details on RF structure and evaluation. See the data processing scheme in Figure 1.

### Post hoc analysis of employment status

An analysis of the classification results revealed a strong association between the misclassification of individual subjects and employment status. We, therefore, investigated the association between employment status, group membership (BD patients or HCs), and individual actigraphy feature values. A set of linear models was built that the parameter value was a linear combination of BD patients/HC group status, employment status, and intercept:

$$\text{feature} \sim 1 + \text{BD/HC} + \text{employment status}. \quad (1)$$

The model was fitted using a least-square means approach with robust bi-square weights, and the significance of the coefficient values was evaluated using standard T-statistic. Based on the results, the identified features independent of employment status were used for training classification model C.

## Results

### Subject characteristics

The BD patients and HC group characteristics, after exclusions, are shown in Table 1. Ten (29%) of the 35 BD patients enrolled in the study were excluded: five subjects were excluded due to insufficient length of recorded interepisode data ( $\leq 50$  days—four for depression episodes and one for psychosis after childbirth), four for an excessive amount of missing actigraphy data (due to wearable removal or malfunction of the wearable device), and one resigned from the study upon personal request, resulting in 25 DB patients

in the final set. All of the subjects were attending a standard BD treatment program and were using clinician's choice medication. Among 26 HCs, one subject was excluded due to an excessive amount of missing data, resulting in 25 HCs in the final set. The HC set was recruited after the BD set, when deterioration in the same patient's state was observed. A lower dropout rate in HCs was expected.

### Statistical comparison

In terms of LTTV, compared with HCs, BD patients showed significantly greater variation in the IV feature ( $t(48) = -4.71$ ,  $P_{\text{corr}} = .0005$ ,  $\text{AUC} = 0.85$ ), greater variability in the activity-peak-time (M10-time;  $z = 3.24$ ,  $P_{\text{corr}} = .0107$ ,  $\text{AUC} = 0.77$ ), and greater variability in the L5 time ( $t(48) = -2.88$ ,  $P_{\text{corr}} = .0500$ ,  $\text{AUC} = 0.75$ ). In the IS feature, the variability had higher predictive capacity than the mean value (both nonsignificant). For actual differences, see Table 1, and for effect sizes see, Table 2.

When evaluating individual averages (Tables 1 and 2), compared to HCs, BD was associated with lower ADA ( $t(48) = 6.06$ ,  $P_{\text{corr}} < .0001$ ,  $\text{AUC} = 0.90$ ), longer sleep duration ( $z = -4.35$ ,  $P_{\text{corr}} = .0002$ ,  $\text{AUC} = 0.86$ ), and lower CQ ( $z = -4.25$ ,  $P_{\text{corr}} = .0002$ ,  $\text{AUC} = 0.85$ ). However, in some features, mainly the overall averages, the observed differences were highly associated with BD employment status. For more details and posthoc analysis on the effect of employment status, see "Effect of employment status" section.

### Classification of BD and HCs

The full actigraphy-based model (model A) was successful in distinguishing between people with interepisode BD and HCs. Accuracy was around 88% with specificity 91%—see Table 3.

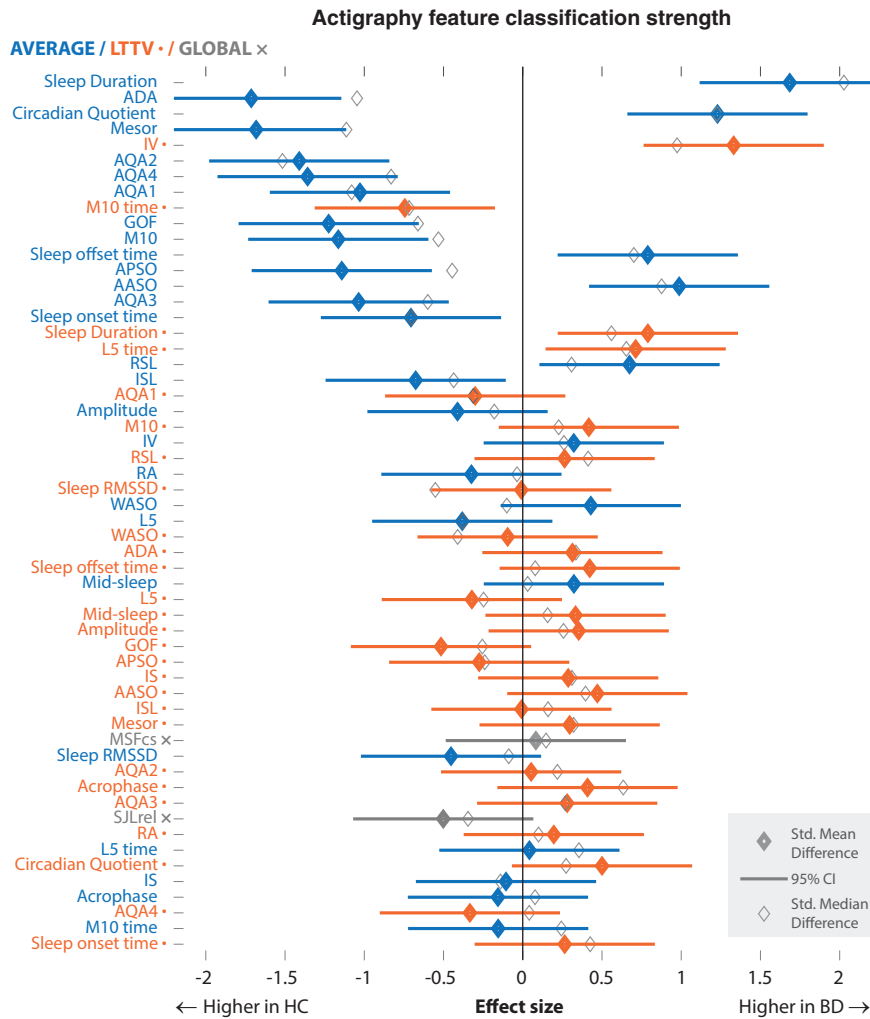
When only time variability of the actigraphy features were used (model B), the classification accuracy dropped, mainly due to a higher HC misclassification rate (ie, a drop in specificity). The accuracy drop in the B model was apparently also due to the removal of the strongest feature, which was the average SDur

**Table 3.** Random Forest Classifier Model Results in Participants Whose Data Were Not Used During Model Training

Model	Accuracy Mean (SD)	Sensitivity Mean (SD)	Specificity Mean (SD)	Misclassification in BD Patients Based on Employment <sup>a</sup> —Full-Time/Part-Time/Pensioned
A. All features	87.8 (2.6)%	84.8 (3.5)%	91.0 (4.0)%	29%/17.2%/0%
B. Time variations	78.5 (4.2)%	77.7 (5.2)%	79.3 (5.8)%	36%/26.7%/3%
C. Employment status independent features	78.7 (3.4)%	76.2 (5.3)%	81.2 (4.2)%	33%/29%/7%

Abbreviations: SD, standard deviation; BD, bipolar disorder.

<sup>a</sup>The number of bipolar disorder (BD) patients working full-time was 6 (therefore, 1 patient corresponds to 16.7%) part-time working n = 12 (1 patient ~ 8.3%) and unemployed n = 7 (1 patient ~ 14.3%).



**Figure 2.** All features used in classification the (Model A) ordered by their classification strength, showing the effect size for each feature (non-normalized data), LTTVs in blue, average values in orange, and global features in gray. The effect size (with a 95% confidence interval) is an approximation because the distribution was not always normal. The grey diamond shows the effect size estimated by the median values and shows how precise the blue approximation is.

(BD  $8.97 \pm 1.22$  hours vs HC  $7.40 \pm 0.51$ —this corresponds to adult sleep duration).<sup>47</sup> Most of the misclassifications were in full-time-/part-time-working BD patients (see the last column in Table 3). For model A, in the working patients, 1.7 out of 6 were on average misclassified; in the part-time-working patients, 2.1 out of 12 were misclassified, while there were no misclassifications in the unemployed/pensioned patients. For model B, 0.2 out of 7 unemployed/pensioned patients were misclassified. (For model C, which

uses features that do not show dependency on employment status, see “Effect of employment status” section.)

Based on the out-of-bag estimation (see Supplementary Material), we assessed the importance of each feature in the classification task. Figure 2 shows features ordered by their average classification strength, depicting their approximate effect sizes based on model A. Models B and C differ by not including the unused features (the order of classification strength does not change).

**Table 4.** Categories of Features Based on Employment Status Effects

Category	Category Description	Time-Variability of the Feature	Average of the Feature
1	Features affected exclusively by BD/HC difference (used for model C).	<b>M10 time, IV, SDur</b>	<b>M10, APSO, AASO</b>
2	Features affected exclusively by employment status.	L5 time	L5 time
3	Features affected by both BD group and employment status. The $\uparrow\uparrow$ reduces the BD effect (are in the same direction and combine into a bigger difference) and the $\uparrow\downarrow$ feature supports the BD effect due to employment status (is in the opposite direction and the difference is greater).		$\uparrow\uparrow$ CQ, $\uparrow\uparrow$ ADA, $\uparrow\uparrow$ SDur <sup>a</sup> , $\uparrow\downarrow$ RSL
4	Features not significantly affected either by BD or by working status.	CQ, M10, IS, ADA, L5, APSO, AASO, RSL	L5, M10 time, IS, IV, MSFcs

Abbreviations: AASO, activity 2 hours after sleep onset; ADA, average daily activity; APSO, activity 2 hours prior to sleep onset; BD, bipolar disorder; CQ, circadian quotient; HC, healthy control; IS, interdaily Sstability; IV, intradaily Vvariability; RSL, restless sleep; SDur, sleep duration.

<sup>a</sup>For sleep duration, the effect of the disease is twice as strong, resulting in the finding that even working BD patients differed significantly from HCs.

### Effect of employment status

Using linear models, we identified four types of features based on their association with employment status (see Table 4). We trained a new random forest distinguishing BD patients and HCs, using exclusively the variables that were most affected by the BD patients/HC group difference and not by employment status. Model C, which used only type 1 features LTTV in M10 time, IV, and SDur and averages of M10, and activities prior and after sleep onset reached 78.7% accuracy, for details see Table 3—model C.

### Discussion

This study shows that a machine-learning model using only actigraphic recording was capable of distinguishing between interepisode BD patients and HCs with 88% accuracy on the test data. In addition, when the effect of working status was suppressed by empirically derived feature selection, our results indicated that actigraphic data on motor activity patterns in BD may contain a clinically informative and scalable biosignal that differentiates between BD patients and HCs. In an article, Ortiz et al.<sup>48</sup> used machine learning for forecasting a clinical episode based on patient-perceived energy during the evening. The motor activity is associated with future mood and energy<sup>11</sup>; therefore, long-term actigraphy may be promising for relapse forecasting.

When compared to existing actigraphy-based machine-learning studies Krane-Gartiser et al.<sup>25</sup> and Faedda et al.,<sup>24</sup> our model using all features is more accurate than both. When only features with low dependency on employment status are used, the results are slightly better than the results by Krane-Gartiser (acc. 79% vs 78%) and lower than Faedda's (acc. 83%) with higher sensitivity (76% vs 64%) and lower specificity (81% vs 92%). These results have to be considered bearing in mind that our remission criteria were more lenient than in the Krane-Gartiser paper, whose dataset was matched for employment status. Faedda used children with a similar daily regimen (school) and without any medication treatment but also cleaned noisy data based on additional information obtained from parents. In addition to actigraphy, the Krane-Gartiser<sup>25</sup> employed MADRS as an additional predictor variable as well. Post hoc analyses demonstrated that the inclusion of this psychopathology score contributed critically to the overall efficacy of the model. When MADRS was excluded, the accuracy-based selectively on motor activity dropped to 70%,<sup>25</sup> which is lower than our results.

Matching the BD patients and HC groups based on employment status, Krane-Gartiser et al.<sup>25</sup> reduced the confounding effect of

differences in social engagement. This approach has been substantiated by the identification of employment status as a significant confounder. The fact that HCs are typically employed and, at the same time, many BD patients are either unemployed or pensioned, may itself introduce a significant bias, due to the systematic effect of the dissimilar social clock and demands in the two groups. To address this problem at least partially, some studies have used shift work as an exclusion criterion.<sup>6,7,49</sup> Only a small number of actigraphy studies have attempted to match HCs on employment status.<sup>14,17,25,50</sup> Unfortunately, even using age-matched HC groups with a similar rate of unemployment may introduce a different type of bias,<sup>6</sup> due to the reasons causing a healthy person of productive age to be unemployed.

To control specifically for these potential biases, we identified and modeled a set of actigraphic features with low dependency on employment status and possibly also other aspects affecting motor activity during the day, such as family status and type of employment. The contribution of these different factors to the BD-specific characteristics of motor activity patterns is beyond the presented dataset and has to be evaluated in a separate study. According to our analysis, LTTV in interdaily variability feature, LTTV in M10 time, LTTV in SDur and average M10, and average activity before sleep onset and after sleep onset fulfil these requirements. In a post hoc analysis, the model incorporating exclusively features with low dependency on employment status achieved predictive accuracy of 79% in discriminating between BD patients and HCs.

### Long-term temporal variability

Recent evidence suggests that not only previously reported changes in the sleep and activity of BD patients but also, in particular, the temporal variability of these parameters may be a disease-specific trait marker.<sup>51</sup> Despite this promising report, only a few studies have been able to specifically address time variation features in actigraphy. The reasons for this situation are mainly technical, as the analysis typically requires long-term continuous actigraphy. These limited number of studies found for BD: (a) increased standard deviation and RMSSD in actigraphy,<sup>22</sup> (b) a positive correlation between mood variability and variability in activity,<sup>52</sup> and (c) greater variability in afternoon activity (BDI) and in nighttime activity (BDII), without differences in peak time variation,<sup>51</sup> but (d) greater variability in peak activity time.<sup>9</sup> For a review of variability in actigraphy, see Bei.<sup>21</sup>

Consistently, our analysis of the long-term time variability of actigraphy and sleep features revealed a significantly higher variability in the IV feature and in day-peak and day-trough activity



times (M10 time and L5 time) in BD patients vs HCs. In sleep features, we observed a difference in SDur time variability (although not significant after correcting for multiple comparisons  $P < .1$ , see the limitation of the power analysis in the Supplementary Material).

### Average actigraphy and sleep

As in previous studies, a lower overall activity (ADA) and flattening in rhythmicity (CQ) were detected in BD patients vs HCs. Lower activity is a widely reported trait-marker of BD, even in remitted cases.<sup>7,14–16,19,26,49,53</sup> Unfortunately, ADA also showed significant dependency on employment status. A lower daily activity peak was observed by prior studies,<sup>19,53</sup> where it was connected with worsening of the disease.

In contrast to previous studies,<sup>1,9,13</sup> we did not observe a between-group difference in chronotype based on motor activity, although all subjects were evaluated at approximately the same time of year. Similarly, we did not observe any later activity onset in BD patients vs HCs, as had been observed previously.<sup>9,15,50,51,53</sup>

Prolonged SDur (in our study for >1 hour) has been observed by some<sup>5,6,18</sup> but not other studies.<sup>7,14,17</sup> It is possible that the observed difference may be caused (a) by persistent subdepressive symptoms because even between-episodes BD patients show more depression-related symptoms,<sup>5,54</sup> (b) by medication, whereby especially atypical antipsychotics are related to hypersomnia,<sup>30</sup> and/or (c) by the difference in employment status, as already has been mentioned.

Other commonly observed differences in BD are lower sleep efficiency<sup>6,16</sup> and prolonged sleep latency.<sup>5,6,17,18</sup> These values cannot be estimated without the use of sleep diaries or patient markings of sleep time, which were not collected in our study. Our fully automatic approximation of these features are RSL, for sleep efficiency, and decline in activity on sleep onset, measured by APSO and AASO, for sleep latency. The between-group difference in RSL was not significant after corrections for multiple comparisons. Further, a slower decline in activity was observed in BD vs HCs during sleep onset was observed (APSO was lower and AASO was higher in BD patients vs HCs).

### Limitations

Results of this study need to be interpreted considering the following limitations. First, the relatively small sample size can reduce the power of the statistical tests (see the Supplementary Material for a power analysis). Although the sample size was small, it is in line with many previous actigraphy studies, each of which had a much shorter follow-up duration than our 90-day period.

Second, we had a relatively high dropout/exclusion rate of about 29% in BD patients, due to loss of interest in participating in the study, occurrence of a relapse, or due to technical difficulties. However, a comparable dropout rate is not exceptional in this type of study. For example, Krane-Gartiser et al.<sup>25</sup> had a dropout rate of 54% in the BD patients group, as a consequence of very strict exclusion criteria.

Third, BD patients and HCs were not matched for employment status. There are reasons that might cause unemployment in “healthy” productive age individuals. However, contaminating the HC group with the risk of morbidity might add a different type of bias as the oversampling employed BD patients. To address this issue, we did not select the sample on the basis of employment status. Instead, we conducted a sensitivity analysis, creating a

model, whereby actigraphy features that were highly correlated with employment status were removed, showing the robustness of our discrimination/prediction models.

Fourth, all BD patients should have been using their prescribed medication. There are reported effects of medication on sleep,<sup>55</sup> and effects on activity can also be expected, although Jones et al.<sup>14</sup> stated that “no evidence was found for a significant association between medication use and any of the circadian activity measures” and Shou et al.<sup>51</sup> did not observe any association between psychotropic medication and levels of activity. It has been shown that mood stabilizers can affect several circadian parameters.<sup>56</sup> The assumed major mechanism is through the regularization (normalization) of the sleep and circadian rhythm as it has been shown for lithium (seven patients in our study) and valproate (three patients in our study).<sup>5</sup> Considering the combination of medications, Gonzalez et al.<sup>28</sup> observed that individual medication type (mood stabilizers, antidepressants, antipsychotics, etc.) had higher association with motor activity changes than the number of medications from each type. The medications may nonetheless impact the results and therefore present a limitation of the study. However, withdrawal from medication during the follow-up period is unjustified due to the risk of relapse and related ethical issues.

Fifth, the BD subjects were not fully euthymic, and residual symptoms may have affected the results. Our relapse threshold allowed the presence of subclinical symptoms in the examined sample, for example, residual depression.<sup>54</sup> Monthly clinical assessments may also miss, or may underestimate, briefer but clinically relevant mood shifts.

Sixth, there are findings of high prevalence of comorbidities in BD.<sup>57</sup> Although many are hard to distinguish from symptoms of BP itself (sleep disorders, anxiety disorders, borderline personality disorder), there are other diseases that have higher prevalence in the BD group such as drug/alcohol abuse, asthma, hypothyroidism, migraine, etc., which may also affect circadian rhythm and motor activities throughout the day. These were not matched with the HC group and thus present a possible confounder and a limitation of the study.

Finally, we did not include patients with psychiatric disorders other than BD in order to evaluate the degree to which the identified actigraphic biosignature is specific to BD or is more globally a marker of mental illness. Thus, future studies should include psychiatric control groups to investigate this issue.

### Conclusions

There are significant differences in activity patterns between BD patients and HCs. A clinically applicable, cost-effective, and scalable classifier-based approach was able to distinguish BD patients from HCs with approximately 88% accuracy, which is better than previous studies by a large margin. Some of the strongest discriminants, for example, ADA and SDur, could be closely associated with differences in employment status and also with differences in the use of medications. The time variance in some features (intra-daily variability, peak activity time, SDur) showed lower dependency on employment status and may therefore be a preferable actigraphy biomarker candidate. When only features that are less dependent on employment status were used, the model was still able to distinguish BD patients from HCs with approximately 79% accuracy, which is still comparable with the best results obtained by other groups.<sup>24,25</sup> Future studies are needed in order to identify actigraphy features that are global trait-markers of mental illness

from those that are more specific to BD and, eventually, to identify features (state-markers) that may be associated with an impending relapse.

**Financial support.** The research presented here was supported by the Ministry of Health of the Czech Republic (grant number IGA NT 14387). The work at NIMH was supported by the Ministry of Education Youth and Sports of the Czech Republic (project number LO1611—NPU I program). The work of JS has been supported by the student grant agency of the Czech Technical University in Prague (grant number SGS19/171/OHK3/3 T/13). The sponsors of the study had no role in the design or conduct of this study; in the collection, management, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript.

**Disclosures.** The authors, Ing. Jakub Schneider, Ing. Eduard Bakštein, PhD., Ing. Pavel Vostatek, PhD., and doc. Ing. Daniel Novák, PhD. are associated, as consultants, advisors and/or data analysts, with MINDPAX, the company that focuses on management of bipolar and schizophrenia disorder and mental health data analyses. The companies that Dr. Correll has been a consultant and/or advisor to, or has received honoraria from: Alkermes, Allergan, Angelini, Boehringer-Ingelheim, Gedeon Richter, Gerson Lehrman Group, Indivior, IntraCellular Therapies, Janssen/J&J, LB Pharma, Lundbeck, MedAvante-ProPhase, Medscape, Merck, Neurocrine, Noven, Otsuka, Pfizer, Recordati, Rovi, Servier, Sumitomo Dainippon, Sunovion, Supernus, Takeda, and Teva. He has provided expert testimony for Bristol-Myers Squibb, Janssen, and Otsuka. He served on a Data Safety Monitoring Board for Boehringer-Ingelheim, Lundbeck, Rovi, Supernus, and Teva. He received royalties from UpToDate and grant support from Janssen and Takeda. He is also a shareholder in LB Pharma.

**Authorship Contributions.** All authors met the ICMJE criteria for authorship. Especially, J.S., E.B., F.S., and D.N. contributed to the design of the study. J.S., E.B., and P.V. contributed to the statistical analysis and checking the integrity of the data. J.S., M.K., and F.S. contributed to data collection. J.S., E.B., M.K., F.S., and C.C., contributed to the interpretation of the results. All authors contributed to writing the manuscript and approved the final version.

**Supplementary Materials.** To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/S1092852920001777>.

## References

- Alloy LB, Ng TH, Titone MK, Boland EM. Circadian rhythm dysregulation in bipolar spectrum disorders. *Curr Psychiatry Rep.* 2017;**19**(4):21. doi:10.1007/s11920-017-0772-z.
- Murray G, Harvey A. Circadian rhythms and sleep in bipolar disorder. *Bipolar Disord.* 2010;**12**(5):459–472. doi:10.1111/j.1399-5618.2010.00843.x.
- Gold A, Sylvia L. The role of sleep in bipolar disorder. *Nat Sci Sleep.* 2016;**8**:207–214. doi:10.2147/NSS.S85754.
- Geoffroy PA, Scott J, Boudebessé C, et al. Sleep in patients with remitted bipolar disorders: a meta-analysis of actigraphy studies. *Acta Psychiatr Scand.* 2015;**131**(2):89–99. doi:10.1111/acps.12367.
- Geoffroy PA, Boudebessé C, Bellivier F, et al. Sleep in remitted bipolar disorder: a naturalistic case-control study using actigraphy. *J Affect Disord.* 2014;**158**:1–7. doi:10.1016/j.jad.2014.01.012.
- Millar A, Espie CA, Scott J. The sleep of remitted bipolar outpatients: a controlled naturalistic study using actigraphy. *J Affect Disord.* 2004;**80**(2–3):145–153. doi:10.1016/S0165-0327(03)00055-7.
- St-Amand J, Provencher MD, Bélanger L, Morin CM. Sleep disturbances in bipolar disorder during remission. *J Affect Disord.* 2013;**146**(1):112–119. doi:10.1016/j.jad.2012.05.057.
- Gershon A, Kaufmann CN, Depp CA, et al. Subjective versus objective evening chronotypes in bipolar disorder. *J Affect Disord.* 2018;**225**:342–349. doi:10.1016/j.jad.2017.08.055.
- Kaufmann CN, Gershon A, Depp CA, Miller S, Zeitzer JM, Ketter TA. Daytime midpoint as a digital biomarker for chronotype in bipolar disorder. *J Affect Disord.* 2018;**241**:586–591. doi:10.1016/j.jad.2018.08.032.
- Milhiet V, Etain B, Boudebessé C, Bellivier F. Circadian biomarkers, circadian genes and bipolar disorders. *J Physiol.* 2011;**105**(4–6):183–189. doi:10.1016/j.jphysparis.2011.07.002.
- Merikangas KR, Swendsen J, Hickie IB, et al. Real-time mobile monitoring of the dynamic associations among motor activity, energy, mood, and sleep in adults with bipolar disorder. *JAMA Psychiatry.* 2019;**76**(2):190–198. doi:10.1001/jamapsychiatry.2018.3546.
- Scott J, Murray G, Henry C, et al. Activation in bipolar disorders. *JAMA Psychiatry.* 2017;**74**(2):189. doi:10.1001/jamapsychiatry.2016.3459.
- Scott J. Clinical parameters of circadian rhythms in affective disorders. *Eur Neuropsychopharmacol.* 2011;**21**(Suppl. 4):S671–S675. doi:10.1016/j.euro-neuro.2011.07.006.
- Jones SH, Hare DJ, Evershed K. Actigraphic assessment of circadian activity and sleep patterns in bipolar disorder. *Bipolar Disord.* 2005;**7**(2):176–186. doi:10.1111/j.1399-5618.2005.00187.x.
- Salvatore P, Ghidini S, Zita G, et al. Circadian activity rhythm abnormalities in ill and recovered bipolar I disorder patients. *Bipolar Disord.* 2008;**10**(2):256–265. doi:10.1111/j.1399-5618.2007.00505.x.
- Harvey AG, Schmidt DA, Scarnà 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**(1):50–57. doi:10.1176/appi.ajp.162.1.50.
- Gershon A, Thompson WK, Eidelman P, McGlinchey EL, Kaplan KA, Harvey AG. Restless pillow, ruffled mind: sleep and affect coupling in interepisode bipolar disorder. *J Abnorm Psychol.* 2012;**121**(4):863–873. doi:10.1037/a0028233.
- Ritter PS, Marx C, Lewtschenko N, et al. The characteristics of sleep in patients with manifest bipolar disorder, subjects at high risk of developing the disease and healthy controls. *J Neural Transm.* 2012;**119**(10):1173–1184. doi:10.1007/s00702-012-0883-y.
- McKenna BS, Drummond SPA, Eyer LT. Associations between circadian activity rhythms and functional brain abnormalities among euthymic bipolar patients: a preliminary study. *J Affect Disord.* 2014;**164**:101–106. doi:10.1016/j.jad.2014.04.034.
- Sebela A, Kolenic M, Farkova E, Novak T, Goetz M. Decreased need for sleep as an endophenotype of bipolar disorder: an actigraphy study. *Chronobiol Int.* 2019;**36**(9):1227–1239. doi:10.1080/07420528.2019.1630631.
- Bei B, Wiley JF, Trinder J, Manber R. Beyond the mean: a systematic review on the correlates of daily intraindividual variability of sleep/wake patterns. *Sleep Med Rev.* 2016;**28**:108–124. doi:10.1016/j.smrv.2015.06.003.
- Krane-Gartiser K, Henriksen TEG, Morken G, Vaaler A, Fasmer OB. Actigraphic assessment of motor activity in acutely admitted inpatients with bipolar disorder. *PLoS One.* 2014;**9**(2). doi:10.1371/journal.pone.0089574.
- Mullin BC, Harvey AG, Hinshaw SP. A preliminary study of sleep in adolescents with bipolar disorder, ADHD, and non-patient controls. *Bipolar Disord.* 2011;**13**(4):425–432. doi:10.1111/j.1399-5618.2011.00933.x.
- Faedda GL, Ohashi K, Hernandez M, et al. Actigraph measures discriminate pediatric bipolar disorder from attention-deficit/hyperactivity disorder and typically developing controls. *J Child Psychol Psychiatry Allied Discip.* 2016;**57**(6):706–716. doi:10.1111/jcpp.12520.
- Krane-Gartiser K, Scott J, Nevoret C, et al. Which actigraphic variables optimally characterize the sleep-wake cycle of individuals with bipolar disorders? *Acta Psychiatr Scand.* 2019;**139**(3):0–2. doi:10.1111/acps.13003.
- Janney CA, Fagiolini A, Swartz HA, Jakicic JM, Holleman RG, Richardson CR. Are adults with bipolar disorder active? Objectively measured physical activity and sedentary behavior using accelerometry. *J Affect Disord.* 2014;**152–154**(1):498–504. doi:10.1016/j.jad.2013.09.009.
- Gonçalves BSB, Adamowicz T, Louzada FM, Moreno CR, Araujo JF. A fresh look at the use of nonparametric analysis in actimetry. *Sleep Med Rev.* 2015;**20**:84–91. doi:10.1016/j.smrv.2014.06.002.
- Gonzalez R, Suppes T, Zeitzer J, et al. The association between mood state and chronobiological characteristics in bipolar I disorder: a naturalistic, variable cluster analysis-based study. *Int J Bipolar Disord.* 2018;**6**(1):5. doi:10.1186/s40345-017-0113-5.

29. De Crescenzo F, Economou A, Sharpley AL, Gormez A, Queded DJ. Actigraphic features of bipolar disorder: a systematic review and meta-analysis. *Sleep Med Rev.* 2017;**33**:58–69. doi:10.1016/j.smrv.2016.05.003.
30. Ng TH, Chung KF, Ho FYY, Yeung WF, Yung KP, Lam TH. Sleep-wake disturbance in interepisode bipolar disorder and high-risk individuals: a systematic review and meta-analysis. *Sleep Med Rev.* 2015;**20**:46–58. doi:10.1016/j.smrv.2014.06.006.
31. Kaplan KA, Talbot LS, Gruber J, Harvey AG. Evaluating sleep in bipolar disorder: comparison between actigraphy, polysomnography, and sleep diary. *Bipolar Disord.* 2012;**14**(8):870–879. doi:10.1111/bdi.12021.
32. Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry.* 1979;**134**(4):382–389. doi:10.1192/bjp.134.4.382.
33. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry.* 1978;**133**(5):429–435. doi:10.1192/bjp.133.5.429.
34. APA. *Diagnostic and Statistical Manual of Mental Disorders : DSM-5.* Arlington, VA, Washington, DC: American Psychiatric Association; 2013.
35. Tohen M, Frank E, Bowden CL, *et al.* The International Society For Bipolar Disorders (ISBD) task force report on the nomenclature of course and outcome in bipolar disorders. *Bipolar Disord.* 2009;**11**(5):453–473. doi:10.1111/j.1399-5618.2009.00726.x.
36. Lecrubier Y, Sheehan D, Weiller E, *et al.* The mini international neuropsychiatric interview (mini). a short diagnostic structured interview: reliability and validity according to the CIDI. *Eur Psychiatry.* 1997;**12**(5):224–231. doi:10.1016/S0924-9338(97)83296-8.
37. Macfadden W, Alphas L, Haskins JT, *et al.* A randomized, double-blind, placebo-controlled study of maintenance treatment with adjunctive risperidone long-acting therapy in patients with bipolar I disorder who relapse frequently. *Bipolar Disord.* 2009;**11**(8):827–839. doi:10.1111/j.1399-5618.2009.00761.x.
38. Juda M, Vetter C, Roenneberg T. The Munich chronotype questionnaire for shift-workers (MCTQ<sup>Shift</sup>). *J Biol Rhythm.* 2013;**28**(2):130–140. doi:10.1177/0748730412475041.
39. Cornelissen G. Cosinor-based rhythmometry. *Theor Biol Med Model.* 2014;**11**(1):16. doi:10.1186/1742-4682-11-16.
40. Witting W, Kwa IH, Eikelenboom P, Mirmiran M, Swaab DF. Alterations in the circadian rest-activity rhythm in aging and Alzheimer's disease. *Biol Psychiatry.* 1990;**27**(6):563–572. doi:10.1016/0006-3223(90)90523-5.
41. Vostatek P. Mindpax sleep classification. *Mindpax White-papers.* 2018;1–6. [https://www.mindpax.me/assets/docs/Validation\\_study\\_appendix.pdf](https://www.mindpax.me/assets/docs/Validation_study_appendix.pdf).
42. Boudebesse C, Leboyer M, Begley A, *et al.* Comparison of five actigraphy scoring methods with bipolar disorder. *Behav Sleep Med.* 2013;**11**(4):275–282. doi:10.1080/15402002.2012.685997.
43. Kosmadopoulos A, Sargent C, Darwent D, Zhou X, Roach GD. Alternatives to polysomnography (PSG): a validation of wrist actigraphy and a partial-PSG system. *Behav Res Methods* 2014;**46**(4):1032–1041. doi:10.3758/s13428-013-0438-7.
44. Santisteban JA, Brown TG, Gruber R. Association between the Munich chronotype questionnaire and wrist actigraphy. *Sleep Disord.* 2018;**2018**:1–7. doi:10.1155/2018/5646848.
45. Holm S. Board of the Foundation of the Scandinavian Journal of Statistics. A simple sequentially rejective multiple test procedure. *Scand J Stat.* 1979;**6**(2):65–70.
46. Breiman L. *Random Forests.* Berkley, VA: Statistics Department University of California; 2001:1–33. doi:10.1023/A:1010933404324.
47. Roenneberg T, Kuehnle T, Juda M, *et al.* Epidemiology of the human circadian clock. *Sleep Med Rev.* 2007;**11**(6):429–438. doi:10.1016/j.smrv.2007.07.005.
48. Ortiz A, Bradler K, Hintze A. Episode forecasting in bipolar disorder: is energy better than mood? *Bipolar Disord.* (2018);**20**:470–476. doi:10.1111/bdi.12603.
49. Bullock B, Murray G. Reduced amplitude of the 24 hour activity rhythm. *Clin Psychol Sci.* 2014;**2**(1):86–96. doi:10.1177/2167702613490158.
50. Gershon A, Ram N, Johnson SL, Harvey AG, Zeitzer JM. Daily actigraphy profiles distinguish depressive and interepisode states in bipolar disorder. *Clin Psychol Sci.* 2016;**4**(4):641–650. doi:10.1177/2167702615604613.
51. Shou H, Cui L, Hickie I, *et al.* Dysregulation of objectively assessed 24-hour motor activity patterns as a potential marker for bipolar I disorder: results of a community-based family study. *Transl Psychiatry.* 2017;**7**(8):e1211. doi:10.1038/tp.2017.136.
52. Carr O, Saunders KEA, Tsanas A, *et al.* Variability in phase and amplitude of diurnal rhythms is related to variation of mood in bipolar and borderline personality disorder. *Sci Rep.* 2018;**8**(1):1–11. doi:10.1038/s41598-018-19888-9.
53. Grierson AB, Hickie IB, Naismith SL, Hermens DF, Scott EM, Scott J. Circadian rhythmicity in emerging mood disorders: state or trait marker? *Int J Bipolar Disord.* 2016;**4**(1):3. doi:10.1186/s40345-015-0043-z.
54. Judd LL. The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Arch Gen Psychiatry.* 2002;**59**(6):530–537. doi:10.1001/archpsyc.59.6.530.
55. Monti JM. The effect of second-generation antipsychotic drugs on sleep parameters in patients with unipolar or bipolar disorder. *Sleep Med.* 2016;**23**:89–96. doi:10.1016/j.sleep.2016.04.020.
56. Hwang JY, Choi JW, Kang SG, Hwang SH, Kim SJ, Lee YJ. Comparison of the effects of quetiapine XR and lithium monotherapy on actigraphy-measured circadian parameters in patients with bipolar II depression. *J Clin Psychopharmacol.* 2017;**37**(3):351–354. doi:10.1097/JCP.0000000000000699.
57. Hossain S, Mainali P, Bhimanadham NN, Imran S, Ahmad N, Patel RS. Medical and psychiatric comorbidities in bipolar disorder: insights from national inpatient population-based study. *Cureus.* 2019;**11**(9):e5636. doi:10.7759/cureus.5636.