



Mood, activity, and sleep measured via daily smartphone-based self-monitoring in young patients with newly diagnosed bipolar disorder, their unaffected relatives and healthy control individuals

Sigurd Arne Melbye^{1,2} · Sharleny Stanislaus¹ · Maj Vinberg^{1,2,4} · Mads Frost² · Jakob Eyvind Bardram^{2,3,5} · Kimie Sletved¹ · Klara Coello¹ · Hanne Lie Kjærstad¹ · Ellen Margrethe Christensen¹ · Maria Faurholt-Jepsen¹ · Lars Vedel Kessing^{1,2}

Received: 14 April 2020 / Accepted: 27 July 2020 / Published online: 2 August 2020
© The Author(s) 2020

Abstract

Diagnostic evaluations and early interventions of patients with bipolar disorder (BD) rely on clinical evaluations. Smartphones have been proposed to facilitate continuous and fine-grained self-monitoring of symptoms. The present study aimed to (1) validate daily smartphone-based self-monitored mood, activity, and sleep, against validated questionnaires and clinical ratings in young patients with newly diagnosed BD, unaffected relatives (UR), and healthy controls persons (HC); (2) investigate differences in daily smartphone-based self-monitored mood, activity, and sleep in young patients with newly diagnosed BD, UR, and HC; (3) investigate associations between self-monitored mood and self-monitored activity and sleep, respectively, in young patients with newly diagnosed BD. 105 young patients with newly diagnosed BD, 24 UR and 77 HC self-monitored 2 to 1077 days (median [IQR]=65 [17.5–112.5]). There was a statistically significantly negative association between the mood item on Hamilton Depression Rating Scale (HAMD) and smartphone-based self-monitored mood ($B = -0.76$, 95% CI -0.91 ; -0.63 , $p < 0.001$) and between psychomotor item on HAMD and self-monitored activity ($B = -0.44$, 95% CI -0.63 ; -0.25 , $p < 0.001$). Smartphone-based self-monitored mood differed between young patients with newly diagnosed BD and HC ($p < 0.001$), and between UR and HC ($p = 0.008$) and was positively associated with smartphone-based self-reported activity ($p < 0.001$) and sleep duration ($p < 0.001$). The findings support the potential of smartphone-based self-monitoring of mood and activity as part of a biomarker for young patients with BD and UR. Smartphone-based self-monitored mood is better to discriminate between young patients with newly diagnosed BD and HC, and between UR and HC, compared with smartphone-based activity and sleep.

Trial registration clinicaltrials.gov NCT0288826

Keywords Bipolar disorder · Smartphones · Mood · Activity · Sleep

Background

Bipolar disorder (BD) is a recurrent heterogeneous, and disabling disorder, with a prevalence of 1%, and a heritability of 60–80% [1]. BD is characterized by episodes of mood disturbances with depressive episodes or manic/mixed episodes intervened by euthymic episodes [2]. BD often has an onset of symptoms during young age and may cause substantial functional impairments also before the time of diagnosis, and a progression of severity during years of untreated illness, stressing the need for early diagnosis and intervention [3–5]. An increasing focus has recently been on coexisting alterations in mood, activity and sleep, and loss of function, both during affective episodes, but also between affective

✉ Sigurd Arne Melbye
sigurd.arne.melbye@regionh.dk

¹ The Copenhagen Affective Disorder Research Center (CADIC), Psychiatric Center Copenhagen, Rigshospitalet, 2100 Copenhagen, Denmark

² Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

³ Mosenso ApS, Langelinie Allé 47, Copenhagen, Denmark

⁴ Psychiatric Research Unit, Psychiatric Centre North Zealand, Hillerød, Denmark

⁵ Department of Health Technology, Technical University of Denmark, Kgs. Lyngby, Denmark

episodes [6–8]. Currently, no objective biomarkers like blood tests or radiologic findings are available to diagnose or monitor BD. Thus, symptom monitoring, diagnostic evaluations, and early interventions rely on clinical evaluations typically made with large intervals between outpatient visits, making them vulnerable to recall bias [9] and not capturing daily fluctuations.

Smartphones are owned by 95% of teens in the USA [10], and are always available to most people during naturalistic settings as a natural part of everyday life, especially for young people [11]. Thus, smartphones may facilitate continuous and fine-grained daily self-monitoring of symptoms like mood, activity, and sleep.

Alteration of mood is a central core feature of BD. Many patients with BD experience subsyndromal mood swings on a daily basis that are associated with increased risk of relapse and hospitalization [12–14]. Our group has in recent studies showed that smartphone-based daily self-monitored mood alterations are associated with increased perceived stress and decreased quality of life and functioning [15, 16] and that increased mood instability may behave as a genetic vulnerability trait for BD being present in remitted BD patients [15] and in their unaffected relatives (UR) [16].

Alterations in the level of activity (psychomotor activity) has in the diagnostic and statistical manual of mental disorders (DSM-5) been highlighted as a core symptom similar to mood in BD [17], and elevated activity levels are associated with (hypo)mania, and reduced activity levels are associated with depression [18]. Furthermore, regardless the affective state, patients with newly diagnosed BD have been shown to have lower activity levels compared with healthy individuals [19, 20]. Several studies on adult patients with BD have found smartphone-based self-monitoring of activity feasible [21–23]. A recent study from our group found smartphone-based self-monitored activity to be statistically significantly associated with activity-related sub-items in the Hamilton Depression Rating Scale (HAMD) [24] and the Young Mania Rating Scale (YMRS) [25, 26].

(Hypo)manic episodes are often associated with a substantial reduced sleep duration, while depressive episodes often are associated with sleep disturbances either as a reduced sleep duration or an increased sleep duration [27]. Sleep assessment is often based on prospective reporting in questionnaires or clinical interviews and is, therefore, prone to recall bias [28]. Thus, a tool for a ubiquitous self-monitoring of sleep might eliminate the risk of recall bias. As sleep disturbance is a key symptom in the beginning of, or preceding an affective episode [29], continuous monitoring of sleep can facilitate early identification and intervention of relapsing affective episodes [30].

First-degree relatives to patients with BD have an increased risk of developing psychiatric disorders [31]. Results from a prior study suggest that 20–30% of healthy

first-degree relatives of patients with BD will develop affective illness, compared to 2–5% among healthy individuals [32]. Therefore, it is likely that UR to patients with BD will show alterations in mood, activity, and sleep intermediate between young patients with BD and healthy controls (HC), maybe as an expression of prodromal symptoms [32]. Additionally, if smartphone-based self-monitoring is able to identify subsyndromal symptoms in UR, it might have the potential to identify prodromal symptoms in relation to relapse of affective episodes in patients with BD [33], and also identify early symptoms in individuals undergoing diagnostic assessment for BD. Smartphone-based self-monitoring has been tested as a screening-tool for BD in a community sample, with promising results [34]. Also, one study showed that 38% of participants showing positive results in a smartphone-based screening for depression, consulted health care professionals about their results [35].

Smartphone-based self-monitored data have shown to validly reflect illness activity in patients with BD [15, 21, 36–38]. Thus, smartphone-based self-monitoring may improve the timeliness in monitoring symptoms, functioning, and fluctuations in symptoms, and might, therefore, be a promising supportive diagnostic tool in individuals where BD is suspected.

A recent review, conducted by the authors of the present study, showed that smartphone-based self-monitoring in young people with psychiatric disorders generally has both high adherence and acceptability [39]. More specifically, a recent study on electronic ecological momentary assessment in young patients with BD showed adherence rates of 80.4% among young patients with BD [40]. Also, prior research has found smartphone monitoring to be useful, easy to use, and feasible for adult patients with BD [21, 38, 41, 42]. However, no studies on smartphone-based monitoring of young patients with newly diagnosed BD, UR, and HC have been conducted.

Aims

The present study aimed to (1) validate daily smartphone-based self-monitored mood, activity, and sleep, against validated questionnaires and clinical ratings in young patients with newly diagnosed BD, UR, and HC; (2) investigate differences in daily smartphone-based self-monitored mood, activity, and sleep in young patients with newly diagnosed BD, UR, and HC; (3) investigate associations between self-monitored mood and self-monitored activity and sleep, respectively in young patients with newly diagnosed BD.

We hypothesized that (1) daily smartphone-based self-monitored mood, activity, and sleep would be associated with validated and standardized questionnaires and clinical ratings, (2) daily smartphone-based self-monitored mood,

activity, and sleep differs between young patients with newly diagnosed BD compared to UR and HC with UR being an intermediary group; (3) daily smartphone-based self-monitored mood would be positively associated with activity and negatively associated with sleep duration in young patients with newly diagnosed BD.

Methods

The present study included data collected as part of the Bipolar Illness Onset study (the BIO study) [43], which is an ongoing longitudinal observational study including patients with newly diagnosed BD, their UR and HC [44, 45].

Study design, participants, and settings

As part of the BIO study, we recruited young patients, from 18 to 25 years old, with newly diagnosed BD, their UR, and HC without a psychiatric family history. In the BIO study, all participants are assessed clinically at baseline and annually, combined with blood tests, and at some visits MRI-scan and cognitive tests in addition to annual visits. The young patients with newly diagnosed BD were recruited from the Affective Disorder Clinic, Rigshospitalet, Copenhagen, Denmark. The Affective Disorders Clinic is a highly specialized clinic offering a 2-year course of treatment to every individual newly diagnosed with BD in the larger region of Copenhagen. We also included some younger patients with newly diagnosed BD, aged 15–17, from the Child and Adolescent Mental Health Center in Copenhagen. For the UR group, we included among the first-degree relatives of the patients included in the study. Inclusion criteria for UR were no history of a psychiatric disorder requiring treatment. HC were recruited among blood donors from the Blood Bank at Rigshospitalet, Copenhagen. Inclusion criteria were no history of a psychiatric disorder requiring treatment, personally or in a first-degree relative, and 25 years old or younger at the time of inclusion. All participants signed an informed consent form before starting in the project; also for participants under the age of 18, we obtained informed consent from one or both legal parents.

Diagnostic assessment

To ensure that participants fulfilled the inclusion criteria to participate in the designated groups, all participants underwent a diagnostic interview using Schedules for Clinical Assessment in Neuropsychiatry (SCAN) [46], during the baseline interview. Trained researchers performed the SCAN interviews; for the BD and UR groups, baseline interviews were performed by Ph.D students in medicine or psychology,

whereas the HC were either examined by the same Ph.D students, or medical or psychology students.

Baseline interview and follow-up

All participants attended a baseline interview with collection of general information about work status, educational status, in addition to diagnostic and clinical assessment. For young patients with newly diagnosed BD, we collected information about time for onset of symptoms, diagnosis, start of treatment, and number and duration of affective episodes. All participants also completed a number of questionnaires. After baseline, participants were asked to attend an annual follow-up interview for a maximum of 5 years. For young patients with newly diagnosed BD, we endeavored to arrange interviews with the participants each time they had a change from one affective episode to another, from an affective episode to euthymic stage, or from a euthymic stage to an affective episode.

Clinical ratings

At baseline and follow-up, the following clinical assessments were conducted: a clinical assessment of the severity of depressive symptoms was performed using the Hamilton Depression Rating Scale 17-item (HAMD) [24, 47], and of (hypo-)manic symptoms using the Young Mania Rating Scale (YMRS) [26]. From the HAMD, we included the total scores from HAMD 17-item and 6-item as well as the scores on sub-item 1 addressing level of decreased mood, sub-item 4–6 addressing sleep disturbances, sub-item 8 addressing the level of psychomotor retardation, and sub-item 9 addressing the level of psychomotor agitation, for the past 3 days. From the YMRS, we included the total score as well as the scores from sub-item 1 addressing the level of elevated mood, sub-item 2 addressing the level of increased motor activity, sub-item 4 reflecting reduced sleep duration, and sub-item 6 addressing increased talkativeness, for the past 3 days.

Further, a clinical assessment of functioning was performed using the Functioning Assessment Short Test (FAST) [48], which is a test developed explicitly for BD and addresses six areas of functioning: autonomy, occupational functioning, cognitive functioning, financial issues, interpersonal relationship, and leisure time. All 24 items are rated from 0 (no difficulties) to 3 (severe difficulties) and assess the last 2 weeks up to the rating. FAST has a high test–retest reliability and has been validated against the Global assessment of functioning scale (GAF) [48].

Questionnaires

All participants completed several questionnaires at baseline and follow-up. The questionnaires used for analyses in this

article were the following: severity of self-assessed depressive symptoms according to the Major Depression Inventory (MDI) with scores on the MDI ranging from 0 to 50, where a score below 20 indicates no depression and a score over 29 indicates severe depression [49]; severity of self-assessed manic symptoms according to the Altman Self-Rating Mania Scale (ASRM), with scores on the ASRM ranging from 0 to 20, where a score of 6 or higher indicates a high probability of hypomanic or manic state [50]; self-assessed physical activity according to the International Physical Activity Questionnaire (IPAQ), where participants report how many minutes of physical activity in different intensities they had the prior week, these are then converted to metabolic equivalent task (MET) minutes per week, which then adds up to the total MET score [51]; and self-assessed sleep quality according to the Pittsburgh Sleep Quality Index (PSQI), with scores on the PSQI ranging from 0 to 21, where a score over 5 indicates sleep disturbances [52]. These questionnaires were chosen as they reflect the level of mood alteration, activity and sleep quality for the participants. We did not include other questionnaires in the analyses, and all hypotheses and analyses were specified in advance.

Smartphone-based self-monitoring

The Monsenso system is a smartphone-based self-monitoring system developed by the authors allowing for daily self-monitoring of mood, sleep, and activity [53, 54]. The Monsenso application also collect automatically generated data on phone usage, social activity (usage of calls and text messages) and physical activity; however, only self-monitored data are investigated in this article. The system is available for both Android and iOS; however, collection of automatically generated data is only for Android phones. These automatically generated data were the only ones which were collected due to technical practicalities.

At the time of inclusion, every participant was registered in the Monsenso system and instructed to start using the system 3 days before the baseline interview. Participants installed the app on their phone, and participants not owning a phone compatible with Monsenso were offered to borrow an Android smartphone (LG Nexus 5). Young patients with newly diagnosed BD were asked to use the system for at least 3 months, and the remaining participants were asked to use the system for at least 1 month. Participants were not instructed on what time of the day to perform the self-monitoring. When participants were booked for follow-up interviews with clinical assessments, they were encouraged to start using the system again.

When installed on the phone, the application prompted the user to fill out the self-monitoring once a day. As default, prompts were set for 8 pm daily, but participants could set their preferred time for prompts, or deactivate

prompts completely. When prompted (or reminded by memory), the participant opened the application and completed as many features of the self-assessment as wanted. Figure 1 shows screenshots of the interface. The self-monitoring took 2–4 min each day.

For self-monitoring of mood, young patients with newly diagnosed BD, UR and HC scored how their days had been on a 7-point-scale (−3, −2, −1, 0, 1, 2, 3). In this way, all participants daily scored how good or how bad their day had been. Additionally, young patients with newly diagnosed BD scored their mood on a 9-point scale with scores from depressive to manic (−3, −2, −1, −0.5, 0, 0.5, 1, 2, 3), as a more fine-grained scale has been found useful for individuals with BD [53]. For self-assessment of activity, participants scored their daily level of activity on a 7-point-scale (−3, −2, −1, 0, 1, 2, 3). All participants also registered their bedtime and the time they got out of bed. Self-monitoring of mood, activity and sleep were part of the default interface in the application, and the only features included in the analyses; however, it was possible to add other prefabricated features, or even create custom made features.

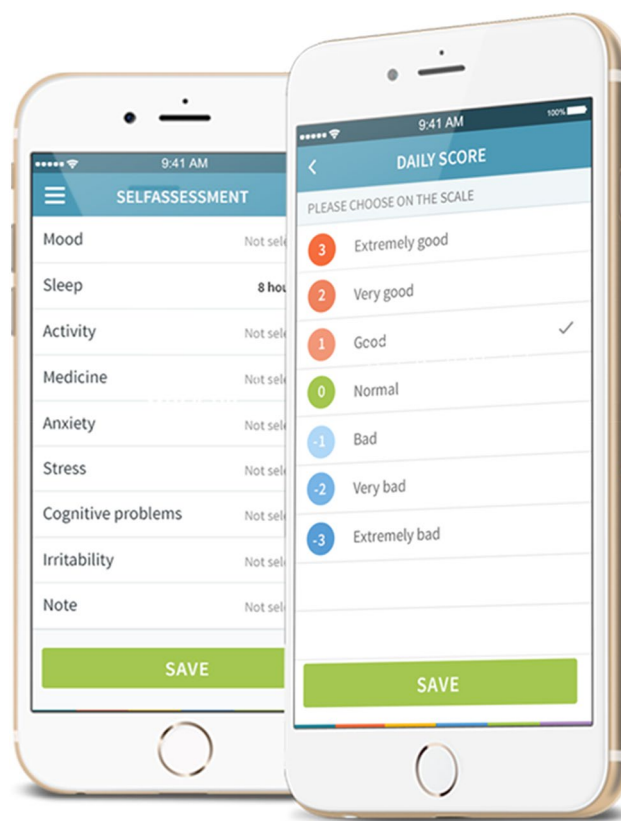


Fig. 1 Screenshots from the Monsenso interface

Statistical analyses

The hypotheses and statistical analyses for the present study were defined a priori. For background characteristics, linear mixed effect models were used for analyzing the differences in the mean of continuous variables between the three groups. For between-group differences of categorical data, Chi-square tests were conducted. Concerning analyses for aims 1–3, for each measure of interest, a linear mixed effect model, which accommodates both the variation of the variables of interest within young patients with newly diagnosed BD (intra-individual variation) and between individuals (inter-individual variation) was employed. Because the data used in this article are part of a larger still ongoing longitudinal study [43], we had data from both baseline and, for some participants, up to a maximum of nine follow-ups. Thus, we used a linear mixed effect model analysis to account for repeated measurements within each participant. As mixed effect models are more robust in dealing with missing data in repeated measurements, we could include all data from all of our participants, regardless of large differences in the amount of data from each participant. We identified all participants with a unique identification (ID) number and a family number. Each patient was given a family number, and if they had a relative in the study, they were given the same family number as the patient. Thus, the familial relationship was accounted for by adding family number as a random factor in the linear mixed effect model. For aim (1), we used a linear mixed effect model to analyze the association between smartphone-based self-assessed scores and scores from clinical rating scales and questionnaires. For analyses on associations between smartphone-based monitoring and clinical ratings and questionnaires averages of smartphone-based score ratings were calculated for the current day and 3 days before ratings with the HAMD and the YMRS, 7 days prior for the ASRM, and 14 days prior for the MDI and the FAST, and 30 days prior for the PSQI. For aim (2), we used a linear mixed effect models to compare means between smartphone-based self-assessments scores, scores from clinical ratings, and scores from questionnaires.

For aim (3), we used a linear mixed effect model to analyze the associations between smartphone-based self-assessed mood and smartphone-based self-assessed activity and sleep, respectively. For all analyses, we added individual ID number and family number as random factors.

For aim (1) and (2), we analyzed pooled data from all participants; for aim (3), we analyzed data on young patients with newly diagnosed, only. Analyses were conducted in models adjusted for age and sex.

As few prior studies have investigated differences in daily smartphone-based self-monitored symptoms between young patients with newly diagnosed BD, UR, and HC, statistical power analyses prior to the study were not performed. Data

were collected as part of a larger longitudinal observational study, and thus the sample size was defined according to this. Model assumptions were checked visually by means of residuals and QQ plots for each of the statistical analyses. We used SPSS version 25 (Statistical Package for the Social Sciences) for all analyses. p values ≤ 0.05 (two sided) were considered statistically significant.

Results

Demographic and clinical characteristics

A total of 401 young patients with newly diagnosed BD, 92 UR, and 200 HC from September 2016 to November 2019 were included in the BIO study. Of these, a total of 105 young patients with newly diagnosed BD, 24 UR, and 77 HC provided smartphone-based self-monitoring data and were 25 years old or younger by the time of inclusion and thus eligible to be included in the analyses for the present study. There were 3 pairs of relatives among the 23 UR. The rate of participants taking part of the study after talking to one of the researchers was very high ($< 85\%$); however, these were already screened by clinicians regarding interest in participating. Further background characteristics are presented in Table 1.

Follow-up and smartphone-based self-assessment

There were a total number of 342 visits with clinical assessment among the participants, and for 262 of these visits, the participants provided daily smartphone-based self-assessment. The number of visits for each participant varied from one to nine visits. The number of days where participants performed self-assessment varied from 2 to 1077 days (median [IQR] = 65 [17.5–112.5]) (Fig. 2).

The validity of smartphone-based self-monitored mood, activity, and sleep compared with clinical ratings and questionnaires

Mood

As shown in Table 2, there was a statistically significant negative association between self-monitored mood and scores on subitem 1 of the HAMD ($B = -0.76$, 95% CI -0.90 ; -0.63 , $p < 0.001$); as well as the total score of HAMD-17 ($B = -0.11$, 95% CI -0.12 ; -0.09 , $p < 0.001$) and HAMD-6 rating scale ($B = -0.18$, 95% CI -0.21 ; -0.15 , $p < 0.001$). Thus, for every decrease of 1 point on the HAMD sub-item 1 score, there was a 0.76 increase on the smartphone-based self-monitoring of mood. There was no statistically significant association between self-monitored mood and the

Table 1 Demographics and background characteristics of young patients with newly diagnosed bipolar disorder (BD), unaffected relatives (UR), and healthy control individuals (HC); $N=206$

	BD	UR	HC	BD vs UR <i>p</i>	BD vs HC <i>p</i>	UR vs HC <i>p</i>
Participants	105	24	77			
Age, years	21.74 (21.28; 22.21)	23.08 (21.40; 24.77)	22.87 (22.45; 23.29)	<i>0.012</i>	<i>0.002</i>	0.755
Female, % (<i>n</i>)	75.2 (79)	60.9 (14)	76.6 (59)	0.205	0.829	0.172
Education, years	13 [11.5–14.5]	15 [13–17]	15 [11.5–14.5]	<i>0.004</i>	<i><0.001</i>	0.404
Student, % (<i>n</i>)	60.9 (64)	50 (12)	61 (47)	0.325	0.991	0.338
Full time employment, % (<i>n</i>)	12.4 (13)	37.5 (9)	13 (10)	<i>0.003</i>	0.903	<i>0.007</i>
HAMD-17	8 [2.5–13.5]	1.5 [0–5.5]	0 [0–2]	<i><0.001</i>	<i><0.001</i>	0.107
YMRS	3 [0–7.5]	0 [0–2]	0 [0–1]	<i><0.001</i>	<i><0.001</i>	0.374
FAST, total score	19 [8.5–29.5]	2 [0–5]	1 [0–2]	<i><0.001</i>	<i><0.001</i>	0.079
Bipolar disorder subtype II, % (<i>n</i>)	72.4 (76)	–	–	–	–	–
Age of illness onset, years	15 [13–17]	–	–	–	–	–
Illness duration, years ^a	6 [3–9]	–	–	–	–	–
Number of years untreated ^b	3 [0.5–5.5]	–	–	–	–	–
No. of depressive episodes	5 [1.5–8.5]	–	–	–	–	–
No. of hypomanic episodes	4 [0–8]	–	–	–	–	–
No. of manic episodes	0 [0–1]	–	–	–	–	–
No. of mixed episodes	0 [0–0]	–	–	–	–	–
No. of total episodes	12 [5.5–18.5]	–	–	–	–	–

Continuous variables are presented as median [interquartile range] or mean (SD) and *p* values are calculated based on differences in the mean between the three groups using linear mixed models. Categorical data are presented as % (*n*), and *p* values are calculated using the Chi-square test

HAMD-17: Hamilton Depression Rating Scale with 17 items version; YMRS: Young Mania Rating Scale; FAST: Functional Assessment Short Test

p values < 0.05 are presented in italic to emphasize statistical significance

^aIllness duration was defined as the time from the first episode to the time of inclusion

^bUntreated BD was defined as the time from the first mania, hypomania, or mixed episode to time of diagnosis

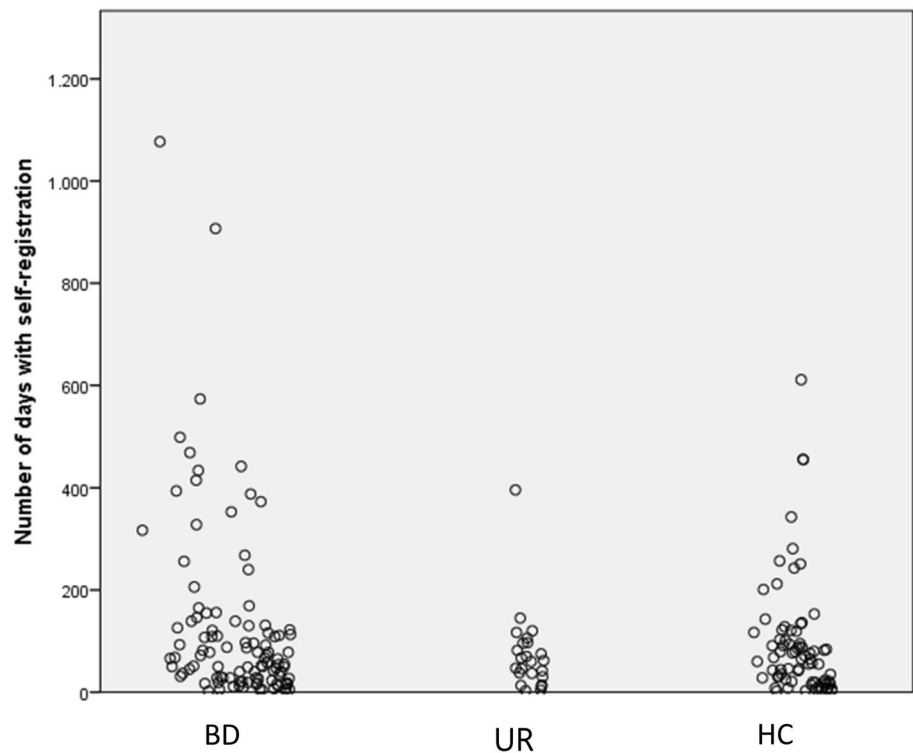
YMRS, neither for subitem 1 or total score. For questionnaires, there was a statistically significant negative association between smartphone-based self-monitored mood and scores on the MDI ($B=-0.05$, 95% CI -0.06 ; -0.04 , $p<0.001$), and a statistically significant positive association between smartphone-based self-monitored mood and scores on the ASRM ($B=0.06$, 95% CI 0.02 ; 0.10 , $p=0.001$). Thus, for every decrease of 1 point on the MDI total score, there was a 0.05 increase on the smartphone-based self-monitoring of mood. For every increase of 1 point on the ASRM score, there was a 0.06 increase on the smartphone-based self-monitoring of mood.

Activity

There was a statistically significantly negatively association between smartphone-based self-monitored activity and the HAMD subitem 8 ($B=-0.44$, 95% CI -0.63 ; -0.25 , $p<0.001$) and 9 ($B=-0.37$, 95% CI -0.58 ; -0.15 , $p=0.001$), and total scores on the FAST ($B=-0.02$, 95% CI -0.03 ; -0.01 , $p<0.001$). Thus, for every increase of

1 point on the HAMD sub-item 8 score, there was a 0.44 decrease on the smartphone-based self-monitoring of activity, for every increase of 1 point on the HAMD sub-item 9 score, there was a 0.37 decrease on the smartphone-based self-monitoring of activity, and for every 1 point increase on the FAST total score, there was a 0.02 decrease on the smartphone-based self-monitoring of activity. Smartphone-based self-monitored activity was significantly positively associated with the YMRS subitem 2 ($B=0.21$, 95% CI 0.02 ; -0.39 , $p=0.028$). Thus, for every increase of 1 point on the YMRS sub-item 2 score, there was a 0.21 increase on the smartphone-based self-monitoring of activity. There was no statistically significant association between smartphone-based self-monitored activity and the YMRS subitem 6. There was a statistically significant positive association between self-monitored activity and total score for the IPAQ ($B=3.94E-5$, 95% CI $1.09E-5$; $6.79E-5$, $p=0.007$).

Fig. 2 The number of days with smartphone-based self-monitoring in young patients with newly diagnosed bipolar disorder (BD), unaffected relatives (UR), and healthy control individuals (HC). Each circle represents a participant



Sleep

There was no statistically significant association between smartphone-based self-monitored sleep duration and scores on the HAMD subitems 4, 5, or 6. Nevertheless, there was a statistically significant negative association between smartphone-based self-monitored sleep duration and the YMRS subitem 4 ($B = -29.65$, 95% CI -43.51 ; -15.80 , $p < 0.001$), as well as between smartphone-based self-monitored sleep duration and total scores on the PSQI ($B = -2.94$, 95% CI -5.07 ; -0.82 , $p = 0.007$). Thus, for every increase of 1 point on the YMRS sub-item 4 score, there was a 29.65 min decrease on the smartphone-based self-monitoring of sleep duration, and for every increase of 1 on the PSQI total score, there was a 2.94 min increase on the smartphone-based self-monitoring of sleep duration.

Differences in daily smartphone-based self-monitoring of mood, activity, and sleep between young patients with newly diagnosed bipolar disorder, unaffected relatives and healthy controls

Table 3 shows the differences in means of scores of smartphone-based self-monitored mood, activity and sleep, and some of the clinical ratings between young patients with newly diagnosed BD, UR and HC. There was no statistically significant difference in smartphone-based self-monitored

mood between young patients with newly diagnosed BD and UR ($p = 0.107$). There was a statistically significant difference between young patients with newly diagnosed BD and HC ($p < 0.001$), and between UR and HC ($p = 0.008$). For smartphone-based self-monitored activity, there was no statistically significant difference between any of the groups. For smartphone-based self-monitored sleep duration, there was a statistically significant difference between young patients with newly diagnosed BD and HC ($p = 0.003$), and no other statistically significant between-group difference.

There were, however, statistically significant differences between groups with clinical ratings and questionnaires, mainly between BD and UR and between BD and HC.

Association between smartphone-based self-monitored mood, activity, and sleep

Table 4 shows the associations between smartphone-based self-monitored mood, activity and sleep duration, respectively, in young patients with newly diagnosed BD. Analyses both with 7-point mood scale and the finer grained 9-point mood scale (both ranging from -3 to 3) were employed. There was a statistically significant positive association between smartphone-based self-monitored mood and activity for both the 7-point ($B = 0.37$, 95% CI 0.36 ; 0.39 , $p < 0.001$) scale and 9-point scale ($B = 0.16$, 95% CI 0.15 ; 0.17 , $p < 0.001$). Thus, for every increase of 1 point on the smartphone-based self-monitoring of activity, there was

Table 2 Associations between smartphone-based self-monitored mood, activity and sleep (dependent variables), and observer-based rating scales for depression, mania, functioning, and questionnaires for sleep and physical activity (data pooled for young patients with newly diagnosed bipolar disorder, unaffected relatives, and healthy control individuals), $N=190$

	<i>B</i>	95% CI	<i>p</i>
Smartphone-based self-monitored mood			
HAMD-17 total ^a	-0.11	-0.12; -0.09	<i><0.001</i>
HAMD-6 total ^b	-0.18	-0.21; -0.15	<i><0.001</i>
HAMD sub-item 1 ^c	-0.76	-0.90; -0.63	<i><0.001</i>
YMRS total ^d	-0.02	-0.05; 0.02	0.277
YMRS sub-item 1 ^e	0.16	-0.04; 0.35	0.112
FAST ^f	-0.05	-0.05; -0.04	<i><0.001</i>
MDI ^g	-0.05	-0.06; -0.04	<i><0.001</i>
ASRM ^h	0.06	0.02; 0.10	0.001
Smartphone-based self-monitored activity			
HAMD-17 sub-item 8 ⁱ	-0.44	-0.63; -0.25	<i><0.001</i>
HAMD-17 sub-item 9 ^j	-0.37	-0.58; -0.15	0.001
YMRS sub-item 2 ^k	0.21	0.02; 0.39	0.028
YMRS sub-item 6 ^l	0.05	-0.08; 0.19	0.429
FAST ^f	-0.02	-0.03; -0.01	<i><0.001</i>
IPAQ ^m	3.94E-5	1.09E-5; 6.79E-5	0.007
Smartphone-based self-monitored sleep duration			
HAMD-17 sub-item 4 ⁿ	-8.72	-18.365; 0.917	0.076
HAMD-17 sub-item 5 ^o	-3.37	-13.616; 6.867	0.517
HAMD-17 sub-item 6 ^p	-0.89	-13.035; 11.260	0.886
YMRS sub-item 4 ^q	-29.65	-43.514; -15.793	<i><0.001</i>
FAST ^f	0.377	-0.195; 0.949	0.196
PSQI ^r	-2.94	-5.065; -0.819	0.007

Model adjusted for age and gender

p values <0.05 are presented in italic to emphasize statistical significance

^aHAMD-17: The Hamilton Depression Rating Scale 17 total score

^bHAMD-6 Hamilton depression Rating scale 6 total score

^cHAMD sub-item 1—level of decreased mood

^dYMRS: The Young Mania Rating Scale total score

^eYMRS sub-item 1—level of elevated mood

^fFAST: The Functional Assessment Short Test total score

^gMDI: The Major depression Inventory total score

^hASRM: The Altman Self-Rating Mania Scale total score

ⁱHAMD sub-item 8—level of psychomotor retardation

^jHAMD sub-item 9—level of psychomotor agitation

^kYMRS sub-item 2—level of increased motor activity

^lYMRS sub-item 6—increased talkativeness

^mIPAQ: The Physical Activity Questionnaire—short form total score

ⁿHAMD sub-item 4—problems falling asleep

^oHAMD sub-item 5—problems with mid-sleep—wake-ups

^pHAMD sub-item 6—problems with early morning awakening

^qYMRS Sub-item 4—reduced sleep duration

^rPSQI: The Pittsburgh Sleep Quality Index total score

a 0.37 increase on self-reported mood on the 7-point scale, and a 0.16 increase on self-reported mood on the 9-point scale. There was a statistically significant negative association between smartphone-based self-monitored mood and self-monitored sleep duration for the 9-point mood scale ($B=-0.0007$, 95% CI -0.0008 ; -0.0006 , $p<0.001$), and for the 7-point mood scale ($B=-0.0007$, 95% CI -0.0009 ; -0.0005 , $p<0.001$). Thus, for every increase of 1 min on self-reported sleep duration, there was 0.0007 decrease on smartphone-based self-monitoring of mood.

Discussion

This observational study is the first to investigate smartphone self-monitoring of mood, activity, and sleep duration in young patients with newly diagnosed BD, UR and HC. The study included 206 participants who used the Monsenso smartphone application for self-monitoring for up to 3 years, in addition to multiple clinical assessments.

Our results suggest that smartphone-based self-monitoring is a valid presentation of mood in young patients with depressive symptoms, and self-monitored activity gives a valid presentation of the level of activity detected in clinical ratings.

Aim 1: In accordance with our hypotheses, we found a statistically significant association between smartphone-based self-monitored mood and scores on the mood item 1 and the total HAMD. This finding is in line with findings from similar studies on adults with BD [15, 21, 36, 37], and indicates that smartphone-based self-monitoring can identify depressive symptoms and thus function as a fine-grained monitoring tool for young patients with BD, and as a potential diagnostic tool for patients where BD or unipolar disorder is suspected. There was no significant association between self-monitored mood and scores on the YMRS. This may be due to a low number of observations of young patients with newly diagnosed BD clinically assessed during a manic episode, and due to UR and HC rating high on self-reported mood, without having manic symptoms. Thus, the latter results should be interpreted with caution.

Also, in accordance with our hypotheses, self-monitored activity was associated with the HAMD subitems concerning activity and with one of two activity related YMRS items, showing that this monitoring feature also may serve as a monitoring tool for young patients with BD. There was a significant association between self-monitored sleep duration and the sleep-related YMRS subitem; however, no significant correlations between self-monitored sleep duration and sleep-related HAMD subitems. A reason for this can be that depression is associated with both increased and decreased sleep duration, and that self-monitoring of

Table 3 Estimated differences in mood, activity, and sleep indexes between young patients with newly diagnosed bipolar disorder (BD, $n = 105$), unaffected relatives (UR, $n = 24$) and control individuals (HC, $n = 77$), $N = 206$

	BD		UR		HC		BD/UR <i>p</i>	BD/HC <i>p</i>	UR/HC <i>P</i>
	Mean	95% CI	Mean	95% CI	Mean	95% CI			
Smartphone-based self-monitored mood									
Mood ^a	0.41	0.15; 0.68	0.89	0.35; 1.44	1.71	1.42; 1.10	0.107	<0.001	0.008
Mood measures in HAMD and YMRS									
HAMD subitem 1 ^b	0.63	0.51; 0.74	0.06	-0.17; 0.28	0.06	-0.07; 0.20	<0.001	<0.001	0.982
YMRS subitem 1 ^c	0.55	0.44; 0.66	0.17	-0.05; 0.39	0.22	0.09; 0.35	0.003	<0.001	0.694
Total score in MDI and ASRM									
MDI ^d	20.28	18.59; 21.97	9.76	6.54; 12.97	5.50	3.55; 7.45	<0.001	<0.001	0.023
ASRM ^e	4.13	3.50; 4.75	2.61	1.43; 3.80	2.94	2.23; 3.66	0.024	0.010	0.631
Smartphone-based self-monitored activity									
Activity score ^f	0.41	0.14; 0.67	0.76	0.24; 1.27	0.70	0.41; 0.98	0.226	0.124	0.823
Activity measures on HAMD and YMRS									
HAMD sub-item 8 ^g	0.35	0.26; 0.43	0.01	-0.15; 0.18	0.00	-0.10; 0.10	<0.001	<0.001	0.894
HAMD sub-item 9 ^h	0.34	0.26; 0.42	0.17	0.01; 0.32	0.02	-0.07; 0.11	0.048	<0.001	0.111
YMRS sub-item 2 ⁱ	0.48	0.36; 0.60	0.10	-0.14; 0.33	0.11	-0.03; 0.25	0.004	<0.001	0.903
YMRS sub-item 6 ^j	0.78	0.63; 0.94	0.04	-0.27; 0.33	0.10	-0.08; 0.28	<0.001	<0.001	0.697
International Physical Activity Questionnaire									
FAST total score ^k	18.02	16.30; 19.73	3.55	0.18; 6.93	1.12	-0.85; 3.09	<0.001	<0.001	0.209
Functional Assessment Short Test									
IPAQ total score ^l	3416	2664; 4168	2725	1381; 4069	3415	2593; 4236	0.356	0.997	0.377
Smartphone-based self-monitored sleep									
Sleep duration (hours:min) ^m	6:48	6:27; 7:09	7:04	6:23; 7:44	7:31	7:09; 7:54	0.489	0.003	0.225
Sleep measures on HAMD ^f and YMRS									
HAMD sub-item 4–6 ⁿ	2.45	2.18; 2.72	1.03	0.51; 1.55	0.69	0.38; 0.10	<0.001	<0.001	0.251
YMRS sub-item 4 ^o	0.46	0.35; 0.58	0.11	-0.11; 0.33	0.08	-0.05; 0.21	0.005	<0.001	0.805
Pittsburgh sleep Quality Index									
Total PSQI score ^p	8.66	8.09; 9.22	5.42	4.41; 6.43	4.38	3.76; 4.99	<0.001	<0.001	0.075

Adjusted for age and sex

p values < 0.05 are presented in italic to emphasize statistical significance^aSmartphone-based self-monitored mood: Averages of smartphone-based day score^bHAMD sub-item 1—level of decreased mood^cYMRS sub-item 1—level of elevated mood^dMDI: The Major depression Inventory total score^eASRM: The Altman Self-Rating Mania Scale total score^fSmartphone-based self-monitored activity: averages of smartphone-based activity rating^gHAMD sub-item 8—level of psychomotor retardation^hHAMD sub-item 9—level of psychomotor agitationⁱYMRS sub-item 2—level of increased motor activity^jYMRS sub-item 6—increased talkativeness^kFAST: The Functional Assessment Short Test total score^lIPAQ: The Physical Activity Questionnaire—short form total score^mSmartphone-based self-monitored sleep duration: Averages of smartphone-based sleep length (minutes from bedtime to time they got out of bed)ⁿHAMD sub-item 4–6—sum of scores regarding problems with sleep^oYMRS sub-item 4—reduced amount of sleep last 3 days^pPSQI: The Pittsburgh Sleep Quality Index total score

Table 4 Associations between smartphone-based self-monitored mood (dependent variable), and smartphone-based self-monitored sleep and activity, respectively, for young patients with newly diagnosed bipolar disorder ($N=105$)

	<i>B</i>	95% CI	<i>p</i>
Smartphone-based self-monitored mood (7-point scale) ^a			
Smartphone-based self-monitored activity ^b	0.37	0.36; 0.39	< 0.001
Smartphone-based self-monitored sleep duration ^c	−0.0007	−0.0009; −0.0005	< 0.001
Smartphone-based self-monitored mood (9-point scale) ^d			
Smartphone-based self-monitored activity	0.16	0.15; 0.17	< 0.001
Smartphone-based self-monitored sleep duration	−0.0007	−0.0008; −0.0006	< 0.001

Adjusted for age and gender

p values < 0.05 are presented in italic to emphasize statistical significance

^aSmartphone-based self-monitored mood rated on a 7-point scale from −3 to +3

^bSmartphone-based self-monitored activity rated on a 7-point scale from −3 to +3

^cSmartphone-based self-monitored sleep duration reported in minutes from bedtime to the time they got out of bed

^dSmartphone-based self-monitored mood rated on a 9-point scale from −3 to +3

sleep needs more detailed features like e.g., number of times woken up, quality of sleep etc.

Aim 2: For self-monitored mood, there was a statistically significant difference both between UR and HC, and between young patients with newly diagnosed BD and HC, which indicates that UR can have intermediary symptoms. Nevertheless, there was no significant difference between young patients with newly diagnosed BD and UR. The overall findings suggests that the self-monitoring can be used to identify subsyndromal symptoms, which can be very useful as a diagnostic tool and in early intervention, as it potentially can identify prodromal symptoms which sometimes patients with BD have difficulties identifying themselves [33].

Aim 3: The significant positive correlation between self-monitored mood and self-monitored activity and between self-monitored mood and self-monitored sleep duration in young patients with newly diagnosed BD, supports the clinically known associations between mood, and activity and sleep, respectively, and thus suggest that these features may constitute a part of an electronic biomarker, as it also gives more insight to the patients complex distribution of symptoms, potentially identifying therapeutic targets [55].

Limitations

First, in addition to the smartphone-based self-monitoring data, participants in the present study provided blood samples, cognitive tests, and MRI as part of the BIO study, making smartphone-based self-monitoring an additional part of the study participation. This might have resulted in a lower adherence, compared to if the study focused on smartphone-based monitoring only. Thus, we saw a big variety in adherence reflecting the difference in the participant's dedication for the smartphone-based part of the BIO study. Second, the number of UR in this study was rather small, partly due to

the criteria of no prior history of mental disorders and partly due to patients with BD not approving us to contact their relatives. Third, it was challenging to get the young patients with newly diagnosed BD to revisit for clinical assessment during more severe affective episodes. The Monsenso system, also collects automatically generated data which gives an even more fine-grained and objective information regarding activity and sleep; however, this is only available for Android phones. As the share of Android users among our participants in the present study was low, the number of UR contributing with automatically generated data was very low to include these data in the analysis; also, the focus of this article was on self-monitored data.

Our population consisted of 74% female, and among the young patients with newly diagnosed BD, 72% had BD type 2. However, we believe that our results are generalizable to young patients with newly diagnosed BD in general. Lastly, our HC were recruited among blood donors, which might represent a 'super healthy' population [17], although other and more suited control groups are difficult to identify and recruit.

Strengths

The present study systematically recruited 206 young people comprising young patients with newly diagnosed BD, their UR, and HC. Also, the young patients with newly diagnosed BD were diagnosed at a mood disorder clinic and the diagnosis, or lack of diagnoses were verified for all participants with a SCAN-interview conducted by trained MD and M.Sc. psychology Ph.D student assessors. Furthermore, participants were clinically assessed using clinically validated observer-based rating scales like the HAMD, the YMRS, and the FAST. The smartphone-based self-monitoring system (the Monsenso system) used in the present study is well

validated, useful and fulfills safety of data storage and privacy requirements.

Perspectives

The emerging research on smartphone-based monitoring of psychiatric symptoms in young people, has shown this technology to be both feasible and acceptable [39], and our findings suggest that it also produce valid clinical information in young patients with BD and potentially UR. Future studies should investigate the use of automatically generated data, which supplements with more continuous, fine-grained, objective, and real-time clinical information. This provides unique possibilities for a fine-grained real-time monitoring tool that can give clinicians an extended insight into the symptoms of patients and UR and level of function over time and facilitate targeted intervention. Additionally, valid identification of symptoms may improve diagnosing and may result in patients getting the correct diagnoses with less latency.

Conclusions

Smartphone-based self-monitoring of mood and activity represents valid measurements for young patients with BD and their UR, specifically with depressive symptoms according to validated clinical ratings and self-reported questionnaires. Smartphone-based self-monitored mood was, compared with self-reported activity and sleep, better at discriminating between young patients with newly diagnosed BD and HC, and between UR and HC, with patients scoring lowest, and UR intermediary lower mood scores. Moreover, self-monitored mood correlates with self-reported activity and sleep duration. Overall, the study shows the potential of using smartphone-based self-monitoring as part of an electronic biomarker for BD, as well as for investigating the impact on UR. The findings indicate that UR have a symptom burden intermediary between young patients with newly diagnosed BD and HC regarding mood, whereas activity and sleep do not differ between UR and HC.

Acknowledgements Sigurd Arne Melbye is an early stage researcher of technology-enabled mental health (TEAM) for young people. This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant agreement No. 722561 and the Copenhagen Center for Health Technology (CACHET). A special thanks goes out to all the participants in the study and the Copenhagen Affective Disorder Clinic, Psychiatric Center Copenhagen, Copenhagen University Hospital, Rigshospitalet, Denmark. Also, we wish to acknowledge the BIO team for assisting in recruiting of participants to the cohorts.

Funding This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie

Skłodowska-Curie Grant agreement No. 722561 and the Copenhagen Center for Health Technology (CACHET).

Availability of data and materials The study is ongoing; therefore, the research data are not shared.

Conflict of interest SM, SS, KC, HLK, KS, and MFJ have no conflicts to declare. LVK has, within the preceding 3 years, been a consultant for Lundbeck. MV has the last 3 years been a consultant for Sunovion, Janssen, and Lundbeck. EMC has been a consultant for Eli Lilly, Astra Zeneca, Servier, Bristol-Myers Squibb, Lundbeck, and Medilink within the last 3 years. JEB and MF are co-founders and shareholders of Monsenso ApS.

Ethics approval and consent to participate The Bipolar Illness Onset (BIO) study has been approved by the ethics committee in the Capital Region, Copenhagen, Denmark (Ref. nr. H-7-2014-007) and the Danish Data Protection Agency, Capital Region of Copenhagen (Protocol no.: RHP-2015-023).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Yatham LN, Berettini WH (2010) Genetics of bipolar disorder: clinical and neurobiological foundations. In: Yatham LN, Maj M (eds) *Bipolar disorder: clinical and neurobiological foundations*, vol 110-23. Wiley-Blackwell, Singapore
2. Frederick KG, Kay Redfield J (2007) *Manic-depressive illness*. Oxford University Press, New York
3. Kessing LV, Hansen HV, Christensen EM, Dam H, Gluud C, Wetterslev J (2014) Do young adults with bipolar disorder benefit from early intervention? *J Affect Disord* 152–154:403–408. <https://doi.org/10.1016/j.jad.2013.10.001>
4. Baldessarini RJ, Tondo L, Baethge CJ, Lepri B, Bratti IM (2007) Effects of treatment latency on response to maintenance treatment in manic-depressive disorders. *Bipolar Disord* 9(4):386–393. <https://doi.org/10.1111/j.1399-5618.2007.00385.x>
5. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE (2005) Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Arch Gen Psychiatry* 62(6):593–602. <https://doi.org/10.1001/archpsyc.62.6.593>
6. Hacımusalar Y, Dogan ES (2019) Assessment of the functioning levels and related factors in patients with bipolar disorder during remission. *Noro psikiyatri arsivi* 56(3):213–218. <https://doi.org/10.29399/npa.23539>
7. St-Amand J, Provencher MD, Belanger L, Morin CM (2013) Sleep disturbances in bipolar disorder during remission. *J Affect Disord* 146(1):112–119. <https://doi.org/10.1016/j.jad.2012.05.057>

8. MacQueen GM, Marriott M, Begin H, Robb J, Joffe RT, Young LT (2003) Subsyndromal symptoms assessed in longitudinal, prospective follow-up of a cohort of patients with bipolar disorder. *Bipolar Disord* 5(5):349–355. <https://doi.org/10.1034/j.1399-5618.2003.00048.x>
9. Wareham NJ, Jakes RW, Rennie KL, Mitchell J, Hennings S, Day NE (2002) Validity and repeatability of the EPIC-norfolk physical activity questionnaire. *Int J Epidemiol* 31(1):168–174. <https://doi.org/10.1093/ije/31.1.168>
10. Pew Research Center (2018) Teens, Social Media & Technology 2018
11. Rideout V (2016) Measuring time spent with media: the common sense census of media use by US 8- to 18-year-olds. *J Child Media JOCAM NEXT* 10(1):138–144. <https://doi.org/10.1080/17482798.2016.1129808>
12. Strejilevich SA, Martino DJ, Murru A, Teitelbaum J, Fassi G, Marengo E, Igoa A, Colom F (2013) Mood instability and functional recovery in bipolar disorders. *Acta Psychiatr Scand* 128(3):194–202
13. Patel R, Lloyd T, Jackson R, Ball M, Shetty H, Broadbent M, Geddes JR, Stewart R, McGuire P, Taylor M (2015) Mood instability is a common feature of mental health disorders and is associated with poor clinical outcomes. *BMJ Open* 5(5):e007504. <https://doi.org/10.1136/bmjopen-2014-007504>
14. O'Donnell LA, Ellis AJ, Van de Loo MM, Stange JP, Axelson DA, Kowatch RA, Schneck CD, Miklowitz DJ (2018) Mood instability as a predictor of clinical and functional outcomes in adolescents with bipolar I and bipolar II disorder. *J Affect Disord* 236:199–206. <https://doi.org/10.1016/j.jad.2018.04.021>
15. Faurholt-Jepsen M, Frost M, Busk J, Christensen EM, Bardram JE, Vinberg M, Kessing LV (2019) Is smartphone-based mood instability associated with stress, quality of life, and functioning in bipolar disorder? *Bipolar Disord* 21(7):611–620. <https://doi.org/10.1111/bdi.12796>
16. Stanislaus S, Faurholt-Jepsen M, Vinberg M, Coello K, Kjaerstad HL, Melbye S, Sletved KSO, Christensen EM, Frost M, Bardram JE, Kessing LV (2020) Mood instability in patients with newly diagnosed bipolar disorder, unaffected relatives, and healthy control individuals measured daily using smartphones. *J Affect Disord* 271:336–344. <https://doi.org/10.1016/j.jad.2020.03.049>
17. Golding J, Northstone K, Miller LL, Davey Smith G, Pembrey M (2013) Differences between blood donors and a population sample: implications for case–control studies. *Int J Epidemiol* 42(4):1145–1156. <https://doi.org/10.1093/ije/dyt095>
18. Kupfer DJ, Weiss BL, Foster G, Detre TP, McPartland R (1974) Psychomotor activity in affective states. *Arch Gen Psychiatry* 30(6):765–768. <https://doi.org/10.1001/archpsyc.1974.01760120029005>
19. Faurholt-Jepsen M, Brage S, Vinberg M, Christensen EM, Knorr U, Jensen HM, Kessing LV (2012) Differences in psychomotor activity in patients suffering from unipolar and bipolar affective disorder in the remitted or mild/moderate depressive state. *J Affect Disord* 141(2–3):457–463. <https://doi.org/10.1016/j.jad.2012.02.020>
20. Scott J, Murray G, Henry C, Morken G, Scott E, Angst J, Merikangas KR, Hickie IB (2017) Activation in bipolar disorders: a systematic review. *JAMA Psychiatry* 74(2):189–196. <https://doi.org/10.1001/jamapsychiatry.2016.3459>
21. Faurholt-Jepsen M, Vinberg M, Frost M, Christensen EM, Bardram JE, Kessing LV (2015) Smartphone data as an electronic biomarker of illness activity in bipolar disorder. *Bipolar Disord* 17(7):715–728. <https://doi.org/10.1111/bdi.12332>
22. Bauer MS, Crits-Christoph P, Ball WA, Dewees E, McAllister T, Alahi P, Cacciola J, Whybrow PC (1991) Independent assessment of manic and depressive symptoms by self-rating. Scale characteristics and implications for the study of mania. *Arch Gen Psychiatry* 48(9):807–812. <https://doi.org/10.1001/archpsyc.1991.01810330031005>
23. Tsanas A, Saunders K, Bilderbeck A, Palmius N, Goodwin G, De Vos M (2017) Clinical insight into latent variables of psychiatric questionnaires for mood symptom self-assessment. *JMIR Ment Health* 4(2):e15. <https://doi.org/10.2196/mental.6917>
24. Hamilton M (1960) A rating scale for depression. *J Neurol Neurosurg Psychiatry* 23:56–62. <https://doi.org/10.1136/jnnp.23.1.56>
25. Stanislaus S, et al (2020) Daily self-reported and smartphone-generated activity in patients with bipolar disorder, unaffected relatives and healthy control individuals (**In review**)
26. Young RC, Biggs JT, Ziegler VE, Meyer DA (1978) A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry* J Ment Sci 133:429–435. <https://doi.org/10.1192/bjp.133.5.429>
27. Harvey AG (2008) Sleep and circadian rhythms in bipolar disorder: seeking synchrony, harmony, and regulation. *Am J Psychiatry* 165(7):820–829. <https://doi.org/10.1176/appi.ajp.2008.08010098>
28. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR (2003) Patient compliance with paper and electronic diaries. *Control Clin Trials* 24(2):182–199. [https://doi.org/10.1016/s0197-2456\(02\)00320-3](https://doi.org/10.1016/s0197-2456(02)00320-3)
29. la Cour Karotki NF, Coello K, Stanislaus S, Melbye S, Kjaerstad HL, Sletved KS, Kessing LV, Vinberg M (2020) Sleep and physical activity in patients with newly diagnosed bipolar disorder in remission, their first-degree unaffected relatives and healthy controls. *Int J Bipolar Disord* 8:1–16. <https://doi.org/10.1186/s40345-020-00181-6>
30. Jackson A, Cavanagh J, Scott J (2003) A systematic review of manic and depressive prodromes, vol 74. Elsevier, Amsterdam. [https://doi.org/10.1016/s0165-0327\(02\)00266-5](https://doi.org/10.1016/s0165-0327(02)00266-5)
31. Barnett JH, Smoller JW (2009) The genetics of bipolar disorder. *Neuroscience* 164(1):331–343. <https://doi.org/10.1016/j.neuroscience.2009.03.080>
32. Vinberg M, Miskowiak K, Kessing LV (2013) Risk markers for affective disorder, a seven-years follow up study of a twin cohort at low and high risk for affective disorder. *J Psychiatr Res* 47(5):565–571. <https://doi.org/10.1016/j.jpsychires.2013.01.013>
33. Sahoo MK, Chakrabarti S, Kulhara P (2012) Detection of prodromal symptoms of relapse in mania and unipolar depression by relatives and patients. *Indian J Med Res* 135(2):177–183
34. Woo YS, Bahk W-M, Hong J, Yoon B-H, Hwang T-Y, Kim M-D, Jon D-I (2016) Use of a smartphone application to screen for bipolar spectrum disorder in a community sample. *Health Inform J* 22(3):779–788. <https://doi.org/10.1177/1460458215589601>
35. BinDhim NF, Shaman AM, Trevena L, Basyouni MH, Pont LG, Alhawassi TM (2015) Depression screening via a smartphone app: cross-country user characteristics and feasibility. *J Am Med Inform Assoc JAMIA* 22(1):29–34. <https://doi.org/10.1136/amiajnl-2014-002840>
36. Faurholt-Jepsen M, Frost M, Christensen EM, Bardram JE, Vinberg M, Kessing LV (2019) The association between mixed symptoms, irritability and functioning measured using smartphones in bipolar disorder. *Acta Psychiatr Scand* 139(5):443–453. <https://doi.org/10.1111/acps.13021>
37. Faurholt-Jepsen M, Frost M, Christensen EM, Bardram JE, Vinberg M, Kessing LV (2019) The validity of daily patient-reported anxiety measured using smartphones and the association with stress, quality of life and functioning in patients with bipolar disorder. *J Affect Disord* 257:100–107. <https://doi.org/10.1016/j.jad.2019.07.029>
38. Stanislaus S, et al (2020) Daily self-reported and automatically generated smartphone-based sleep measurements in patients with newly diagnosed bipolar disorder, unaffected first-degree relatives, and healthy control individuals. Accepted for publishing in *Evid Based Ment Health*

39. Melbye S, et al (2020) Smartphone-based self-monitoring, treatment, and automatically generated data in children, adolescents, and young adults with psychiatric disorders: a systematic review. Resubmitted after review
40. Gershon A, Kaufmann CN, Torous J, Depp C, Ketter TA (2019) Electronic Ecological Momentary Assessment (EMA) in youth with bipolar disorder: demographic and clinical predictors of electronic EMA adherence. *J Psychiatr Res* 116:14–18. <https://doi.org/10.1016/j.jpsychires.2019.05.026>
41. Tsanas A, Saunders KEA, Bilderbeck AC, Palmius N, Osipov M, Clifford GD, Goodwin GM, De Vos M (2016) Daily longitudinal self-monitoring of mood variability in bipolar disorder and borderline personality disorder. *J Affect Disord* 205:225–233. <https://doi.org/10.1016/j.jad.2016.06.065>
42. Matthews M, Abdullah S, Murnane E, Volda S, Choudhury T, Gay G, Frank E (2016) Development and evaluation of a smartphone-based measure of social rhythms for bipolar disorder. *Assessment* 23(4):472–483. <https://doi.org/10.1177/1073191116656794>
43. Kessing LV, Munkholm K, Faurholt-Jepsen M, Miskowiak KW, Nielsen LB, Frikke-Schmidt R, Ekstrom C, Winther O, Pedersen BK, Poulsen HE, McIntyre RS, Kapczinski F, Gattaz WF, Bardram J, Frost M, Mayora O, Knudsen GM, Phillips M, Vinberg M (2017) The Bipolar Illness Onset study: research protocol for the BIO cohort study. *BMJ Open* 7(6):e015462. <https://doi.org/10.1136/bmjopen-2016-015462>
44. Coello K, Kjaerstad HL, Stanislaus S, Melbye S, Faurholt-Jepsen M, Miskowiak KW, McIntyre RS, Vinberg M, Kessing LV, Munkholm K (2019) Thirty-year cardiovascular risk score in patients with newly diagnosed bipolar disorder and their unaffected first-degree relatives. *Aust N Z J Psychiatry* 53(7):651–662. <https://doi.org/10.1177/0004867418815987>
45. Kjaerstad HL, Mistar N, Coello K, Stanislaus S, Melbye SA, Harmer CJ, Vinberg M, Miskowiak K, Kessing LV (2019) Aberrant cognition in newly diagnosed patients with bipolar disorder and their unaffected relatives. *Psychol Med*. <https://doi.org/10.1017/s0033291719001867>
46. Wing JK, Babor T, Brugha T, Burke J, Cooper JE, Giel R, Jablenski A, Regier D, Sartorius N (1990) SCAN. Schedules for clinical assessment in neuropsychiatry. *Arch Gen Psychiatry* 47(6):589–593. <https://doi.org/10.1001/archpsyc.1990.01810180089012>
47. Bech P, Gram LF, Dein E, Jacobsen O, Vitger J, Bolwig TG (1975) Quantitative rating of depressive states. *Acta Psychiatr Scand* 51(3):161–170. <https://doi.org/10.1111/j.1600-0447.1975.tb00002.x>
48. Rosa AR, Sanchez-Moreno J, Martinez-Aran A, Salamero M, Torrent C, Reinares M, Comes M, Colom F, Van Riel W, Ayuso-Mateos JL, Kapczinski F, Vieta E (2007) Validity and reliability of the Functioning Assessment Short Test (FAST) in bipolar disorder. *Clin Pract Epidemiol Ment Health* 3:5. <https://doi.org/10.1186/1745-0179-3-5>
49. Bech P, Rasmussen NA, Olsen LR, Noerholm V, Abildgaard W (2001) The sensitivity and specificity of the major depression inventory, using the Present State Examination as the index of diagnostic validity. *J Affect Disord* 66(2–3):159–164. [https://doi.org/10.1016/s0165-0327\(00\)00309-8](https://doi.org/10.1016/s0165-0327(00)00309-8)
50. Altman EG, Hedeker D, Peterson JL, Davis JM (1997) The altman self-rating mania scale. *Biol Psychiat* 42(10):948–955. [https://doi.org/10.1016/s0006-3223\(96\)00548-3](https://doi.org/10.1016/s0006-3223(96)00548-3)
51. Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE, Pratt M, Ekelund U, Yngve A, Sallis JF, Oja P (2003) International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 35(8):1381–1395. <https://doi.org/10.1249/01.mss.0000078924.61453.fb>
52. Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ (1989) The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 28(2):193–213. [https://doi.org/10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4)
53. Bardram J, Frost M, Szántó K, Faurholt-Jepsen M, Vinberg M, Kessing LV (2013) Designing mobile health technology for bipolar disorder: a field trial of the monarca system. *ACM*. <https://doi.org/10.1145/2470654.2481364>
54. Frost M, Marcu G, Hansen R, Szaántó K, Bardram JE (2011) The MONARCA self-assessment system: persuasive personal monitoring for bipolar patients. In: 2011 5th international conference on pervasive computing technologies for healthcare (PervasiveHealth) and workshops, 23–26 May 2011, pp 204–205. <https://doi.org/10.4108/icst.pervasivehealth.2011.246050>
55. Merikangas KR, Swendsen J, Hickie IB, Cui L, Shou H, Merikangas AK, Zhang J, Lamers F, Crainiceanu C, Volkow ND, Zipunnikov V (2019) Real-time mobile monitoring of the dynamic associations among motor activity, energy, mood, and sleep in adults with bipolar disorder. *JAMA Psychiatry* 76(2):190–198. <https://doi.org/10.1001/jamapsychiatry.2018.3546>