Which actigraphic variables optimally characterize the sleep-wake cycle of individuals with bipolar disorders? Significant outcomes

K. Krane-Gartiser, J. Scott, C. Nevoret, V. Benard, C. Benizri, H. Brochard, P. A. Geoffroy, S. Katsahian, J. Maruani, S. Yeim, et al.

Which actigraphic variables optimally characterize the sleep-wake cycle of individuals with bipolar disorders?

Krane-Gartiser K, Scott J, Nevoret C, Benard V, Benizri C, Brochard H, Geoffroy PA, Katsahian S, Maruani J, Yeim S, Leboyer M, Bellivier F, Etain B. Which actigraphic variables optimally characterize the sleep-wake cycle of individuals with bipolar disorders?

Objective: To examine which combination of objectively measured actigraphy parameters best characterizes the sleep-wake cycle of euthymic individuals with bipolar disorder (BD) compared with healthy controls (HC).
Methods: Sixty-one BD cases and 61 matched HC undertook 21 consecutive days of actigraphy. Groups were compared using discriminant function analyses (DFA) that explored dimensions derived from mean values of sleep parameters (Model 1); variability of sleep parameters (2); daytime activity (3); and combined sleep and activity parameters (4). Exploratory within-group analyses examined characteristics associated with misclassification.
Results: After controlling for depressive symptoms, the combined model (4) correctly classified 75% cases, while the sleep models (1 and 2) correctly classified 87% controls. The area under the curve favored the combined model (0.86). Age was significantly associated with misclassification among HC, while a diagnosis of BD-II was associated with an increased risk of misclassifications of cases.
Conclusion: Including sleep variability and activity parameters alongside measures of sleep quantity improves the characterization of cases of euthymic BD and helps distinguish them from HC. If replicated, the findings indicate that traditional approaches to actigraphy (examining mean values for the standard set of sleep parameters) may represent a suboptimal approach to understanding sleep-wake cycles in BD.

K. Krane-Gartiser[1,2,3], J. Scott[1,4,5,6], C. Nevoret[7,8,9], V. Benard[3], C. Benizri[10], H. Brochard[10,11], P. A. Geoffroy[3,5,12,13], S. Katsahian[7,8,9], J. Maruani[12], S. Yeim[5,12], M. Leboyer[10,13,14,15], F. Bellivier[3,5,12,13], B. Etain[3,5,6,12,13]

Significant outcomes

- This clinical study of a matched case–control sample demonstrated that dimensional approaches to analysis of actigraphy parameters may help to characterize the sleep-wake cycle of individuals with bipolar disorder and more reliably differentiate them from healthy controls (compared to traditional approaches).
- A combined model (comprised of actigraphic recordings of sleep quantity and variability alongside activity parameters) improved the classification of cases of bipolar disorders compared to mean sleep values alone.
- Knowledge and understanding of sleep-wake cycle abnormalities that are key markers of bipolarity could be improved if actigraphy studies avoid convenience sampling and limit reliance on univariate analyses of mean values of individual sleep parameters.

Limitations

- Few actigraphy studies of bipolar disorders report daytime activity parameters, and there is less consensus on the optimal markers of this construct; as such, different combinations of activity parameters could show dissimilar discriminatory abilities.
- Although this is one of the largest clinical studies of actigraphy in bipolar disorders, there was insufficient statistical power to stratify cases according to bipolar disorder subtype or classes of prescribed medication for the exploratory within-group analysis of misclassifications.

[1]Department of Mental Health, NTNU, Norwegian University of Science and Technology, [2]Department of Psychiatry, St. Olav's University Hospital, Trondheim, Norway, [3]INSERM U1144, Paris, France, [4]Academic Psychiatry, Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, UK, [5]Sorbonne Paris Cité, Université Paris Diderot, Paris, France, [6]Centre for Affective Disorders, Institute of Psychiatry, London, UK, [7]INSERM, UMR_S 1138, Université Paris Descartes, Sorbonne Universités, UPMC Université Paris 06, UMR_S 1138, Centre de Recherche des Cordeliers, Paris, France, [8]Assistance Publique – Hôpitaux de Paris, Hôpital Européen Georges-Pompidou, Unité d'Épidémiologie et de Recherche Clinique, Paris, France, [9]INSERM, Centre d'Investigation Clinique 1418, Module Épidémiologie Clinique, Paris, France, [10]Equipe Psychiatrie Translationnelle, INSERM U955, Créteil, France, [11]Pôle sectoriel, Centre Hospitalier Fondation Vallée, Gentilly, France, [12]Département de Psychiatrie et de Médecine Addictologique, AP-HP, GH Saint-Louis – Lariboisière – F. Widal, Paris, France, [13]Fondation FondaMental, Créteil, France, [14]AP-HP, Hôpitaux Universitaires Henri Mondor, DHU Pepsy, Pôle de Psychiatrie et d'Addictologie, Créteil, France and [15]Université Paris Est Créteil, Creteil, France
Key words: actigraphy; bipolar disorder; classification; variability; activity

Jan Scott, Institute of Neuroscience, Newcastle University, Newcastle, NE7 6RU, UK.
E-mail: Jan.scott@newcastle.ac.uk

Introduction

Actigraphy is a frequently recommended, ecologically valid, objective measure of sleep and circadian disruptions (1), and it is increasingly employed in research into sleep and daytime patterns in bipolar disorders (BD) (2–4).

Limitations in clinical research designs and methodologies may have undermined the validity of some of the findings regarding actigraphy in BD (2–5). For example, the standard set of variables derived from actigraphy recordings are most relevant to investigations of insomnias and related sleep disorders (1). Current actigraphy software automatically provides a set of so-called sleep parameters of which the most frequently examined are total sleep time (TST), sleep onset latency (SOL), wake after sleep onset (WASO), sleep efficiency (SE), and Fragmentation Index (FI). Most actigraphy research in BD reports the mean values for these variables, and individual studies and meta-analyses have examined these quantitative aspects of sleep in BD (3–9). However, this research usually gives equal weighting to each sleep parameter and reports analyses of each variable separately. Overall, findings are inconsistent regarding the utility of individual parameters for identifying BD cases (3–9). Despite this, few investigations explore whether combinations of variables or sleep dimensions are better able to characterize the sleep-wake cycle in BD cases and differentiate these from healthy controls (HC) (10, 11).

Systematic reviews and meta-analyses of actigraphy in BD (2, 4, 5) reveal significant sources of study bias, including small sample sizes (median ~30), limited statistical power to detect group differences, short duration of actigraphic monitoring, and over-reliance on convenience sampling of comparator groups with demographic and health profiles that differ markedly from participants with BD. This can be problematic, as age differences between BD cases and HC influence the effect size for TST differences (4), and sex and body mass index (BMI) are associated with differences in mean values for several sleep parameters (12, 13). A further issue is heterogeneity within BD subsamples, such as inclusion of individuals with spectrum disorders and/or mixed samples of asymptomatic and symptomatic BD cases (2, 14).

As well as the need to consider sampling and methodological differences in previous BD studies (2, 4, 5, 15), investigators have advocated for the extension of actigraphy to examine sleep variability and daytime activity (2, 16, 17). Existing studies suggest that measuring variability of sleep patterns may increase our ability to characterize euthymic BD cases or identify individuals with or without a family history of BD (5, 16, 18, 19). However, these studies usually benefit from more extended periods of actigraphy monitoring (about two weeks is usually required to capture variability) (20). Additionally, recent studies have begun to examine regularity/intensity of daytime activity in BD, with some of these measures being employed as proxies of circadian dysrhythmias (2, 3, 11, 21–25). The relative lack of BD studies that have addressed these activity-based parameters might be explained by need for additional expertise to derive them (most are not automatically generated by the software) (2, 22, 23). To date, the available literature suggests that daytime activity patterns in currently euthymic BD cases are lower and less stable than healthy and other non-BD controls and that baseline values of relative amplitude of activity have predictive validity for future onset of BD and major depression (26–30).

To summarize, the extant literature on actigraphy in BD repeatedly demonstrates differences between cases and controls. However, publications consistently highlight that greater attention to potential biases in study design is required

and researchers may need to go beyond the reporting mean values of actigraphic sleep parameters and consider including sleep variability and markers of daytime activity patterns. Also, it is timely to analyze whether subsets of parameters can improve the utility of actigraphy for both characterizing the sleep-wake cycle dimensions in BD cases and for determining which actigraphic parameters (from within each dimension) are the best predictors of the classification of cases or controls.

Aims of the study

To recruit a sample of BD cases and closely matched HC, and

i) to use multivariate modeling, namely discriminant function analysis (DFA) to identify sleep and activity dimensions and then examine whether any dimensions can be used to characterize BD cases;
ii) to explore the specific sleep and activity parameters that significantly predict group membership (i.e. case or control), and
iii) to identify potential confounders of the classification of cases or controls.

Material and methods

Sample

Ethics: The study was approved by the Comité de Protection des Personnes from La Pitié-Salpétrière Hospital (reference: P111002-IDRCB2008-AO146 5-50) in Paris, France. All participants provided written informed consent prior to inclusion. This study is part of the GAN (genetics–actigraphy–neuropsychiatry) project which includes the collection of objective actigraphy recording of sleep and circadian rhythms from BD cases attending clinical services and from HC recruited from the general population (see below).

Selection Procedure: Cases and controls were selected for this study according to a three-stage process.

Stage 1: Recruitment. Individuals with BD were recruited from psychiatric clinics affiliated with the University of Paris between 2012 and 2017. The HC were recruited through several avenues including via advertisements (targeted at members of the public and those attending institutions, e.g. universities or government departments) and from individuals attending the blood donor service at a general hospital affiliated to the university. All participants were aged >18 years. Additional details of the sampling methods are described elsewhere (31).

Eligibility Criteria for BD cases: Individuals with BD were included if they reported a mood disorder that met DSM-IV criteria for a BD-I or BD-II diagnosis assessed using the Diagnostic Interview for Genetic Studies (DIGS) administered by trained psychiatrists (32); they were currently euthymic (defined as a score ≤ 8 on both the MADRS (33) and Young Mania Rating Scale (YMRS) (34)); they had been in remission for ≥ 3 months (according to the ISBD task force criteria (35)); and they had no recorded changes in the class(es) and dose(s) of prescribed psychotropic medication(s) during this period. Patients were excluded if they fulfilled the criteria for an alcohol misuse (abuse or dependence) disorder in the last two years (according to the DIGS assessment).

Eligibility Criteria for HC: HC were excluded if the individual met DSM-IV criteria (according to the DIGS assessment as administered by a trained psychiatrist) for a lifetime affective, schizophrenic, or alcohol misuse (abuse or dependence) disorder and/or had a lifetime history of suicide attempts. Also, HC were excluded if a first-degree relative had a lifetime history of affective or schizophrenic disorder and/or suicide attempts (according to the Family Interview for Genetic Studies (36)).

Stage 2: Actigraphy procedure. Participants were excluded from participating in the actigraphy study if, in the previous three months, they had

i) been hospitalized for any reason;
ii) received treatment(s) or taken drugs for somatic conditions that could affect their sleep pattern;
iii) experienced severe sleep disruption due to a somatic condition and/or experienced any life event that could have altered their sleep-wake habits (assessed using a checklist that included shift work, recent trans-meridian travel (with a >3-h time difference), pregnancy, childbirth, recent bereavement or somatic conditions).

Stage 3: Matching procedure. Eighty HC and 135 individuals with BD met eligibility criteria for study entry and commenced actigraphy recordings. At this stage, individuals were excluded from the current protocol (prior to matching cases and controls) if they

i) had an increased risk of obstructive sleep apnea (OSA) according to the Berlin Questionnaire (37);
ii) failed to complete three consecutive weeks of actigraphy monitoring;

iii) met criteria for a depressive or (hypo)manic episode during actigraphic monitoring (according to criteria described by the International Society of Bipolar Disorder (ISBD) task force (35)).

At the final step, we matched each eligible BD case with a HC according to demographic and anthropometric measures known to be associated with different sleep and rest-activity patterns, namely age (±5 years), sex (1:1), and body mass index (BMI; ±2 points). Any participant who could not be allocated to a case–control pair was excluded from the analyses. The sample for the current study comprised of 122 individuals (i.e., 61 case: control pairs).

Clinical assessment and actigraphy: procedure and parameters

Clinical assessment visits. At inclusion, participants were given a written information letter about the protocol and signed the informed consent form. They were then clinically assessed (using the DIGS for all participants plus the FIGS for HC). Participants completed the Berlin Questionnaire, and their BMI was estimated. The MADRS and the YMRS were rated at inclusion, and then, each participant was given instructions about how to use the actiwatch (see below).

After 21 days of recording, participants attended a follow-up appointment to check the completeness of the actigraphy data (this was done by visual inspection of the actigraphy output and via interviewing participants to clarify whether there were any time periods when they did not wear the device). If data were incomplete, participants were excluded from the study. Mood symptoms were recorded using the MADRS and the YMRS.

Actigraphy procedure. All participants wore an actigraph (AW-7 CamNtech) on the wrist of their non-dominant hand for 21 consecutive days. The participants were asked to wear the actiwatch continuously (as the Actiwatch AW7 is waterproof up to 6 BAR of pressure, it is not necessary to remove it when showering or swimming, etc). Participants were instructed to press the event marker on the actiwatch when they intended to go to sleep at night and when they got up in the morning. Participants completed a sleep diary during the monitoring period, which allowed concordance between diary and actigraphic recordings of bedtimes and rise times to be established. Actigraphy data (sampled in one-minute epochs) were analyzed using the Actiwatch Activity and Sleep Analysis software (CamNtech 7.28).

Parameters. We selected sleep and rest-activity parameters that best represent the variables found to be important in distinguishing BD cases from HC and/or from other comparator groups in previous publications (reported in the Introduction and summarized in Table 1).

i) Sleep parameters: The key sleep parameters derived from actigraphy software are TST, SOL, WASO, SE, and FI. These were explored from two perspectives:
   - Mean values: Means were estimated for the three weeks of monitoring for the five standard actigraphic sleep variables; this is the most common approach to report sleep data in actigraphy studies of BD.
   - Variability: As in previous studies, we calculated standard deviations (SDs) of the standard sleep parameters to obtain a measure of within-individual variability (7, 38).

   All these parameters were automatically calculated using the Actiwatch Activity and Sleep Analysis software (CamNtech 7.28). Standard deviations (SD) of sleep parameters across the 21 days period were calculated using SPSS (version 23).
ii) Daytime activity parameters: We identified four measures from the Van Someren's nonparametric variables (39) that are frequently reported in published studies of BD and also best represent the spectrum of activity (timing, amplitude, and stability/regularity). As shown in Table 1, the selected parameters were interdaily stability (IS); intradaily variability (IV); M10 onset; and relative amplitude (RA). Interdaily stability and intradaily variability were selected primarily because of the frequency of reporting in BD studies, while M10 onset and RA were selected as the best combination of parameters measuring regularity, timing, and quantity of daytime activity. The daytime activity parameters were automatically calculated from the software (Non-Parametric Circadian Rhythm Analysis (NPCRA)); thus, no extraction of raw data was required.

Statistical analysis

All analyses were undertaken using SPSS (version 23). We followed the recommendations of Mueller et al. (40) to undertake the power calculation (using the online G*Power program) and estimated that assuming an Eigenvalue of ≥0.3 (the magnitude of the discriminant function) for each model,

Table 1. Sleep, sleep variability, and activity cycle parameters most frequently used in publications on bipolar disorders and/or representing the key actigraphy measures (see text for details)

Actigraphy measures	Parameter	Definition
Sleep parameters—mean values	TST	Total sleep time (time between reported sleep onset and offset measured in hours and minutes)
	WASO	Wake after sleep onset (measured in minutes)
	SOL	Sleep onset latency (time between reported bedtime and sleep onset measured in minutes)
	SE	Sleep efficiency; the total sleep time divided by the total time spent in bed
	FI	Fragmentation index; a measure of sleep continuity. The FI is calculated as the amount of time associated with movement (restlessness) during the sleep period expressed as a percentage. A higher FI indicates more disrupted sleep.
Sleep parameters—variability	SD of TST	Standard deviation of total sleep time
	SD of WASO	Standard deviation of wake after sleep onset
	SD of SOL	Standard deviation of sleep onset latency
	SD of SE	Standard deviation of sleep efficiency
	SD of FI	Standard deviation of fragmentation index
Activity cycle parameters	IS	Interdaily stability is a measure of similarity in the diurnal pattern (hourly averaged activity) and quantifies the degree of regularity in the rest-activity patterns from day to day (range 0–1, with 1 = stable rhythms)
	IV	Intradaily variability quantifies the degree of fragmentation of rest-activity periods. Healthy individuals typically show a single prolonged activity period and a single prolonged rest period per 24-h cycle. The value of IV ranges from 0 to 2 (higher values indicate higher fragmentation of rest vs. activity).
	M10 onset	The M10 provides the average activity level for the sequence of the most active 10-h period during 24 h. The M10 onset indicates the starting time and so it provides an indication of the phase of the most active hours.
	RA	Amplitude (AMP) represents the difference between the average activity level in the most active 10-h period (M10) and least active 5-h period (L5). The RA is used to remove sensitivity to overall activity level, as it represents the AMP divided by the sum of all activity during these 15 h (using the formula RA = (M10−L5)/(M10 + L5)). RA has a theoretical range of 0–1 (higher values indicating a rhythm with higher amplitude)

TST, total sleep time; WASO, wake time after sleep onset; SOL, sleep onset latency; SE, sleep efficiency; FI, fragmentation index; IS, interdaily stability; IV, intradaily variability; L5, least active 5 h; M10, most active 10 h; RA, relative amplitude.

then the current sample size has 80% power to detect between-group differences with statistical significance of $P < 0.05$ in multivariate analyses.

The planned analyses proceeded in three stages. First, we compared BD cases and HC on characteristics used in the matching procedure and on symptom rating scale scores; then, we examined which sleep and daytime activity variables optimally classified BD cases. Finally, we undertook an exploratory analysis of any characteristics that identified those BD cases or HC that were consistently misclassified.

Descriptive analyses. Categorical data are reported as counts and percentages; as some continuous variables showed non-normal distributions, these are reported as medians and interquartile ranges (IQR). We report Fisher's exact tests and Mann–Whitney *U*-tests to demonstrate the quality of matching of cases and controls (i.e., *P*-values are reported only to show the nearness of matching).

Discriminant function analyses. Discriminant function analysis is a multiple regression technique that determines the best weighting of variables to maximize the differences among groups and predict group membership (41). We employed the standardized residual scores for each sleep and activity parameter as predictors in the model, and we report the proportion of cases correctly classified by the DFA. We examined the reliability of this classification using a 'leave-one-out' approach cross-validation analysis (41).

We undertook four separate stepwise DFA (10) to determine which set of sleep and/or activity parameters from within each dimension were most useful for identifying individuals as BD cases or HC. All four DFA included age, sex, BMI, and MADRS scores as covariates. To avoid multicollinearity, we included only the MADRS as a symptom severity score (the MADRS and YMRS were significantly correlated). For all models, the probabilities of F for actigraphy parameters to be entered and removed were $P = 0.15$ and $P = 0.20$ respectively. Each DFA introduced a different set of actigraphy parameters: (i) Basic set included mean values for the sleep parameters; (ii) variability set included the SDs of the sleep parameters; (iii) daytime activity set included activity cycle parameters (which also serve as proxies for circadian rhythmicity); and (iv) combined sleep and activity set included all the selected actigraphy parameters. Lastly, we repeated the DFA for the final combined model but excluded the MADRS score to determine whether there was a change in variables included in the model or the classification rate.

In the main text, we summarize the DFA models and those parameters that contributed to the final step of the DFA model. In Table 4, we report the magnitude of the discriminant function (Eigen values) and the classification rate for cases and

controls, while the supplementary materials provide additional details for the canonical discriminant functions. Finally, the group membership probabilities generated by each DFA were used to calculate receiver operating characteristics (ROC) and the area under the curve (AUC) for each model. We performed post hoc testing of the differences between the AUC (see Figure S1).

Exploratory analysis of misclassifications. We identified the number of BD cases or HC who were incorrectly classified at least once by DFA and those that were always incorrectly classified. Within-group (hypothesis-free) univariate analyses were undertaken to explore which baseline characteristics best identified individuals who were misclassified (compared with those who were correctly classified) using Mann–Whitney *U*-test and Fisher's exact tests. Within the BD group, we also explored whether misclassification was associated with BD subtype, duration of illness, number of BD episodes (in total and by polarity), and classes of medication prescribed.

Results

Sample characteristics

The sample comprised of 61 individuals with BD and 61 HC who were matched for age, gender, and BMI. Table 2 shows the baseline characteristics for BD cases and HC (raw scores for actigraphy parameters are shown in Table S1). Matching of BD cases and HC was complete for sex, and the total sample showed a slight preponderance of females (54%). Matching was also acceptable for age ($P = 0.196$) and BMI ($P = 0.603$). Although the median scores were zero for the MADRS and YMRS, BD cases were statistically significantly more likely than HC to report some symptoms. Further analysis demonstrated that the means and standard deviations for the MADRS (BD cases: mean = 1.25, SD = 1.93; HC: mean = 0.21, SD = 43) were slightly higher than for YMRS (BD cases: mean = 0.89, SD = 1.72; HC: mean = 0.13, SD = 0.3), but that all symptom levels were very low.

As shown in Table 3, the median age at onset of BD was about 23, and individuals had been ill for about 14 years at study entry. Three quarters of cases met diagnostic criteria for BD-I ($n = 46$). About a third had a lifetime history of alcohol misuse (though not currently); >50% reported use of nicotine (data not shown). The most commonly prescribed mood stabilizers were lithium and anticonvulsants; just over half of the cases ($n = 33$; 54%) were receiving mood stabilizer monotherapy.

Table 2. Baseline characteristics of BD cases and healthy controls

	BD cases ($n = 61$)	Healthy controls ($n = 61$)	Sig.†
Characteristics used for matching			
Median age in years (IQR)	38 (31, 54)	34 (28, 53)	0.196
Number of females (%)	33 (54%)	33 (54%)	1.000
Median body mass index in kg/m^2 (IQR)	24.3 (21.8, 27.2)	23.4 (20.7, 27.7)	0.603
Median scores on symptom rating scales (IQR)*			
Montgomery–Asberg Depression Rating Scale	0 (0, 2.5)	0 (0, 0)	<0.001
Young Mania Rating Scale	0 (0, 1)	0 (0, 0)	<0.001

IQR, interquartile range; % are reported to the nearest whole number.
*Ratings completed at the end of the three-week monitoring period.
†Mann–Whitney *U*-test for continuous variables, Fisher's exact tests for categorical data (*P* values quoted for matching variables are provided to give an indication of any trends).

Discriminant function analyses

A summary of the key DFA findings is shown in Table 4 (with additional information on the summary statistics in Table S2). As shown, significant predictors in the basic model were MADRS score, WASO mean, and FI mean with borderline significance for the TST mean ($P = 0.06$); and the DFA correctly classified 61% of BD cases and more than four-fifths of HC (87%). The most significant predictors in the variability model were MADRS score and FI variability (with borderline significance for WASO variability; $P < 0.054$). This DFA also classified 87% of HC, but it was less reliable than the basic model for classifying BD cases (52%). In the third DFA (activity model), MADRS score and M10 onset were the significant predictors; this model demonstrated similar

Table 3. Clinical characteristics of individuals with bipolar disorders (BD cases)

Variable	BD cases ($n = 61$)*
Number with bipolar I disorder	46 (75%)
Age of onset in years	22.5 (19.0, 30.0)
Duration of illness in years	14.0 (9.8, 20.5)
Number of episodes	6 (4, 8)
Prescribed medication at study entry	
Number of psychotropics	2 (1, 3)
Lithium treatment	36 (59%)
Anticonvulsant treatment	30 (49%)
Antipsychotic treatment	21 (34%)
Antidepressant treatment	15 (25%)
Hypnotics	7 (12%)

*Medians with interquartile ranges (IQR) for continuous variables; Numbers are counts with percentages (% are reported to the nearest whole number).

classification rates for HC (85%) and BD cases (56%) as the DFA models based on sleep parameters.

The combined model provided the best classification of BD cases. The most significant predictors in this DFA were MADRS score, FI mean, FI variability, WASO variability, and IS and M10 onset, with TST mean showing borderline significance ($P = 0.09$). These variables correctly classified similar proportions of BD cases (80%) and HC (75%). If the MADRS score is excluded from the analysis, the actigraphy variables correctly classified 67% of cases and controls.

All four DFAs produced a highly significant function ($P < 0.001$), but the highest Eigenvalue (a measure of the discriminant ability of the function) and optimal classification of BD cases and HC was provided by the combined model. Further, level of depressive symptoms (MADRS score) significantly contributed to all the DFAs. In contrast, SE and SOL (mean or variability) did not contribute to any models in which they were considered. Likewise, age, gender, and BMI did not contribute to any DFAs.

Using estimated group membership probabilities, we produced ROC curves for each DFA (see Figure S1). The AUC was 0.86 for the combined model, compared with 0.82 for the variability model, 0.79 for the activity model, and 0.78 for the basic model.

Exploratory analyses of misclassifications

Within-group analysis of the BD cases demonstrated that 23% (14 individuals) were consistently misclassified by the DFA models, but there were no significant associations between misclassification and any clinical, treatment, or demographic variables. Twelve of the 15 cases with BD-II were misclassified at least once, and there was a trend ($P = 0.09$) for BD-II cases to be misclassified more often than BD-I cases.

Twenty-one HC (33%) were misclassified at least once, but only three individuals were misclassified in all DFA models. Misclassified HC were significantly older than correctly classified HC (median age 54 vs. 31; $P = <0.001$).

Discussion

The present study used a matched case–control design and a multivariate modeling approach to identify sleep and daytime activity dimensions and to determine the actigraphic parameters that best characterized BD cases and differentiated them from HC. The methodology included steps to

Table 4. Summary of key outputs of the four discriminant function analysis models (see text and Table S1 for additional details)

Grouping variables	Selected parameters	Basic model	Variability model	Activity model	Combined model
Demographic, anthropometric, and clinical characteristics	MADRS	**<0.001**	**<0.001**	**<0.001**	**<0.001**
	Age	ns	ns	ns	ns
	Gender	ns	ns	ns	ns
	BMI	ns	ns	ns	ns
Actigraphic measures of sleep (means)	TST	0.06			0.09
	WASO	**0.025**			ns
	SOL	ns			ns
	SE	ns			ns
	FI	**0.004**			**0.016**
Actigraphic measures of sleep (variability)	SD of TST		ns		0.13
	SD of WASO		0.054		**0.001**
	SD of SOL		ns		ns
	SD of SE		ns		ns
	SD of FI		**0.002**		**<0.001**
Actigraphic measures of activity	IS			ns	**0.032**
	IV			ns	ns
	M10 onset			**0.036**	**0.038**
	RA			ns	ns
Eigenvalue	Discriminant function	0.32	0.33	0.27	0.55
Classification rates*	Cases	61%	52%	56%	75%
	Controls	87%	87%	85%	80%

MADRS, Montgomery Asberg Depression Rating Scale score; BMI, body mass index; TST, total sleep time; WASO; wake time after sleep onset; SOL, sleep onset latency; SE, sleep efficiency; FI, fragmentation index; IS, interdaily stability; IV, intradaily variability; M10 onset, start time of the most active 10 h; RA, relative amplitude. Values are *P*-values unless otherwise is noted; *P*-values in bold font identify the variables that contribute significantly to the discriminant function; *P*-values are reported for other variables if they were included in the final step of the analysis; ns indicates the variable made no significant contribution to that model.

*Percentages refer to the number of individuals that are correctly classified by each DFA model.

minimize confounding, and we performed within-group analyses to explore which variables could explain misclassifications of BD cases and HC. We suggest that there are four key findings with clear implications for the field of actigraphy research in BD. First, extending studies to include activity parameters alongside measures of sleep quantity and variability (the combined model) may be useful for characterizing BD cases. Also, this DFA model (which included activity parameters, measures of sleep quantity, and variability) provided the best classification of euthymic BD cases and HC. Second, if the primary goal of a study is to exclude non-cases (i.e., correctly identify HC), focusing on sleep (mean values) or sleep variability dimensions may be most appropriate. Third, the traditional approach to using actigraphy in BD (focusing on mean values of the standard set of sleep parameters) produces a basic model that shows the lowest AUC of the four models we examined. If confirmed by further research, this may indicate that the most commonly used analytic strategy in actigraphy studies of BD cases and controls (univariate comparisons between groups of mean values for sleep parameters) has the lowest utility and may be the least valid. Fourth, while the within-group analyses were exploratory, they emphasize the need to consider the potential for biases to be introduced into studies because of heterogeneity within the control group (e.g., wide age range in HC). Also, the trend toward misclassification of BD-II as compared with BD-I cases suggests it may be worthwhile further investigating underlying differences in the actigraphic characteristics of these diagnostic subtypes. This is timely, given the findings from a recently published, large-scale community study that reported that individuals with BD-I show greater variability in activity during the afternoon, while BD-II show greater variability at nighttime (16). In the following paragraphs, we expand the discussion of the findings from the current study in the context of past and future research strategies and address its strengths and limitations.

To our knowledge, this study is the first to suggest an incremental value of combining sleep quantity, sleep variability, and activity cycle parameters to improve the characterization of BD cases and differentiate them from HC. The ROC analyses highlight that the DFA for the combined model has the best discriminant ability and high accuracy (AUC 0.86). In this model, parameters from each of the three sleep-wake cycle domains contributed to the discriminant function; of most importance were two variability measures (FI and WASO), two activity measures (IS and M10 onset) and one quantitative sleep measure (mean FI). Conversely, other parameters (SOL, SE) that were shown to be significant in previously published univariate analyses did not contribute to the classification of cases and controls when dimensional approaches were used and key covariates were considered (e.g., level of residual depressive symptoms). Studies such as this one may help researchers in the search for a set of actigraphy parameters that might be the most relevant to explore in future studies of BD (without giving priority to only one sleep or activity domain).

The existing literature suggests that sleep variability may be a more prominent feature of euthymic BD than mean values of sleep parameters (18, 38). Our findings indicate that there may be a need for a more cautious interpretation, namely that the absence of variability in the HC made this a useful dimension for identifying controls. Further, on its own, the sleep variability dimension did not improve on the classification of BD cases offered by the basic model (52% vs. 61% respectively). We suggest that to fully understand these findings and expand on recent research (e.g., (31)), future studies might report both sleep quantity and variability (especially in situations where activity parameters cannot be examined).

The relevance of activity parameters for characterizing BD cases is understudied. Three systematic reviews that have synthesized the limited data available all indicate anomalies in intensity, regularity, and timing of activity in BD populations (2, 3, 5). In our study, the activity model showed that BD was associated only with objective alterations in M10 onset (a phase marker). However, it was notable that the findings regarding the activity dimension were less impressive for case–control classification than when it was combined with sleep measures in the combined model. Unlike a recent population-based study (30), we did not find that RA contributed to the combined model, which opens an intriguing possibility that some activity (or circadian markers) may be useful in discriminating individuals at risk of BD from those who are not, while others better characterize established cases (5, 8, 19, 42, 43).

This study has several strengths. First, we assessed a broad spectrum of sleep abnormalities (quantity and variability) as well as daytime activity (timing and regularity, etc). Second, we employed multivariate modeling as our primary statistical approach; this is noteworthy because DFA assumes that combinations of parameters are more relevant than each individual variable for determining group membership. Third, we undertook an extended period of 21 consecutive days of

actigraphy recordings, which increases the possibilities for reliably capturing sleep variability and irregular activity. Fourth, we matched clinical BD cases (for age, sex, and BMI) with HC recruited from a similar geographic location. Also, we tried to ensure that key demographic and anthropometric variables and depressive symptoms did not confound the analyses. These methodological approaches may partly explain why we found that residual symptoms (as measured by the MADRS) were a necessary but not sufficient contributor to the differentiation of BD cases from HC, in contrast to a recent community study of BD-I (14). However, we recognize that the MADRS score significantly and consistently contributed to all the DFAs, highlighting that even low-grade depressive symptoms should be included as a covariate in analyses of sleep-wake patterns in BD. This is emphasized by the decrease in the overall classification rate for the combined model when the MADRS score is excluded.

There are several limitations to the present study that should be addressed in future investigations. For instance, the estimation of statistical power for multivariate approaches can be complex (40) and the trends toward statistical significance for some parameters in our DFA models suggest that we cannot rule out false-negative findings. Likewise, the within-group exploratory analyses of BD were at risk of false-positive findings. In the event, we found only trends toward the greater likelihood of misclassification of BD-II cases, highlighting the fact that future studies not only need to consider matching of cases and controls, but also what proportion of individuals with BD-I (or BD-II) would allow analyses to be stratified by BD subtype. Also, there is no consensus on the most appropriate combination of activity variables for exploration in BD; therefore, studies focusing on a different set of activity parameters may yield different findings from those reported here. Further, we could not examine the impact of habitual patterns of daytime activity or entrained routines. As there is emerging evidence that these issues influence findings in ecological momentary assessment research (2, 19, 44), they will need careful consideration in the future. Likewise, screening for OSA relied on a questionnaire rather than more sophisticated methods (45). An enduring issue for BD case–control studies is that between-group differences in actigraphy parameters may be partly explained by exposure to medications (46) and mood stabilizers can affect several circadian parameters (47). Therefore, some potential case–control differences might have been masked by the fact that all patients with BD were currently being prescribed medications that could modify their circadian rhythms. While the recruitment of unmedicated patients (including those who are euthymic) is difficult in clinical settings, it is important to bear in mind the difficulty in determining the relative contributions of medications as compared to illness characteristics to differences in actigraphy recordings (46).

Similarly, current exposure to tobacco and alcohol might contribute to the differences observed between patients and controls and future studies should explore the influence of these substances on the actigraphy profile of patients with BD.

In conclusion, the current research suggests that a dimensional approach is beneficial as it identifies five sleep-wake cycle parameters that, when combined together, offered the optimal prediction of BD caseness. Also, the study methodology is relatively easy to reproduce, so it should be possible for other researchers to explore these sleep-wake cycle dimensions and determine whether the findings can be replicated in independent samples.

Funding

Authors acknowledge the support provided by INSERM (Research Protocol C0829), Assistance Publique des Hôpitaux de Paris (Research Protocol GAN12). This research was also supported by the Investissements d'Avenir program managed by the ANR under reference ANR-11-IDEX-0004 and Fondation Fondamental (RTRS Santé Mentale).

Declaration of interest

The authors declare no financial or other conflict of interests in relation to the topics addressed in this paper.

References

1. Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP. The role of actigraphy in the study of sleep and circadian rhythms. Sleep 2003;**26**:342–392.
2. Scott J, Murray G, Henry C et al. Activation in bipolar disorders: a systematic review. JAMA Psychiatry 2017;**74**:189–196.
3. De Crescenzo F, Economou A, Sharpley AL, Gormez A, Quested DJ. Actigraphic features of bipolar disorder: a systematic review and meta-analysis. Sleep Med Rev 2017;**33**:58–69.
4. Geoffroy PA, Scott J, Boudebesse C et al. Sleep in patients with remitted bipolar disorders: a meta-analysis of actigraphy studies. Acta Psychiatr Scand 2015;**131**:89–99.
5. Ng TH, Chung KF, Ho FY, 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;**20C**:46–58.
6. Harvey AG, Schmidt DA, Scarna A, Semler CN, Goodwin GM. Sleep-related functioning in euthymic patients with bipolar disorder, patients with insomnia, and subjects without sleep problems. Am J Psychiatry 2005;**162**:50–57.

7. 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**:425–432.
8. 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**:1173–1184.
9. Kaplan KA, Talbot LS, Gruber J, Harvey AG. Evaluating sleep in bipolar disorder: comparison between actigraphy, polysomnography, and sleep diary. Bipolar Disord 2012;**14**:870–879.
10. Scott J, Vaaler AE, Fasmer OB, Morken G, Krane-Gartiser K. A pilot study to determine whether combinations of objectively measured activity parameters can be used to differentiate between mixed states, mania, and bipolar depression. Int J Bipolar Dis 2017;**5**:5.
11. 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 Dis. 2018;**6**:5.
12. Mansour HA, Wood J, Chowdari KV et al. Circadian phase variation in bipolar I disorder. Chronobiol Int 2005;**22**:571–584.
13. Boudebesse C, Geoffroy PA, Henry C et al. Links between sleep and body mass index in bipolar disorders: an exploratory study. European Psychiatry 2015;**30**:89–93.
14. Verkooijen S, Van Bergen AH, Knapen SE et al. An actigraphy study investigating sleep in bipolar I patients, unaffected siblings and controls. J Affect Disord 2017;**208**:248–254.
15. Geoffroy PA, Scott J, Boudebesse C et al. Reply: Sleep in patients with remitted bipolar disorders: analyses stratified on actigraphy devices, age and gender. Acta Psychiatr Scand 2015;**131**:400.
16. 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 Psychiat 2017;**7**:e1211.
17. Gershon A, Kaufmann CN, Depp CA et al. Subjective versus objective evening chronotypes in bipolar disorder. J Affect Disord 2018;**225**:342–349.
18. 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**:863–873.
19. Scott J, Naismith S, Grierson A et al. Sleep-wake cycle phenotypes in young people with familial and non-familial mood disorders. Bipolar Disord 2016;**18**:642–649.
20. Ancoli-Israel S, Martin JL, Blackwell T et al. The SBSM guide to actigraphy monitoring: clinical and research applications. Behav Sleep Med 2015;**13**(Suppl 1):S4–S38.
21. Gonzalez R, Tamminga CA, Tohen M, Suppes T. The relationship between affective state and the rhythmicity of activity in bipolar disorder. J Clin Psychiatry 2014;**75**:e317–e322.
22. 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**:e89574.
23. Hadaeghi F, Hashemi Golpayegani MR, Jafari S, Murray G. Toward a complex system understanding of bipolar disorder: a chaotic model of abnormal circadian activity rhythms in euthymic bipolar disorder. Aust N Z J Psychiatry 2016;**50**:783–792.
24. Melo MCA, Abreu RLC, Linhares Neto VB, De Bruin PFC, De Bruin VMS. Chronotype and circadian rhythm in bipolar disorder: a systematic review. Sleep Med Rev 2017;**34**:46–58.
25. Faurholt-Jepsen M, Brage S, Vinberg M et al. Differences in psychomotor activity in patients suffering from unipolar and bipolar affective disorder in the remitted or mild/moderate depressive state. J Affect Disord 2012;**141**:457–463.
26. Jones SH, Hare DJ, Evershed K. Actigraphic assessment of circadian activity and sleep patterns in bipolar disorder. Bipolar Disord 2005;**7**:176–186.
27. Pagani L, St Clair PA, Teshiba TM et al. Genetic contributions to circadian activity rhythm and sleep pattern phenotypes in pedigrees segregating for severe bipolar disorder. Proc Natl Acad Sci USA 2016;**113**:E754–E761.
28. Salvatore P, Ghidini S, Zita G et al. Circadian activity rhythm abnormalities in ill and recovered bipolar I disorder patients. Bipolar Disord 2008;**10**:256–265.
29. McKenna BS, Drummond SP, Eyler LT. Associations between circadian activity rhythms and functional brain abnormalities among euthymic bipolar patients: a preliminary study. J Affect Disord 2014;**164**:101–106.
30. Lyall LM, Wyse CA, Graham N et al. Association of disrupted circadian rhythmicity with mood disorders, subjective wellbeing, and cognitive function: a cross-sectional study of 91 105 participants from the UK Biobank. Lancet Psychiatry 2018;**5**:507–514.
31. Geoffroy PA, Boudebesse C, Bellivier F et al. Sleep in remitted bipolar disorder: a naturalistic case-control study using actigraphy. J Affect Disord 2014;**158**:1–7.
32. Nurnberger JI Jr, Blehar MC, Kaufmann CA et al. Diagnostic interview for genetic studies Rationale, unique features, and training. NIMH Genetics Initiative. Arch Gen Psychiatry 1994;**51**:849–859; discussion 63-4.
33. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. British Journal of Psychiatry 1979;**134**:382–389.
34. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. Brit J Psychiatry 1978;**133**:429–435.
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**:453–473.
36. Maxwell ME. Family interview for Genetic Studies. Clinical Neurogenetics Branch, Intramural Research Program, NIMH 1992.
37. Netzer NC, Stoohs RA, Netzer CM, Clark K, Strohl KP. Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. Ann Intern Med 1999;**131**:485–491.
38. Millar A, Espie CA, Scott J. The sleep of remitted bipolar outpatients: a controlled naturalistic study using actigraphy. J Affect Disord 2004;**80**:145–153.
39. van Someren EJ, Swaab DF, Colenda CC, Cohen W, McCall WV, Rosenquist PB. Bright light therapy: improved sensitivity to its effects on rest-activity rhythms in Alzheimer patients by application of nonparametric methods. Chronobiol Int 1999;**16**:505–518.
40. Muller KE, Lavange LM, Ramey SL, Ramey CT. Power calculations for general linear multivariate models including repeated measures applications. J Am Stat Assoc 1992;**87**:1209–1226.
41. Huberty C, Olejnik S. Applied MANOVA and discriminant analysis. 2nd ed. New York, NY: Wiley-Interscience; 2006.
42. 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 Dis 2016;**4**:3.

43. Jones SH, Tai S, Evershed K, Knowles R, Bentall R. Early detection of bipolar disorder: a pilot familial high-risk study of parents with bipolar disorder and their adolescent children. Bipolar Disord 2006;**8**:362–372.
44. Zhang J, Paksarian D, Lamers F, Hickie IB, He J, Merikangas KR. Sleep patterns and mental health correlates in US adolescents. J Pediatr 2017;**182**:137–143.
45. Soreca I, Levenson J, Lotz M, Frank E, Kupfer DJ. Sleep apnea risk and clinical correlates in patients with bipolar disorder. Bipolar Disord 2012;**14**:672–676.
46. Robillard R, Oxley C, Hermens DF et al. The relative contributions of psychiatric symptoms and psychotropic medications on the sleep-wake profile of young persons with anxiety, depression and bipolar disorders. Psychiatry Res 2016;**243**:403–406.
47. 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**:351–354.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article:

Figure S1. Receiver Operating Characteristic (ROC) curves for the discriminant function probabilities of BD cases versus healthy controls

Table S1. Raw scores of actigraphy parameters for BD cases and Healthy Controls

Table S2. Additional statistical details for the discriminant function Analyses (DFA)