Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder

MARIA FAURHOLT-JEPSEN,¹ MAJ VINBERG,¹ MAD S FROST,² SUNE DEBEL,²
ELLEN MARGRETHE CHRISTENSEN,¹ JAKOB E BARDRAM³ & LARS VEDEL KESSING¹

1 Psychiatric Center Copenhagen, Rigshospitalet, Copenhagen, Denmark

2 The Pervasive Interaction Laboratory (PIT Lab), IT University of Copenhagen, Copenhagen, Denmark

3 DTU Compute, Technical University of Denmark (DTU), Lyngby, Denmark

---

Key words

bipolar disorder, smartphone, illness activity, automatically generated objective data, objective outcome measure

Correspondence

Maria Faurholt-Jepsen,
Psychiatric Center Copenhagen,
Rigshospitalet, Blegdamsvej 9,
DK- 2100 Copenhagen,
Denmark.

Telephone (+45) 3864 7073

Email: maria@faurholt-jepsen.dk

Received 7 September 2015;
revised 20 January 2016;
accepted 2 February 2016

Abstract

Smartphones are useful in symptom-monitoring in bipolar disorder (BD). Objective smartphone data reflecting illness activity could facilitate early treatment and act as outcome in efficacy trials. A total of 29 patients with BD presenting with moderate to severe levels of depressive and manic symptoms used a smartphone-based self-monitoring system during 12 weeks. Objective smartphone data on behavioral activities were collected. Symptoms were clinically assessed every second week using the Hamilton Depression Rating Scale and the Young Mania Rating Scale. Objective smartphone data correlated with symptom severity. The more severe the depressive symptoms (1) the longer the smartphone's screen was "on"/day, (2) more received incoming calls/day, (3) fewer outgoing calls/day were made, (4) less answered incoming calls/day, (5) the patients moved less between cell towers IDs/day. Conversely, the more severe the manic symptoms (1) more outgoing text messages/day sent, (2) the phone calls/day were longer, (3) the fewer number of characters in incoming text messages/day, (4) the lower duration of outgoing calls/day, (5) the patients moved more between cell towers IDs/day. Further, objective smartphone data were able to discriminate between affective states. Objective smartphone data reflect illness severity, discriminates between affective states in BD and may facilitate the cooperation between patient and clinician. Copyright © 2016 John Wiley & Sons, Ltd.

Introduction

Observer-based semi-structured clinical rating scales, e.g. the Hamilton Depression Rating Scale (HDRS-17) (Hamilton, 1967) and the Young Mania Rating Scale (YMRS)

(Young *et al.*, 1978) are used as the golden standard to assess the level of depressive and manic symptoms in bipolar disorder (BD). The use of these rating scales requires a face-to-face patient–clinician encounter, and the level of affective symptoms is assessed during a clinical evaluation.

Study findings may be unreliable when using rating scales as outcome measures due to methodological issues such as unblinding of raters and patients, differences in rater experiences, missing visits for outcome assessments, baseline score inflations and recall bias (e.g. these scales ask the patient to evaluate their symptoms from the previous days under non-real conditions) (Demitrack *et al.*, 1998; Psaty and Prentice, 2010). Thus, more objective methods with real-time data for assessing affective symptoms are warranted.

Research efforts on BD have developed information technology (IT) platforms for continuous electronic self-monitoring of illness activity using regular cell phones (Bopp *et al.*, 2010; Miklowitz *et al.*, 2012); personal digital assistants (PDAs) (Depp *et al.*, 2010; Husky *et al.*, 2010; Mutschler *et al.*, 2012); computers (Barnes *et al.*, 2015; Bauer *et al.*, 2004; Lauder *et al.*, 2015; Lieberman *et al.*, 2010; Todd *et al.*, 2014); and smartphones (Bardram *et al.*, 2012; Depp *et al.*, 2012, 2015; Faurholt-Jepsen *et al.*, 2014, 2015). Our group has developed an IT platform that also collects rather simple automatically generated objective smartphone data on behavioral activities and shown that these reflect the level of illness activity in patients with BD (Bardram *et al.*, 2012; Faurholt-Jepsen *et al.*, 2014, 2015; Frost *et al.*, 2013).

BD exhibits a very complex psychopathology where the level of psychomotor activity (Faurholt-Jepsen *et al.*, 2015, 2016; Kuhs and Reschke, 1992; Kupfer *et al.*, 1974) and changes in the level of engagement in social activities (Weinstock and Miller, 2008) represent central aspects of the level of illness activity, and these aspects can be assessed objectively. Studies investigating these objective aspects have used sensors such as accelerometers worn on the wrist or the thorax or sensors incorporated into the patients clothing (Faurholt-Jepsen *et al.*, 2012; Javelot *et al.*, 2014; Krane-Gartiser *et al.*, 2014). However, these sensors have not been able to collect multiple objective measures reflecting aspects of various behavioral activities. Furthermore, according to “the Hawthorne effect” using a separate “monitoring device” increases the risk of changes in behavioral activities simply due to the awareness of being monitored (Wickström and Bendix, 2000).

Using smartphones to collect real-time automatically generated objective smartphone data including more detailed objective data during long-term naturalistic settings would make continuous monitoring of illness activity in BD possible. Monitoring illness activity using automatically generated objective smartphone data reflects a monitoring method with low level of intrusiveness, and since many people nowadays own a smartphone it would not

require a separate “monitoring device” for this kind of data collection. Monitoring the level of illness activity without the need for continuous self-monitored data delivered by the patients would allow for collection of information on illness activity even when self-monitored data would not be available, which likely could be the case due to the patients current affective state. Further, patients would not be asked to evaluate their symptoms from the previous days thereby eliminating the risk of potential recall bias, and also objective data would be available during relapses where patients are at risk of presenting with a lack of illness insight. In these situations these kinds of electronic objective data could be especially useful. Continuous real-time monitoring in naturalistic settings and collection of automatically generated objective smartphone data that reflect illness activity could facilitate early intervention and have a potential use as objective outcome measures in efficacy trials (Ankers and Jones, 2009; Bopp *et al.*, 2010).

In 2010 an electronic monitoring system for smartphones (the MONARCA system) for patients with BD was developed by our group (Bardram *et al.*, 2012; Faurholt-Jepsen *et al.*, 2013). The MONARCA system allows for collection of daily electronic self-monitored data on items reflecting illness activity (mood, sleep length, activity level, etc.) and collection of automatically generated objective smartphone data on behavioral activities collected by the smartphone. Previous studies on patients with BD using the MONARCA system have showed a high usability and significant correlations between daily electronic self-monitored data and clinically rated depressive and manic symptoms (Faurholt-Jepsen *et al.*, 2014, 2015; Frost *et al.*, 2013). Furthermore, rather simple automatically generated objective smartphone data collected from smartphones (e.g. the number of outgoing phone calls/day; the number of text messages sent/day) have been shown to correlate with the level of clinically rated depressive and manic symptoms in a group of patient with BD with a stable course of illness (Faurholt-Jepsen *et al.*, 2014, 2015).

However, until now automatically generated objective smartphone data collected using the MONARCA system has been rather simple. The expanding possibilities in technology have allowed for more detailed automatically generated objective smartphone data to be collected by the smartphones, and thus updated versions of the system were developed. The new version of the MONARCA system allowed for collection of information on the amount of time the smartphone’s screen was turned “on”/day; the number of times the smartphone’s screen was turned “on”/day; the number of characters in incoming and

outgoing text messages/day and the number of missed calls/day. These new variables were chosen since we believed that they represent aspects of patterns in communication and phone usage that would change according to the severity of depressive and manic symptoms. Consequently, in the present study we recruited a new sample of patients with BD, which had not previously used the MONARCA system. The patients were included at the beginning of their course of treatment without severity criteria applied ($\text{HDRS-17} < 13$ and $\text{YMRS} < 13$) as an additional inclusion criteria as in our previous studies (Faurholt-Jepsen *et al.*, 2014, 2015). Thus, the included patients presented with more severe depressive and manic symptoms than the patients investigated in our previous studies. Nevertheless, none of the patients fulfilled the criteria for rapid cycling at the time of inclusion. The cohort was followed prospectively for 12 weeks while using an updated version of the MONARCA system.

Based on findings from our previous studies (Faurholt-Jepsen *et al.*, 2014, 2015), the objectives of the present study were to test the following hypotheses in patients with BD presenting with more severe depressive and manic symptoms than investigated in previous studies:

- (1) The level of detailed automatically generated objective smartphone data collected from smartphones correlates with the level of depressive and manic symptoms assessed with the HDRS-17 and the YMRS, respectively. The automatically generated objective smartphone data included the following variables: the number of incoming text messages/day; the number of outgoing text messages/day; the duration of phone calls/day; the amount of time the smartphone's screen is turned "on"/day; the number of times the smartphone's screen is turned "on"/day; the number of changes in cell tower IDs/day; the number of characters in incoming text messages/day; the number of characters in outgoing text messages/day; the number of incoming phone calls/day; the number of outgoing phone calls/day; and the number of missed calls/day.
- (2) These automatically generated objective smartphone data collected from smartphones can discriminate between different categories of affective states (depression, mania and euthymia).

Methods

Study participants and settings

Patients with BD were recruited from the Copenhagen Clinic for Affective Disorders, Psychiatric Center Copenhagen, Denmark, during a period from October 2013 to

December 2014. The clinic is a specialized outpatient clinic for patients with mood disorders and treatment at the clinic comprises two years of combined evidence based psychopharmacological treatment and supporting therapy, including group psychoeducation. Further details about the treatment program at the Copenhagen Clinic for Affective Disorders and the effect of this have been described elsewhere (Kessing *et al.*, 2013).

Inclusion criteria were: BD diagnosis according to the International Classification of Diseases 10th revision (ICD-10) using the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) interview (Wing *et al.*, 1990). Exclusion criteria were: pregnancy; lack of Danish language skills; and schizophrenia, schizotypal or delusional disorders according to the SCAN interview. The reason for excluding pregnant patients was that perhaps pregnant women with BD would present with different communicative patterns on their smartphones than usual due to maternity leave. The included patients participated for a 12 weeks study period during the very early phase of their course of treatment at the clinic and received various types, doses and combinations of psychopharmacological treatment during the study. The patients were invited to participate in the present study following referral to the clinic and had not previously used the MONARCA system or participated in previous MONARCA studies. Clinical and socio-demographic data were collected at inclusion.

The MONARCA system for smartphones

Patients were asked to use the MONARCA system for collection of daily electronic self-monitoring data during the 12 weeks study period. The patients either used their own Android smartphone or were offered to loan a "HTC Desire S" or "LG Nexus 5" smartphone free of charge. The patients used their own SIM card during the study. Economic costs due to data traffic from using the MONARCA system were refunded to the patients. Patients did not receive other economic compensation for participating.

At inclusion all patients received training on how to interact with the smartphone and the MONARCA system according to their needs.

Collection of daily electronic self-monitored data

At the beginning of the study patients included in the study were instructed on how to use the MONARCA system for collection of daily electronic self-monitored data orally; by hands-on demonstration and they could also receive a detailed written manual with pictures from the

system presented (the manual can be retrieved by contacting the authors).

On a daily basis at a self-chosen time during the day patients were prompted to provide daily electronic self-monitored data on experienced symptoms (e.g. mood and sleep length) upon hearing an alarm. If the patient forgot to provide daily electronic self-monitored data it was possible to do so retrospectively for up to two days, and it was then noted in the system that data were collected retrospectively.

Due to technical advancements during the study period, an updated version of the MONARCA system was available during the present study. The updated and more technically advanced version allowed for collection of other more detailed automatically generated objective data and a slightly updated version of the daily electronic self-monitored data was also available. One patient used both versions of the MONARCA system during the study period – initially the first version of the MONARCA system and later during the study the updated version. The remaining of the included patients either used the first version of the MONARCA system ($n = 17$) or the updated version of the MONARCA system ($n = 3$) depending on the date of inclusion in the study.

The following daily electronic self-monitored were evaluated on a daily basis by the patients: mood (scored from depressive to manic on a scale from -3 to $+3$, including scores of $+0.5$ and -0.5); sleep length (number of hours slept/night measured in half hours intervals); medication taken (yes/no/changes [the patients were asked to specify these changes]); activity (scored on a scale from -3 to $+3$); alcohol consumption (number of units/day); mixed mood (yes/no); and indication of the presence of individualized early warning signs (yes/no). In the initial version of the MONARCA system used in the present study the following daily electronic self-monitored data were evaluated in addition: irritability (yes/no); cognitive problems (yes/no); stress (scored on a scale $0-3$).

In the updated version of the MONARCA system the following daily electronic self-monitored data were evaluated in addition to the earlier mentioned items: anxiety (scored on a scale $0-2$); irritability (scored on a scale $0-2$); cognitive problems (scored on a scale $0-2$); and stress (scored on a scale $0-2$). Thus, mood; sleep length; activity level; alcohol consumption and indications of the presence of individualized early warning signs were evaluated on the same scales in both versions of the MONARCA system. Tables 1–6 present the results that reflect the samples in each analysis.

The daily electronic self-monitored data from the last two weeks were visualized graphically directly on the

Table 1. Background and clinical characteristics of patients with bipolar disorder using the MONARCA system, $N = 29$ ¹

Age (years)	30.2 (8.8)
Female sex, % (n)	62 (18)
HDRS-17 ² at inclusion	12.6 (5.5)
YMRS ³ at inclusion	4.5 (5.1)
Number of depressive episodes	4 [2–6.5]
Number of (hypo)manic episodes	2.5 [1–5]
Number of hospitalizations	1 [0–2]
Bipolar disorder type I, % (n)	55.2 (16)
Age at onset (years)	21.0 (6.7)
Illness duration (years)	9.3 (6.2)
Years of education after primary school	4.8 (3.1)
Social status	
Employed full time, % (n)	7.1 (2)
Student, % (n)	35.7 (10)
Unemployed, % (n)	32.1 (9)
In relationship, % (n)	51.7 (15)
Psychopharmacological treatment	
Anticonvulsants, % (n)	48.3 (14)
Lithium, % (n)	44.8 (13)
Antipsychotics, % (n)	72.4 (21)
Antidepressants, % (n)	6.9 (2)

¹Data are mean (standard deviation, SD), median [interquartile range, IQR] or proportions (%), (n) unless otherwise stated.

²Hamilton Depression Rating Scale 17-item score

³Young Mania Rating Scale score.

smartphones screen by touching a “visualization” button in the MONARCA software (Figure 1). If the patients wished to get access and visualization of data entered more than two weeks ago they could log on to a secure web-site where all entered personal data were available.

Collection of automatically generated objective data

Automatically generated objective smartphone data reflecting aspects of behavioral activities were collected around the clock. This data collection did not require the patients to interact with the software. The following were collected: the number of incoming text messages/day; the number of outgoing text messages/day; the duration of phone calls/day (collected with a maximum of 120 seconds/call); the amount of time the smartphone's screen was “on”/day; the number of times the smartphone's screen was turned “on”/day; and the number of cell tower ID changes/day.

In an updated version of the MONARCA system the following additional data were collected: the number of

Table 2. Level of depressive and manic symptoms during affective states in patients with bipolar disorder, $N = 182^1$

	Depressive state, $n = 62^4$	Manic or mixed state, $n = 21^5$	Euthymic state, $n = 99^6$
HDRS-17 ²	17.1 (3.7)	11.2 (4.0)	7.0 (3.6)
YMRS ³	3.7 (3.1)	16.9 (3.4)	2.7 (3.0)

¹ N represents the total number of clinical assessments with repeated measurements per patient during a follow-up period of 12 weeks. Data are mean (standard deviation, SD) and raw unadjusted values.

²Hamilton Depression Rating Scale 17-item score.

³Young Mania Rating Scale score.

⁴Defined as HDRS-17 ≥ 13 and YMRS < 13 .

⁵Defined as YMRS ≥ 13 .

⁶Defined as HDRS-17 < 13 and YMRS < 13 .

characters in incoming text messages/day; the number of characters in outgoing text messages/day; the number of incoming calls/day; the number of outgoing calls/day; and the number of missed calls/day.

Patients were instructed to use the smartphones for their usual communicative purposes and to carry it with them during the day.

The collected automatically generated objective smartphone data were chosen because of predefined hypotheses and finding from studies by the authors (Faurholt-Jepsen *et al.*, 2014, 2015) that these would reflect the level of and changes in psychopathology in patients with BD.

Clinical assessments

The BD diagnosis according to ICD-10 was confirmed using SCAN (Wing *et al.*, 1990). Patients visited the researcher (MFJ) scheduled every second week during the 12 weeks study period for clinical assessment. The severity of depressive symptoms was assessed using the HDRS-17 (Hamilton, 1967) and manic symptoms was assessed using the YMRS (Young *et al.*, 1978). These rating scales refer to a time-period of the current day and three previous days, which was applied on all clinical assessments. The researcher (MFJ) did not have access to the automatically generated objective smartphone data and was therefore blinded to these data at the time of the clinical assessments.

Statistical methods

The statistical analyses were defined a priori and computed using linear mixed regression models. These regression models allow for both *within*-individual and *between*-individual variations over

time of the variables considered in the analyses. For each analysis we employed a two-leveled linear mixed regression model with the first level representing the repeated measurements of the smartphone data *within*-individuals and the second level representing the *between*-individuals variation of the smartphone data. All models included a patient specific random intercept and a patient specific random effect to account for patient specific correlations between the dependent variables (the smartphone data) and the independent variables (e.g. the level of symptoms; affective states) over time. Other covariates were specified as fixed effects in the models. Model assumptions were checked visually by means of residuals and QQ plots. We firstly considered an unadjusted analysis, and secondly an analysis adjusted for age and sex as possible confounders. Averages of daily electronic self-monitored data and automatically generated objective smartphone data were analyzed for the day of the clinical assessments and the three previous days since the HDRS-17 and the YMRS address symptoms over the three previous days (Tables 3 and 4). In addition, we analyzed associations between the automatically generated objective smartphone data and sub-items of the HDRS-17 (mood [sub-item 1], psychomotor retardation [sub-item 8], psychomotor agitation [sub-item 9]) and sub-items of the YMRS (mood [sub-item 1], activity [sub-item 2], speech [sub-item 6]), respectively (Table 5). These items were selected as they represent central and objectively measurable parts of depression and mania, and as we previously have shown they correlate with simple automatically generated objective smartphone data (Faurholt-Jepsen *et al.*, 2014, 2015).

Table 3. Correlations between daily electronic self-monitored data¹ collected using the MONARCA system and depressive and manic symptoms assessed with HDRS-17² and YMRS³, respectively

	Model 1 ⁴			Model 2 ⁴		
	Coefficient	95% Confidence interval	<i>p</i>	Coefficient	95% Confidence interval	<i>p</i>
<i>Mood (scale: -3 to +3), N = 29</i>						
HDRS-17	-0.051	-0.066; -0.036	<0.001	-0.049	-0.063; -0.034	<0.001
HDRS-17 sub-item 1 (mood)	-0.33	-0.41; -0.25	<0.001	-0.32	-0.39; -0.24	<0.001
YMRS	0.044	0.028; 0.060	<0.001	0.045	0.030; 0.060	<0.001
YMRS sub-item 1 (mood)	0.38	0.27; 0.49	<0.001	0.38	0.28; 0.49	<0.001
<i>Sleep length (hours/night), N = 29</i>						
HDRS-17	-0.002	-0.041; 0.036	0.91	-0.003	-0.042; 0.035	0.87
YMRS	-0.049	0.090; -0.008	0.018	-0.043	-0.084; -0.003	0.036
<i>Activity level (scale: -3 to +3), N = 29</i>						
HDRS-17	-0.030	-0.052; -0.009	0.006	-0.032	-0.054; -0.010	0.005
YMRS	0.034	0.011; 0.057	0.003	0.036	0.013; 0.059	0.002
<i>Stress level (scale: 0 to +3), N = 17</i>						
HDRS-17	0.062	0.039; 0.080	<0.001	0.062	0.039; 0.084	<0.001
YMRS	0.035	0.010; 0.061	0.007	0.033	0.007; 0.060	0.013
<i>Stress level (scale: 0 to +2), N = 13</i>						
HDRS-17	0.018	-0.004; 0.040	0.11	0.018	-0.002; 0.038	0.080
YMRS	0.007	-0.12; 0.027	0.48	0.010	-0.009; 0.028	0.31
<i>Anxiety level (scale: 0 to +2), N = 13</i>						
HDRS-17	0.034	0.015; 0.052	<0.001	0.033	0.13; 0.052	0.001
YMRS	-0.0003	-0.020; 0.019	0.98	0.003	-0.017; 0.022	0.79
<i>Irritability (scale: 0 to +2), N = 13</i>						
HDRS-17	0.036	0.017; 0.055	<0.001	0.037	0.017; 0.057	<0.001
YMRS	0.010	-0.010; 0.030	0.31	0.011	-0.10; 0.031	0.30
<i>Cognitive difficulties (scale: 0 to +2), N = 13</i>						
HDRS-17	0.062	0.043; 0.080	<0.001	0.056	0.038; 0.073	<0.001
YMRS	0.004	-0.017; 0.024	0.74	0.007	-0.013; 0.028	0.50

¹Averages of the smartphone data were analyzed for the current day and the three previous days before assessment with the HDRS-17 and YMRS, as these rating scales address symptoms over the last four days.

²Hamilton Depression Rating Scale 17-item.

³Young Mania Rating Scale.

⁴Model 1, unadjusted; Model 2, Adjusted for age and sex.

Patients' affective states were categorized into a euthymic state (HDRS-17 < 13 and YMRS < 13); a depressive state (HDRS-17 ≥ 13 and YMRS < 13); and a (hypo-) manic or mixed state (YMRS ≥ 13) (Table 6).

Data were entered using Epidata® (The EpiData Association, Odense, Denmark, Europe) and STATA version 12.1 (StataCorp LP, College Station, TX, USA) was used for the statistical analyses. The significance level of the *p*-values in the statistical models was set to 0.05 (two-sided).

Ethical considerations

The study was approved by the Regional Ethics Committee in the Capital Region of Denmark (H-2-2011-056)

and the Danish Data protection agency (2013-41-1710). All participants were given both written and oral information about the study before obtaining informed consent. Participants were informed that they could withdraw from the study at any time without this having any effect on their course of treatment at the Copenhagen Clinic for Affective Disorder, Psychiatric Center Copenhagen, Denmark.

Results

Background characteristics

During a period from October 2013 to December 2014, 51 eligible patients with a diagnosis of BD were invited to

Table 4. Correlations between automatically generated objective data¹ collected using smartphones and depressive and manic symptoms assessed with HDRS-17² and YMRS³, respectively.

	Model 1 ⁴			Model 2 ⁴		
	Coefficient	95% Confidence interval	<i>p</i>	Coefficient	95% Confidence interval	<i>p</i>
<i>Incoming text messages (numbers/day), N = 29</i>						
HDRS-17	-0.19	-0.41; 0.04	0.10	-0.20	-0.43; 0.03	0.08
YMRS	0.14	-0.10; 0.38	0.25	0.13	-0.11; 0.38	0.29
<i>Outgoing text messages (numbers/day), N = 29</i>						
HDRS-17	0.11	-0.16; 0.38	0.42	0.07	-0.21; 0.35	0.62
YMRS	0.29	0.04; 0.55	0.02	0.31	0.05; 0.56	0.01
<i>Duration screen "on" (seconds/day), N = 29</i>						
HDRS-17	194.8	-4.7; 394.3	0.06	209.6	11.0; 408.2	0.04
YMRS	77.5	-139.4; 294.4	0.48	64.9	-155.3; 285.0	0.56
<i>Number of times screen turned "on" (number/day), N = 29</i>						
HDRS-17	0.15	-2.0; 2.3	0.89	-0.13	-2.2; 1.9	0.90
YMRS	1.39	-0.88; 3.7	0.23	1.05	-1.2; 3.3	0.40
<i>Duration of phone calls (seconds/day), N = 17</i>						
HDRS-17	0.54	-0.37; 1.45	0.25	0.66	-0.18; 1.50	0.12
YMRS	0.39	-0.35; 1.12	0.30	0.63	0.022; 1.23	0.042
<i>Cell tower ID changes (number/day), N = 17</i>						
HDRS-17	-0.57	-0.97; -0.16	0.006	-0.56	-0.94; -0.18	0.004
YMRS	0.55	0.17; 0.94	0.005	0.40	0.006; 0.79	0.04
<i>Number of characters in incoming text messages, N = 13</i>						
HDRS-17	1.53	-0.91; 3.98	0.22	1.62	-0.98; 4.24	0.22
YMRS	-3.49	-5.62; -1.35	0.001	-3.54	-5.71; -1.37	0.001
<i>Number of characters in outgoing text messages, N = 13</i>						
HDRS-17	1.43	-0.93; 3.78	0.24	1.60	-0.78; 3.97	0.19
YMRS	1.81	-0.19; 3.80	0.076	1.55	-0.38; 3.49	0.12
<i>Incoming calls (numbers/day), N = 13</i>						
HDRS-17	0.05	0.01; 0.08	0.01	0.05	0.01; 0.09	0.01
YMRS	-0.02	-0.08; 0.04	0.52	-0.01	-0.07; 0.05	0.69
<i>Outgoing calls (numbers/day), N = 13</i>						
HDRS-17	-0.11	-0.24; 0.02	0.09	-0.12	-0.23; -0.008	0.03
YMRS	0.003	-0.15; 0.16	0.97	-0.007	-0.16; 0.15	0.93
<i>Missed calls (numbers/day), N = 13</i>						
HDRS-17	0.04	0.005; 0.08	0.03	0.05	0.01; 0.08	0.006
YMRS	-0.002	-0.06; 0.06	0.95	0.008	-0.05; 0.07	0.79
<i>Duration of incoming calls (seconds/day), N = 13</i>						
HDRS-17	5.13	-31.1; 41.3	0.78	21.5	-11.1; 54.2	0.20
YMRS	-11.3	-55.7; 33.1	0.62	-19.8	-64.4; 24.7	0.38
<i>Duration of outgoing calls (seconds/day), N = 13</i>						
HDRS-17	-5.1	-20.5; 10.3	0.52	-1.8	-17.7; 14.2	0.83
YMRS	-25.7	-49.4; -2.1	0.030	-24.3	-48.4; -0.20	0.041

¹Averages of the smartphone data were analyzed for the current day and the three previous days before assessment with the HDRS-17 and YMRS, as these rating scales address symptoms over the last three days

²Hamilton Depression Rating Scale 17-item.

³Young Mania Rating Scale.

⁴Model 1, unadjusted; Model 2, adjusted for age and sex.

participate in the present study and 32 (62.7%) agreed to participate. The main reasons for declining were that (1)

it would be too time-consuming to participate ($n=13$) and (2) the system was not available for iPhones ($n=5$).

One patient declined to participate due to the collection of automatically generated objective smartphone data. No differences in socio-demographic characteristics between patients who were not included in the study and patients included in the study were observed. Three patients dropped out immediately after inclusion (changed their minds regarding participation), and thus a total of 29 patients participated in the study. Three of the included patients used their own smartphone during the study. No differences between these patients and the rest of the included patients in clinical characteristics and use of daily electronic self-monitoring were observed. A total of 10.3% (21 out of 203 visits) of the patients visits with the researcher scheduled every second week for HDRS-17 and YMRS ratings were missing (patients who did not show up for the scheduled assessments), leaving a total of 182 available clinical ratings.

The socio-demographic characteristics of patients at inclusion are presented in Table 1. Patients had a mean age of 30.2 (standard deviation [SD] 8.8) years, 62% ($n=18$) were women.

Table 2 present the levels of depressive and manic symptoms according to affective states during the 12 weeks study period represented by raw and unadjusted mean scores of the HDRS-17 and the YMRS, respectively. A total of 182 clinical ratings were available.

Daily electronic self-monitored data and severity of depressive and manic symptoms

Table 3 presents the results from models regarding daily electronic self-monitored data and clinically rated depressive and manic symptoms measured using the HDRS-17 and YMRS, respectively.

Overall, daily electronic self-monitored data on mood and activity level correlated significantly with scores on both the HDRS-17 and the YMRS, respectively. There was a significant negative correlation between daily electronic self-monitored data on mood and scores on the HDRS-17 in both the unadjusted model and the models adjusted for age and sex (adjusted model: $B=-0.049$, 95% confidence interval [CI]: -0.063 ; -0.034 , $p<0.001$). Thus, for every increase of 10 points on the HDRS-17 there was a decrease in self-monitored mood score of -0.49 on a scale from 0 to -3 . Further, there was a significant positive correlation between daily electronic self-monitored data on mood and scores on the YMRS in both the unadjusted and the models adjusted for age and sex (adjusted model: $B=0.045$, 95% CI: 0.030 ; 0.060 , $p<0.001$). Thus, for every increase of 10

points on the YMRS there was an increase in self-monitored mood score of 0.45 on a scale from 0 to +3.

Further, Table 3 presents the results from models regarding daily electronic self-monitored data on sleep length; stress level; anxiety level; level of irritability; and level of cognitive problems in the same way as presented in relation to the models on daily electronic self-monitored data on mood. In relation to depressive symptoms, daily electronic self-monitored data on activity level; stress level; anxiety level; level of irritability and level of cognitive problems all correlated positively and significantly with the HDRS-17, whereas sleep length did not correlate with the HDRS-17.

In relation to manic symptoms, daily electronic self-monitored data on sleep length correlated significantly and negatively with the YMRS and activity level and stress level correlated positively with the YMRS. Daily electronic self-monitored data on anxiety level, level of irritability and level of cognitive problems did not correlate with the YMRS.

Automatically generated objective data and severity of depressive and manic symptoms

Table 4 presents the results from models regarding automatically generated objective data and clinically rated depressive and manic symptoms using the HDRS-17 and the YMRS, respectively.

In relation to depressive symptoms in models adjusted for age and sex and, the number of changes in cell tower IDs/day and the number of outgoing phone calls/day correlated significantly and negatively with the HDRS-17, whereas the number of incoming calls/day and the number of missed phone calls/day correlated positively with the HDRS-17. For example, for every increase of 10 points on the HDRS-17 there was a decrease in the number of changes in cell tower IDs/day of 5.6.

In relation to manic symptoms in models adjusted for age and sex, the number of outgoing text messages/day; the duration of phone calls/day; the number of changes in cell tower IDs/day correlated significantly and positively with the YMRS, whereas the number of characters in incoming text messages/day and the duration of outgoing phone calls/day correlated significantly and negatively with the YMRS. For example, for every increase of 10 points on the YMRS there was an increase in the number of changes in cell tower IDs/day of 4.

Table 5 presents the results from models regarding automatically generated objective data and sub-items on the HDRS-17 and the YMRS, respectively.

Table 5. Correlations between automatically generated objective data¹ collected using smartphones and sub-items² on HDRS-17³ and YMRS⁴, respectively.

	Model 1 ⁵			Model 2 ⁵		
	Coefficient	95% Confidence interval	<i>p</i>	Coefficient	95% Confidence interval	<i>p</i>
<i>Incoming text messages</i>						
<i>(numbers/day), N = 29</i>						
HDRS-17 sub-item 1	-1.32	-2.61; -0.04	0.04	-1.35	-2.65; -0.06	0.04
HDRS-17 sub-item 8	-2.91	-5.64; -0.19	0.03	-3.06	-5.83; -0.29	0.03
<i>Outgoing text messages</i>						
<i>(numbers/day), N = 29</i>						
HDRS-17 sub-item 8	-3.39	-6.55; -0.22	0.03	-4.32	-7.56; -1.08	0.009
HDRS sub-item 9	3.30	1.08; 5.52	0.004	3.21	1.02; 5.41	0.004
YMRS sub-item 2	1.58	0.10; 3.06	0.04	1.64	0.15; 3.13	0.03
YMRS sub-item 6	0.96	-0.005; 1.93	0.05	1.04	0.67; 2.03	0.04
<i>Number of times screen turned</i>						
<i>"on" (number/day), N = 29</i>						
HDRS-17 sub-item 9	21.0	-0.15; 42.1	0.052	21.8	1.5; 42.2	0.04
<i>Duration of phone calls</i>						
<i>(seconds/day), N = 17</i>						
HDRS-17 sub-item 9	4.50	-3.27; 12.27	0.26	7.02	0.63; 13.40	0.031
<i>Cell tower ID changes</i>						
<i>(number/day), N = 17</i>						
HDRS-17 sub-item 1	-4.41	-6.49; -2.32	<0.001	-4.02	-6.05; -2.0	<0.001
YMRS sub-item 1	3.31	0.27; 6.34	0.03	3.29	0.19; 5.79	0.04
YMRS sub-item 6	2.01	0.57; 3.46	0.006	1.48	0.005; 2.97	0.04
<i>Number of characters in incoming text messages, N = 13</i>						
HDRS-17 sub-item 1	12.32	-0.37; 25.01	0.057	12.65	-0.64; 25.95	0.060
HDRS-17 sub-item 8	32.38	7.67; 57.10	0.010	35.61	9.57; 61.64	0.007
YMRS sub-item 1	-17.89	-30.76; -5.01	0.006	-17.99	-31.16; -4.82	0.007
YMRS sub-item 2	-24.10	-38.44; -9.75	0.001	-24.60	-39.32; -9.89	0.001
YMRS sub-item 6	-9.07	-16.72; -1.41	0.020	-9.10	-17.00; -1.21	0.024
<i>Number of characters in outgoing</i>						
<i>text messages, N = 13</i>						
YMRS sub-item 6	10.69	3.78; 17.60	0.002	9.71	2.82; 16.61	0.006
<i>Missed calls (numbers/day), N = 13</i>						
HDRS-17 sub-item 1	0.40	0.18; 0.63	<0.001	0.45	0.24; 0.65	<0.001
YMRS sub-item 6	-0.28	-0.44; -0.12	0.001	-0.26	-0.44; -0.09	0.003
<i>Duration of incoming calls</i>						
<i>(seconds/day), N = 13</i>						
HDRS-17 sub-item 1	217.5	-2.8; 437.8	0.053	281.7	78.8; 484.6	0.007
<i>Duration of outgoing calls (seconds/day), N = 13</i>						
YMRS sub-item 1	-193.9	-358.0; -29.8	0.021	-190.5	-364.4; -16.5	0.032

¹Averages of the smartphone data were analyzed for the current day and the three previous days before assessment with the HDRS-17 and YMRS, as these rating scales address symptoms over the last three days.

²HDRS-17 sub-item 1, mood; HDRS-17 sub-item 8, psychomotor retardation; HDRS-17 sub-item 9, psychomotor agitation; YMRS sub-item 1, mood; YMRS sub-item 2, activity; YMRS sub-item 6, speech.

³Hamilton Depression Rating Scale 17-item.

⁴Young Mania rating Scale.

⁵Model 1, unadjusted; Model 2, adjusted for age and sex

As can be seen, in relation to depressive symptoms, the number of times the smartphone screen was turned “on”/day; the duration of phone calls/day; the number of characters in incoming text messages/day; the number of missed calls/day; and the duration of incoming calls/day correlated significantly and positively with sub-items on the HDRS-17, whereas the number of incoming text messages/day and the number of changes in cell towers IDs/day correlated significantly and negatively with sub-items on the HDRS-17. Further, the number of outgoing text messages/day correlated positively with HDRS-17 sub-item 9 (psychomotor agitation), but negatively with sub-item 8 (psychomotor retardation).

In relation to manic symptoms, the number of outgoing text messages/day; the number of changes in cell towers IDs/day; and the number of characters in outgoing text messages/day correlated significantly and positively with sub-items on the YMRS, whereas the number of characters in incoming text messages/day; the number of missed calls/day; and the duration of outgoing calls/day correlated significantly and negatively with sub-items on the YMRS.

Automatically generated objective data and affective states

Table 6 presents the results from models regarding automatically generated objective data and categories of different affective states, as represented by the HDRS-17 and the YMRS categorized as a euthymic state ($\text{HDRS-17} < 13$ and $\text{YMRS} < 13$); a depressive state ($\text{HDRS-17} \geq 13$ and $\text{YMRS} < 13$); and a (hypo-)manic or mixed state ($\text{YMRS} \geq 13$).

Most of the included variables discriminated between a euthymic state versus a depressive state and a (hypo-)manic or mixed state, respectively (see Table 6). Furthermore, the following variables discriminated between a depressive state and a (hypo-)manic or mixed state: the number of outgoing text messages/day; the number of changes in cell tower IDs/day; and the number of characters in incoming text messages/day.

Discussion

This is the first study investigating correlations between detailed automatically generated objective data and illness activity in patients with BD presenting with more severe levels of depressive and manic symptoms than previously ever investigated. The study used a longitudinal design

allowing for both within-individual and between-individual changes in smartphone data over time. Interestingly and in accordance with our hypotheses a large number of the automatically generated objective data reflecting aspects of behavioral activities correlated with the severity of depressive and manic symptoms assessed with the HDRS-17 and the YMRS, respectively and furthermore discriminated between affective states.

Concerning depressive symptoms, five out of 13 variables of the automatically generated objective data correlated with scores on the HDRS-17. The more severe the depressive symptoms the longer the smartphone's screen was “on”/day, more incoming calls/day were received; fewer outgoing calls/day were made; more incoming calls were missed (not answered)/day; and the less the patients moved around between cell towers IDs/day. In accordance with our hypotheses, the more depressive symptoms the less the patients engaged in communicative activities despite being contacted more, but used the smartphone for other activities than communicating (watching television, browsing etc.) and isolated more.

Concerning manic symptoms, five out of 13 variables of the automatically generated objective data correlated with scores on the YMRS. The more severe the manic symptoms the more outgoing text messages/day were sent; the phone calls/day were longer; the number of characters in incoming text messages/day were fewer; the duration of outgoing calls/day was lower; and the more the patients moved around between cell towers IDs/day. In accordance with our hypotheses, the more severe the manic symptoms the more the patients engaged in communicative activities (that was not met by those who was contacted) and the more they moved more around.

The patterns in communicative activities and use of smartphones thus differed between depressive and manic symptoms. Analyses on the ability of the automatically generated objective data to discriminate between different categories of affective states confirmed these findings (Table 6).

A central feature in BD is abnormalities in psychomotor activity typically with psychomotor retardation during BD depression and increased motor activity during mania (Beigel and Murphy, 1971; Goodwin, and Jamison, 1996; Judd *et al.*, 2012). Changes in the level of psychomotor activity has been suggested a useful tool in the monitoring of illness activity in patients with BD (Kupfer *et al.*, 1974). The level of psychomotor activity, measured using separate sensors (e.g. accelerometers) also seems to reflect the level of illness activity in BD (Faurholt-Jepsen *et al.*, 2016; Jones *et al.*, 2005; Krane-Gartiser *et al.*, 2014).

Table 6. Differences in automatically generated objective data collected using smartphones between affective states¹ in patients with bipolar disorder, $N = 182^2$

	Model 1 ³			Model 2 ³		
	Coefficient	95% Confidence interval	<i>p</i>	Coefficient	95% Confidence interval	<i>p</i>
<i>Outgoing text messages (numbers/day)</i>						
Manic or mixed state/ Euthymic state	5.16	0.76; 9.55	0.02	5.30	0.87; 9.72	0.02
Depressive state/ manic or mixed state	-4.20	-8.84; 0.44	0.07	-5.00	-9.78; -0.22	0.04
<i>Duration screen is "on" (seconds/day)</i>						
Depressive state/ Euthymic state	2338.5	-35.1; 4712.1	0.053	2498.3	131.6; 4864.9	0.04
Manic or mixed state/ Euthymic state	3532.1	-159.2; 7223.4	0.06	3597.4	-59.2; 7253.9	0.054
<i>Number of times screen turned "on" (number/day)</i>						
Manic or mixed state/ Euthymic state	36.8	-0.76; 74.4	0.06	41.3	2.9; 79.6	0.03
<i>Duration of phone calls (seconds/day)</i>						
Manic or mixed state/ Euthymic state	11.15	-1.59; 23.89	0.086	14.25	3.37; 25.11	0.010
<i>Cell tower ID changes (number/day)</i>						
Depressive state/ Euthymic state	-5.03	-10.0; -0.04	0.04	-5.34	-10.1; -0.56	0.02
Manic or mixed state/ Euthymic state	6.51	-0.12; 13.1	0.054	4.55	-1.95; 11.0	0.17
Depressive state/ Manic or mixed state	-11.5	-18.4; -4.67	0.001	-9.90	-16.6; -3.15	0.004
<i>Characters in incoming text messages</i>						
Manic or mixed state/ Euthymic state	-35.74	-68.47; -3.01	0.032	-38.12	-71.26; -4.97	0.024
Depressive state/ Manic or mixed state	44.48	8.01; 80.95	0.017	46.80	9.46; 84.14	0.014
<i>Characters in outgoing text messages</i>						
Depressive state/ Euthymic state	17.13	-3.44; 37.70	0.10	21.44	1.02; 41.86	0.040

¹A euthymic state was defined as a Hamilton Depression Rating Scale 17-items (HDRS-17) score < 13 and a Young Mania Rating Scale (YMRS) score < 13. A depressive state was defined as a HDRS-17 score ≥ 13 and an YMRS score < 13. A manic or mixed state was defined as an YMRS score ≥ 13.

² N represents the total number of clinical assessments. A depressive state, $n = 62$. A manic or mixed state, $n = 21$. A euthymic state, $n = 99$.

³Model 1, unadjusted analyses; Model 2, analyses adjusted for age and sex.

However, this is the first study using smartphones for monitoring the level of psychomotor activity: the number of incoming and outgoing text messages/day; the number of times the smartphones screen was turned "on"/day; the duration of phone calls/day; and the number of characters in incoming text messages/day correlated significantly with the level of clinically assessed psychomotor retardation (HDRS-17 sub-item 8), psychomotor agitation (HDRS-17 sub-item 9) and activity level (YMRS sub-item 2). Interestingly, the number of outgoing text messages/day was negatively correlated with HDRS-17 sub-item 8 reflecting the level of psychomotor retardation, but positively correlated with HDRS-17 sub-item 9 reflecting the level of psychomotor agitation. Furthermore, the number of outgoing text messages/day; the number of changes in cell tower IDs/day; the number of characters in incoming text messages/day and the number

of missed calls/day correlated with YMRS sub-item 6 reflecting the level of pressure of speech.

Overall, the results demonstrate that detailed automatically generated objective data reflecting behavioral activities could be used as an electronic marker of illness activity in patients with BD and may facilitate the cooperation between patient and clinician. Automatically generated objective data may increase illness insight in patients with BD and help improving empowerment.

We have previously shown in patients with BD with a stable course of illness that both daily electronic self-monitored data and rather simple automatically generated objective data (the duration of phone calls/day and the number of changes in cells tower IDs/day) correlate with levels of clinically rated depressive and manic symptoms (Faurholt-Jepsen *et al.*, 2014, 2015). In our

Figure 1. An example of the graphs presented to the patients with visualization of the data acquired by the MONARCA system.

previous studies patients were included if they presented in a remitted state or with low levels of depressive and manic symptoms at inclusion, since at that time we believed that severe levels of depressive and manic symptoms could potentially interfere with the patients' ability to interact with the MONARCA system. In contrast to these previous studies, the 29 patients included in the present study were included regardless the severity of depressive and manic symptoms, and consequently presented with 62 depressive episodes and 21 (hypo-) manic or mixed episodes during the 12 weeks study period. Furthermore, the present study included more detailed information of automatically generated objective data than previous studies were able to and thereby provided more detailed information on changes in behavioral activities during affective states. Apart from these two previous studies by our group, correlations between

automatically generated objective data and illness activity in patients with BD have not been investigated.

A strength of the present study is the use of smartphones to collect information on detailed behavioral activities that otherwise would need the patients to wear different sensors, thus increasing the risk of changes in behavioral activities due to the increased awareness of a sensor (Wickström and Bendix, 2000). The automatically generated objective data were collected during real-time naturalistic settings and did not need the patients to interact with the software on the smartphones.

In addition to the automatically generated objective data we also found significant correlations between daily electronic self-monitored data (e.g. mood and stress level) and the HDRS-17 and the YMRS, respectively. These findings are replications from our previous study (Faurholt-Jepsen *et al.*, 2015), but unlike the previous study the

patient included in the present study were presenting with an unstable course of illness. Thus, despite rather high levels of depressive and manic symptoms the patients were able to validly evaluate their current affective state on several parameters.

As found in the present and our prior study (Faurholt-Jepsen *et al.*, 2015) self-monitored sleep length and depressive symptoms measured using the HDRS-17 did not correlate and may reflect that bipolar depression often presents with heterogeneous and atypical symptoms (Mitchell and Malhi, 2004).

In previous studies by our group, two of the inclusion criteria were a HDRS-17 ≤ 17 and YMRS ≤ 17 since at the time for those studies we believed that the severity of symptoms could complicate interacting with the smartphones (Faurholt-Jepsen *et al.*, 2014, 2015). Consequently, patients included in the previous studies were rather stable and did not present with many affective episodes during the study periods. In the present study patients were included regardless of the severity of depressive and manic symptoms. None of the included patients reported or were observed to experience any difficulties interacting with the smartphones due to depressive or manic symptoms.

Limitations

A small sample of patients was included, but each patient was assessed several times using a longitudinal study design and also the employed statistical models allowed for both within-individual and between-individual variations over time adding to increase the statistical power of the study. However, it cannot be excluded that some of the negative findings could be due to type II errors.

The study period could have been longer, also allowing for the patients to present more severe levels of depressive and manic symptoms. However, the patients were included at the beginning of their course of treatment and presented with a more unstable course of illness than previously investigated.

The patients' affective states were defined according to an ICD-10 diagnosis of BD current episode depressive, manic or mixed combined with the total score of depressive and manic symptoms ≥ 13 according to the HDRS-17 and the YMRS, respectively. The cutoff on the HDRS-17 and YMRS of 13 in contrast to a lower cutoff was chosen a priori to increase the validity of a current affective depressive or manic/mixed state (the more severe, the higher the validity). A current euthymic state was defined as a HDRS-17 and YMRS < 13 and in this way also including affective

states with partial remission. In seven cases the manic states also included depressive symptoms with a HDRS ≥ 13 , i.e. a mixed state. Conversely, during depressive episodes the level of manic symptoms was low.

Some patients declined to participate as the MONARCA system was not available for iPhones and Windows phones. It cannot be excluded that patients using other operating systems than Android represent a different sub-group of patients. Future studies are recommended to develop a system to include iPhones and Windows smartphones.

The patients were instructed to use the smartphones for their usual communicative purposes and to carry the smartphone with them. Perhaps the patients from time to time forgot to carry their smartphone with them in which situations automatically generated objective data on all behavioral activities were not collected. However, as can be seen there were significant correlations between the automatically generated objective data and clinically rated depressive and manic symptoms.

Finally, as the present results are innovative, explorative within the area and hypothesis-generating multiple testing in the statistical analyses was not accounted for.

Implications and perspectives

One patient declined to participate due to the collection of automatically generated objective data. None of the included patients reported that they felt "watched", but rather that the MONARCA system was perceived as a "safety net" in case of deterioration. However, collecting data on behavioral activities raises ethical issues, requires a high degree of trust between the patient and clinician and a high degree of data security. All electronic data collected in the present study were encrypted and sent to secure servers at the hospital.

Using smartphones for monitoring the level of illness activity in BD, a serious and chronic illness, reflects an easy and objective method for monitoring of illness activity during long-term naturalistic settings with a low level of intrusiveness. With the limited access to psychiatrists, using automatically generated objective data to trace illness activity could direct the resources to where needed. Further patients with BD may clinically benefit from being introduced to this tool e.g. during psychoeducation or in Internet lectures. Finally, smartphones are able to collect automatically generated objective data that ultimately may be used as outcome measures in future efficacy trials.

Conclusions

Automatically generated objective data on aspects of behavioral activities collected using smartphones reflect the level of depressive and manic symptoms and discriminate between categories of affective states in patients with BD. Further, these automatically generated objective data can be used to detect illness activity in patients presenting with moderate to severe levels of depressive and manic symptoms. Smartphones seems promising as an electronic biomarker in BD and may be used to facilitate the cooperation between patient and clinician.

Acknowledgements/Acknowledgments

The authors would like to thank the patients for participating in the study and Rie Lambæk Mikkelsen, MD for participating in recruiting patients for the study.

Financial support

The European Union (EU) 7th Frame Program funded the MONARCA I trial together with the Mental Health Services, Copenhagen, Denmark, Trygfonden, the Gert Einar

Joergensens foundation, and the A.P. Moeller and the Hustru Chastine Mc-Kinney Moellers foundation for general purposes. The funders had no role in the trial design, data collection, analyses and preparation of the manuscript.

Declaration of interest statement

MFJ has been a consultant for Eli Lilly and Lundbeck. MV has been a consultant for Eli Lilly, Lundbeck, Astra Zeneca and Servier. EMC has been a consultant for Eli Lilly, Astra Zeneca, Servier, Bristol-Myers Squibb, Lundbeck and Medilink. LVK has within the recent three years been a consultant for Lundbeck and Astra Zeneca. MF and JEB are founders and shareholders of Monsenso. SD has no conflicts of interests.

Contributors

MFJ and LVK performed the statistical analyses. MFJ, MV and LVK wrote the first draft of the manuscript. All authors contributed to and have approved the final version of the manuscript.

References

- Ankers D., Jones S.H. (2009) Objective assessment of circadian activity and sleep patterns in individuals at behavioural risk of hypomania. *Journal of Clinical Psychology*, **65**, 1071–1086.
- Bardram J., Frost M., Szanto K., Margu G. (2012) The MONARCA self-assessment system: a persuasive personal monitoring system for bipolar patients. In Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium (IHI '12), pp. 21–30, New York: ACM.
- Barnes C.W., Hadzi-Pavlovic D., Wilhelm K., Mitchell P.B. (2015) A web-based preventive intervention program for bipolar disorder: outcome of a 12-months randomized controlled trial. *Journal of Affective Disorders*, **174**, 485–492.
- Bauer M., Grof P., Gyulai L., Rasgon N., Glenn T., Whybrow P.C. (2004) Using technology to improve longitudinal studies: self-reporting with ChronoRecord in bipolar disorder. *Bipolar Disorders*, **6**, 67–74.
- Beigel A., Murphy D.L. (1971) Unipolar and bipolar affective illness. Differences in clinical characteristics accompanying depression. *Archives of General Psychiatry*, **24**, 215–220.
- Bopp J.M., Miklowitz D.J., Goodwin G.M., Stevens W., Rendell J.M., Geddes J.R. (2010) The longitudinal course of bipolar disorder as revealed through weekly text messaging: a feasibility study. *Bipolar Disorders*, **12**, 327–334.
- Demitrack M.A., Faries D., Herrera J.M., DeBrotta D., Potter W.Z. (1998) The problem of measurement error in multisite clinical trials. *Psychopharmacology Bulletin*, **34**, 19–24.
- Depp C.A., Mausbach B., Granholm E., Cardenas V., Ben-Zeev D., Patterson T.L., Lebowitz B. D., Jeste D.V. (2010) Mobile interventions for severe mental illness: design and preliminary data from three approaches. *Journal of Nervous and Mental Disease*, **198**, 715–721.
- Depp C.A., Kim D.H., de Dios L.V., Wang V., Ceglowski J. (2012) A pilot study of mood ratings captured by mobile phone versus paper-and-pencil mood charts in bipolar disorder. *Journal of Dual Diagnosis*, **8**, 326–332.
- Depp C.A., Ceglowski J., Wang V.C., Yaghouti F., Mausbach B.T., Thompson W.K., Granholm E.L. (2015) Augmenting psychoeducation with a mobile intervention for bipolar disorder: a randomized controlled trial. *Journal of Affective Disorders*, **174**, 23–30.
- Faurholt-Jepsen M., Brage S., Vinberg M., Kessing L.V. (2016) State related differences in the level of psychomotor activity in patients with bipolar disorder – continuous heart rate and movement monitoring. *Psychiatry Research*, **237**, 166–174.
- Faurholt-Jepsen M., Brage S., Vinberg M., Christensen E.M., Knorr U., Jensen H.M., Kessing L.V. (2012) Differences in psychomotor activity in patients suffering from unipolar and bipolar affective disorder in the remitted or mild/moderate depressive state. *Journal of Affective Disorders*, **141**, 457–463.
- Faurholt-Jepsen M., Vinberg M., Christensen E.M., Frost M., Bardram J., Kessing L.V. (2013) Daily electronic self-monitoring of subjective and objective symptoms in bipolar disorder – the MONARCA trial protocol (MONitoring, treatment and pRediction of bipolar disorder episodes): a randomised controlled single-blind trial. *BMJ Open*, **3**, 7.
- Faurholt-Jepsen M., Frost M., Vinberg M., Christensen E.M., Bardram J.E., Kessing L.V. (2014) Smartphone data as objective measures of bipolar disorder symptoms. *Psychiatry Research*, **217**, 124–127.
- Faurholt-Jepsen M., Vinberg M., Frost M., Christensen E.M., Bardram J.E., Kessing L.V. (2015) Smartphone data as an electronic biomarker of illness activity in bipolar disorder. *Bipolar Disorders*, **17**, 715–728.

- Frost M., Doryab A., Faurholt-Jepsen M., Kessing L.V., Bardram J.E. (2013) Supporting disease insight through data analysis: refinements of the monarca self-assessment system. In Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13), pp. 133–142, ACM: New York.
- Goodwin F.K., Jamison K.R. (1996) *Manic-depressive Illness*, New York: Oxford University Press.
- Hamilton M. (1967) Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology*, **6**, 278–296.
- Husky M.M., Gindre C., Mazure C.M., Brebant C., Nolen-Hoeksema S., Sanacora G., Swendsen J. (2010) Computerized ambulatory monitoring in mood disorders: feasibility, compliance, and reactivity. *Psychiatry Research*, **178**, 440–442.
- Javelot H., Spadazzi A., Weiner L., Garcia S., Gentili C., Kosel M., Bertschy G. (2014) Telemonitoring with respect to mood disorders and information and communication technologies: overview and presentation of the PSYCHE project. *BioMed Research International*, **2014**, 104658.
- Jones S.H., Hare D.J., Evershed K. (2005) Actigraphic assessment of circadian activity and sleep patterns in bipolar disorder. *Bipolar Disorders*, **7**, 176–186.
- Judd L.L., Schettler P.J., Akiskal H., Coryell W., Fawcett J., Fiedorowicz J.G., Solomon D.A., Keller M.B. (2012) Prevalence and clinical significance of subsyndromal manic symptoms, including irritability and psychomotor agitation, during bipolar major depressive episodes. *Journal of Affective Disorders*, **138**, 440–448.
- Kessing L.V., Hansen H.V., Hvenegaard A., Christensen E.M., Dam H., Gluud C., Wetterslev J. (2013) Treatment in a specialised out-patient mood disorder clinic v. standard out-patient treatment in the early course of bipolar disorder: randomised clinical trial. *British Journal of Psychiatry*, **202**, 212–219.
- Krane-Gartiser K., Henriksen T.E.G., Morken G., Vaaler A., Fasmer O.B. (2014) Actigraphic assessment of motor activity in acutely admitted inpatients with bipolar disorder. *PloS One*, **9**, e89574.
- Kuhs H., Reschke D. (1992) Psychomotor activity in unipolar and bipolar depressive patients. *Psychopathology*, **25**, 109–116.
- Kupfer D.J., Weiss B.L., Foster G., Detre T.P., McPartland R. (1974) Psychomotor activity in affective states. *Archives of General Psychiatry*, **30**, 765–768.
- Lauder S., Chester A., Castle D., Dodd S., Gliddon E., Berk L., Chamberlain J., Klein B., Gilbert M., Austin D.W. et al. (2015) A randomized head to head trial of MoodSwings.net.au: an internet based self-help program for bipolar disorder. *Journal of Affective Disorders*, **171**, 13–21.
- Lieberman D.Z., Kelly T.F., Douglas L., Goodwin F.K. (2010) A randomized comparison of on-line and paper mood charts for people with bipolar disorder. *Journal of Affective Disorders*, **124**, 85–89.
- Miklowitz D.J., Price J., Holmes E.A., Rendell J., Bell S., Budge K., Christensen J., Wallace J., Simon J., Armstrong N.M. et al. (2012) Facilitated Integrated Mood Management for adults with bipolar disorder. *Bipolar Disorders*, **14**, 185–197.
- Mitchell P.B., Malhi G.S. (2004) Bipolar depression: phenomenological overview and clinical characteristics. *Bipolar Disorders*, **6**, 530–539.
- Mutschler J., von Zitzewitz F., Rossler W., Grosshans M. (2012) Application of electronic diaries in patients with schizophrenia and bipolar disorders. *Psychiatria Danubina*, **24**, 206–210.
- Psaty B.M., Prentice R.L. (2010) Minimizing bias in randomized trials: the importance of blinding. *JAMA*, **304**, 793–794.
- Todd N.J., Jones S.H., Hart A., Lobban F.A. (2014) A web-based self-management intervention for Bipolar Disorder “living with bipolar”: a feasibility randomised controlled trial. *Journal of Affective Disorders*, **169**, 21–29.
- Weinstock L.M., Miller I.W. (2008) Functional impairment as a predictor of short-term symptom course in bipolar I disorder. *Bipolar Disorders*, **10**, 437–442.
- Wickström G., Bendix T. (2000) The “Hawthorne effect” – what did the original Hawthorne studies actually show? *Scandinavian Journal of Work, Environment & Health*, **26**, 363–367.
- Wing J.K., Babor T., Brugha T., Burke J., Cooper J. E., Giel R., Jablenski A., Regier D., Sartorius N. (1990) SCAN. Schedules for Clinical Assessment in Neuropsychiatry. *Archives of General Psychiatry*, **47**, 589–593.
- Young R.C., Biggs J.T., Ziegler V.E., Meyer D.A. (1978) A rating scale for mania: reliability, validity and sensitivity. *British Journal of Psychiatry*, **133**, 429–435.