Psychological Medicine (2015), **45**, 2691–2704. © Cambridge University Press 2015
doi:10.1017/S0033291715000410

ORIGINAL ARTICLE

Daily electronic self-monitoring in bipolar disorder using smartphones – the MONARCA I trial: a randomized, placebo-controlled, single-blind, parallel group trial

M. Faurholt-Jepsen[1]*, M. Frost[2], C. Ritz[3], E. M. Christensen[1], A. S. Jacoby[1], R. L. Mikkelsen[1], U. Knorr[1], J. E. Bardram[2], M. Vinberg[1] and L. V. Kessing[1]

[1] *The Copenhagen Clinic for Affective Disorder, Psychiatric Centre Copenhagen, Rigshospitalet, Copenhagen, Denmark*
[2] *The Pervasive Interaction Laboratory (PIT Lab), IT University of Copenhagen, Copenhagen, Denmark*
[3] *Department of Basic Sciences and Environment, Faculty of Life Sciences, University of Copenhagen, Copenhagen, Denmark*

Background. The number of studies on electronic self-monitoring in affective disorder and other psychiatric disorders is increasing and indicates high patient acceptance and adherence. Nevertheless, the effect of electronic self-monitoring in patients with bipolar disorder has never been investigated in a randomized controlled trial (RCT). The objective of this trial was to investigate in a RCT whether the use of daily electronic self-monitoring using smartphones reduces depressive and manic symptoms in patients with bipolar disorder.

Method. A total of 78 patients with bipolar disorder according to ICD-10 criteria, aged 18–60 years, and with 17-item Hamilton Depression Rating Scale (HAMD-17) and Young Mania Rating Scale (YMRS) scores ≤17 were randomized to the use of a smartphone for daily self-monitoring including a clinical feedback loop (the intervention group) or to the use of a smartphone for normal communicative purposes (the control group) for 6 months. The primary outcomes were differences in depressive and manic symptoms measured using HAMD-17 and YMRS, respectively, between the intervention and control groups.

Results. Intention-to-treat analyses using linear mixed models showed no significant effects of daily self-monitoring using smartphones on depressive as well as manic symptoms. There was a tendency towards more sustained depressive symptoms in the intervention group ($B = 2.02$, 95% confidence interval −0.13 to 4.17, $p = 0.066$). Sub-group analysis among patients without mixed symptoms and patients with presence of depressive and manic symptoms showed significantly more depressive symptoms and fewer manic symptoms during the trial period in the intervention group.

Conclusions. These results highlight that electronic self-monitoring, although intuitive and appealing, needs critical consideration and further clarification before it is implemented as a clinical tool.

Received 3 November 2014; Revised 10 February 2015; Accepted 10 February 2015; First published online 29 July 2015

Key words: Bipolar disorder, electronic self-monitoring, feedback loop, randomized controlled trial, smartphone, the MONARCA I trial.

Introduction

Bipolar disorder is a long-term and heterogeneous illness with a continued need for treatment and naturalistic follow-up studies suggest that the progressive development of bipolar disorder is not prevented with the present treatment options (Kessing *et al.* 2004; Baldessarini *et al.* 2010). Major reasons for the insufficient effect of present treatment options in clinical practice are delayed intervention for prodromal depressive and manic symptoms as well as decreased adherence to mood stabilizer treatment (Kessing *et al.* 2007; Morriss *et al.* 2007). During the last decade there has been an emerging shift in illness paradigm from a focus on affective episodes in bipolar disorder to an increasing focus on inter-episodic mood instability (MacQueen *et al.* 2003; Bonsall *et al.* 2012). Many patients with bipolar disorder continue to experience subsyndromal mood swings on a daily basis, with euthymic patients with bipolar disorder suffering more from mood instability than healthy subjects (Paykel *et al.* 2006; Henry *et al.* 2008; Bonsall *et al.* 2012). Mood instability at a subclinical level is reported to be associated with impaired global functioning and

* Address for correspondence: Dr M. Faurholt-Jepsen, The Copenhagen Clinic for Affective Disorder, Psychiatric Centre Copenhagen, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark.
(Email: maria@faurholt-jepsen.dk)

high risk of relapse (MacQueen *et al.* 2003; Strejilevich *et al.* 2013).

Different paper-based daily mood charting instruments, such as the National Institute of Mental Health Life Chart Method (NIMH-LCM; Leverich *et al.* 2001), the Systematic Treatment Enhancement Program (Sachs, 2014) and the ChronoSheet (Bauer *et al.* 1991), have been developed and are often used in the treatment of bipolar disorder, thus enabling collection of detailed longitudinal information on daily mood swings and other symptoms relevant to bipolar disorder when the patients are outside the clinical setting. Paper-based mood-charting instruments can be seen as a facilitating tool helping patients with bipolar disorder to gain insight into illness, facilitate patient empowerment, teach patients to recognize early warning signs of recurrence of mania, depression and mixed states and also enable characterization of mood instability in detail. However, several problems regarding paper-based mood-charting instruments have been addressed, such as low compliance and potential recall bias when filling in data retrospectively, i.e. where patients complete batches of daily ratings at a single time (Stone *et al.* 2003; Whybrow *et al.* 2003). Recently, different types of electronic self-monitoring instruments for a variety of illnesses, including bipolar disorder, have been developed. Electronic self-monitoring instruments using computers (Totterdell *et al.* 1996; Whybrow *et al.* 2003), personal digital assistants (Schärer *et al.* 2002; Kreindler *et al.* 2003; Depp *et al.* 2010; Husky *et al.* 2010), text messages (Bopp *et al.* 2010), web interfaces (Lieberman *et al.* 2010), and smartphone applications (Depp *et al.* 2012; Faurholt-Jepsen *et al.* 2014) have been described in the literature, and also a large number of commercial smartphone applications without feedback loop options (feedback on the collected electronic self-monitored data to the clinicians) are available in the App Store and Google Play (e.g. Optimism, MoodRhythm, iMoodJournal, eMoods Bipolar Disorder Tracker, Bipolar Bear, Moody Me, etc.).

Smartphones as a self-monitoring tool offer unique opportunities for continuous and long-term assessment of depressive and manic symptoms and collection of real-time data in naturalistic settings. Furthermore, smartphones eliminate the need for patients to interact with a separate monitoring device because the monitoring application can be installed directly on the smartphone, and since most people carry their cell phone with them during most of the day and use if for normal communicative purposes the risk of stigmatization due to using a separate device for illness monitoring is not present.

An Android smartphone-based self-monitoring system (the MONARCA system) for patients with bipolar disorder including a feedback loop between patient and mental healthcare providers was developed and tested in pilot studies by the authors (Bardram *et al.* 2012, 2013). The MONARCA system included a self-monitoring part where the patients could evaluate their symptoms and a feedback loop. The feedback loop consisted of two levels of loops: (*a*) a feedback loop where the self-monitored data was sent to the clinic allowing for the study nurse to review the data and contact the patients if there were signs of deterioration, thereby allowing for intervention on prodromal depressive and manic symptoms, and (*b*) a feedback loop where the self-monitored data was visualized graphically to the patients themselves providing an overview of the entered data, and thereby providing possibilities for an increased illness insight and understanding. Patients using the MONARCA system found it acceptable to use and adherence to self-monitoring was higher than when monitoring using a paper-based version (Bardram *et al.* 2012, 2013). It has not previously been evaluated in a randomized controlled trial (RCT) whether daily electronic self-monitoring, including a feedback loop from the patient to the clinic and back to the patient and also a feedback loop from the patient back to him/herself, has an effect on depressive and manic symptoms in patients with bipolar disorder. We conducted a RCT, the MONARCA I trial, to evaluate the hypothesis that daily electronic self-monitoring using the MONARCA system developed for Android smartphones, including a clinical feedback loop, in patients with bipolar disorder would reduce depressive and manic symptoms.

Method

The trial is reported according to the CONsolidated Standards Of Reporting Trials and non-pharmacological treatment (CONSORT) statement (Boutron *et al.* 2008; Moher *et al.* 2010). The design of the trial has been described in detail elsewhere (Faurholt-Jepsen *et al.* 2013), thus this paper provides a summary of the design and methods. No changes in methods were made after trial commencement.

Study design, setting and participants

The trial was conducted using a randomized, placebo-controlled, single-blinded, parallel group design. A total of 78 patients were recruited from The Copenhagen Clinic for Affective Disorders, Psychiatric Centre Copenhagen, Rigshospitalet, Denmark, during the period from September 2011 to March 2013 (Fig. 1). The clinic is a specialized outpatient clinic with a catchment area consisting of the

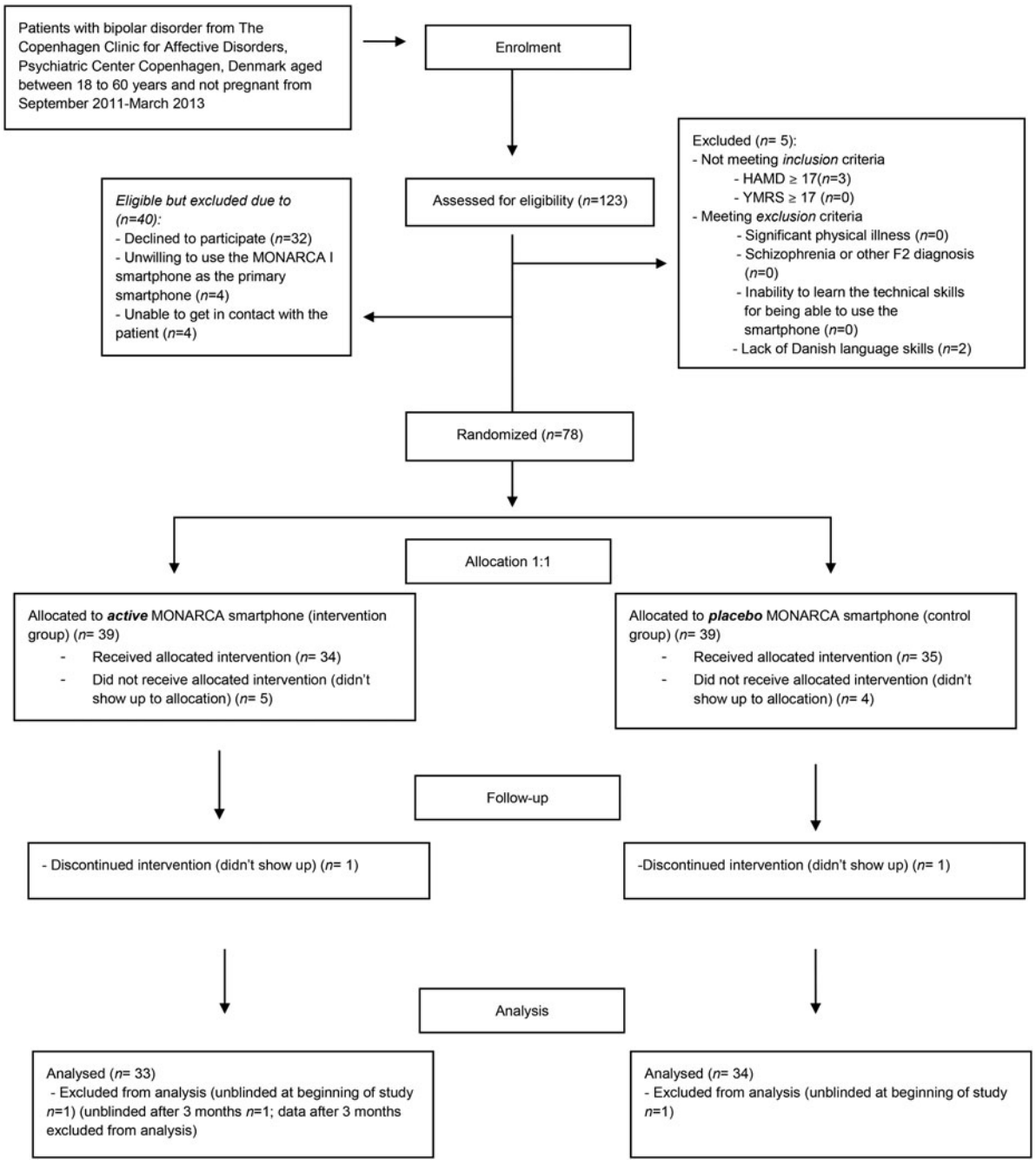

Fig. 1. Study flow chart.

Capital Region in Denmark corresponding to 1.4 million people. Patients with a new diagnosis of bipolar disorder or with treatment-resistant bipolar disorder were referred to the clinic. The staff consists of specialists in psychiatry, psychologists, nurses, and a social worker, all with specific experience and knowledge regarding bipolar disorder. Treatment at the clinic comprises 2 years of combined evidence-based psychopharmacological treatment and supporting therapy, including group psychoeducation. Further details about the treatment programme in The Copenhagen Clinic for Affective Disorders and the effect of this are described elsewhere (Kessing *et al.* 2013).

Patients were invited to participate in the trial following referral to the clinic. Patients were followed by a specialist in psychiatry on a regular basis and did not join psychoeducation groups until after the end of participation in the trial.

Inclusion criteria were: a diagnosis of bipolar disorder according to ICD-10 criteria using the

Schedules for Clinical Assessment in Neuropsychiatry (SCAN) interview (Wing *et al.* 1990), aged between 18 and 60 years, a 17-item Hamilton Depression Rating Scale (HAMD-17) score ≤17 (Hamilton, 1967) and a Young Mania Rating Scale (YMRS) score ≤17 (Young *et al.* 1978) at the time of inclusion.

We excluded patients who were pregnant, who had a lack of Danish language skills, who were unable to learn the technicalities for using a smartphone, who were unwilling to use the trial smartphone as their primary cell phone, and who were severely physically ill or had schizophrenia, schizotypal or delusional disorders according to the SCAN interview.

Randomization, allocation and implementation

Patients were randomized with a balanced ratio of 1:1 to receive either an intervention Android smartphone (the intervention group) or a control Android smartphone (the control group) for a 6-month trial period. A computer-generated list of random allocation numbers was obtained by an independent researcher using http://www.randomisation.com. Randomization was stratified on age (<29 or ≥29 years) and former hospitalization (yes/no) since these were considered to be possible prognostic variables, and a fixed block size of 10 within each stratum was used. The fixed block size was unknown to the clinicians recruiting the patients and the study nurse allocating patients to their randomization group. The allocation sequence was concealed in numbered and opaque envelopes in a locked cabinet of unknown location from the researchers (M.F.J. and A.S.J.) enrolling and assessing the patients.

Blinding

Due to the type of intervention, this trial was single-blinded since blinding of the patients, the clinicians, and the study nurse handling the intervention was not possible. Thus, the patients, the clinicians and the study nurse knew whether a patient was randomized to either the intervention group or the control group. The researchers performing the outcome assessments, data entry, data management, data analyses and interpretation of data were kept blinded to the patients' randomization status at all times during the trial. The trial was therefore single-blinded (researcher-blinded). Four researchers (M.F.J., L.V.K., M.V., C.R.) performed all statistical analyses and interpretations of these analyses before data were unmasked.

Intervention

The intervention group. Patients randomized to the intervention group were provided with a smartphone with the MONARCA system installed (Bardram *et al.* 2013) and were instructed to use the system for self-monitoring on a daily basis during the trial period. The patients were prompted at a self-chosen time during the day by the MONARCA system to evaluate the following subjective items: mood (scored from depressive to manic on a scale from −3 to +3), sleep length (number of hours slept per night), medication taken [yes/no/changes (the patients were asked to specify these)], activity (scored on a scale from −3 to +3), irritability (yes/no), mixed mood (yes/no), cognitive problems (yes/no), alcohol consumption (number of units consumed per day), stress (scored on a scale from 0 to 5), menstruation (yes/no) for women and individualized early warning signs (yes/no). If a patient forgot to enter the items in the MONARCA system it was possible to enter data retrospectively for a maximum of 2 days. It was then noted in the system that data were collected retrospectively.

The MONARCA system also included a feedback loop. The feedback loop comprising two levels was a part of the intervention: (*a*) a feedback loop between the patients and the clinic, where the self-monitored data was sent to the clinic allowing for the study nurse to review the data and contact the patients if there were signs of deterioration, thereby allowing for intervention on prodromal depressive and manic symptoms, and (*b*) a feedback loop within the patients themselves, where the self-monitored data was visualized graphically to the patients providing them with an overview of the entered data, thereby providing possibilities for an increased illness insight and understanding.

As part of the first level of the feedback loop a study nurse with experience in bipolar disorder reviewed the self-reported data from the MONARCA system on a daily basis and if there were signs of deterioration of depressive or manic symptoms, such as changes in sleep length, mood rated as lowered or elevated, changes in adherence to medication, the patients were contacted by text message, phone call or e-mail. The patients could also contact the study nurse in case of deterioration. When contact was established the study nurse then clarified whether the entered data was a sign of depressive or manic symptoms and discussed possible actions to be taken on this matter (e.g. discussing coping strategies, reduce or increase the amount of activities, contact relatives or friends for support, take some extra medication, visit the study nurse for a supporting conversation, visit the clinic for a conversation with their psychiatrist, attend the psychiatric emergency room).

The control group. Patients who were randomized to the control group of the trial were provided with a smartphone without the MONARCA system, which they

had to use for normal communicative purposes during the trial period (a placebo smartphone). This was chosen to control for any effect on depressive or manic symptoms due to simply receiving a smartphone as part of participating in the trial, and also to avoid the risk of unblinding of the researchers due to simply seeing the patients' smartphones.

All patients received treatment at The Copenhagen Clinic for Affective Disorders, Psychiatric Centre Copenhagen, Rigshospitalet, Denmark during the trial period as described above. In addition, patients in the control group were offered a nurse contact in accordance with the treatment programme in the clinic (Faurholt-Jepsen *et al.* 2013).

Thus, the active intervention consisted, in addition to treatment as usual, of the MONARCA system comprising a smartphone app for self-monitoring plus a two-level feedback loop that included contact with a study nurse. The control intervention consisted of treatment as usual including contact with a study nurse plus using a smartphone for normal communicative purposes.

All patients included in the trial were offered the loan of a smartphone free of charge during the trial period and costs due to data traffic were refunded. Thus, patients who did not own a smartphone were provided with one by the trial, and recruitment obstacles due to this issue were therefore not present. The smartphones used in this trial were 'HTC Desire S' which runs on the Android operating system and patients used their own SIM card.

Outcomes and assessments

Primary outcome measures were clinically rated depressive and manic symptoms during the 6-month trial period based on the HAMD-17 (Hamilton, 1967) for depressive symptoms and the YMRS (Young *et al.* 1978) for manic symptoms. The clinical ratings were performed monthly for the entire trial period.

Secondary outcomes that were assessed concurrently with the monthly assessment: Cohen's Perceived Stress Scale (PSS; Cohen *et al.* 1983) as a measure of subjective perceived stress, Functioning Assessment Short test (FAST; Rosa *et al.* 2007) as a measure of psychosocial functioning, The World Health Organization Quality of Life – short version (WHOQOL-BREF; WHO, 1998) as a measures of quality of life, Coping Inventory for Stressful Situations (CISS; Endler & Parker, 1998) as a measure of coping strategies in stressful situations, Major Depression Inventory (MDI; Bech & Olsen, 2001; Bech *et al.* 2001; Olsen *et al.* 2003) to evaluate the severity of self-rated depressive symptoms, Altman Self-Rating Mania Scale (ASRM; Altman *et al.* 1997) to evaluate the severity of self-rated manic symptoms, and the Massachusetts General Hospital Cognitive and Physical Functioning Questionnaire (MASS) as a measure of self-rated cognitive and physical functioning (Fava *et al.* 2009).

Tertiary outcomes that were assessed by the researchers at baseline, after 3 months, and after 6 months were: cognitive function according to the Screen for Cognitive Impairment in Psychiatry (SCIP-S; Guilera *et al.* 2009; Rojo *et al.* 2010) and adherence to prescribed psychopharmacological treatment measured by plasma concentration of medication (mood stabilizers, antipsychotics, antidepressants).

All patients were invited for a monthly visit with the study researchers for outcome assessment. All outcome assessments were conducted by two physicians (M.F.J. or A.S.J.) who were not involved in the treatment of the patients and who were blinded to the patients' group of randomization. No changes were made to trial outcomes after commencement of the trial.

Statistical methods and sample size

The sample size required to detect a clinically relevant difference between the intervention group and the control group defined as a minimum of 3 points on the HAMD-17 and YMRS assuming a standard deviation (S.D.) of 4, a significance level of 0.05, and a power of 0.80 was found to be 28 patients in each of the two groups (sample size calculation for a two-sample t test).

The randomization was stratified according to the two variables: (1) age <29 years or ≥29 years and (2) former hospitalization (yes/no). The reasons for choosing these possible prognostic variables related to the effect of the intervention were: (1) it may be that patients aged <29 years to a larger extent than patients aged ≥29 years have grown up using mobile phones and therefore react differently to electronic self-monitoring, and (2) as former hospitalization is believed to reflect the severity of bipolar disorder perhaps this issue could influence the effect of electronic self-monitoring.

Analyses were conducted as intention-to-treat, including all patients who attended for randomization. Two-sample t tests and χ^2 tests (and Fisher's exact test where appropriate) were used to assess differences in means and differences in proportions, respectively. Model assumptions were checked visually by means of residuals and QQ plots and, logarithmic transformation was applied to the YMRS to ensure model assumptions were met. Back-transformations of treatment differences on the YMRS were done using methods described by others (Laursen *et al.* 2014). For each outcome we specified a two-level linear mixed model with intervention/control group as a fixed effect and included a patient-specific random effect. Linear mixed models allow for both between-individual and

within-individual variations of the specific outcomes over several time-points and takes missing values into account. We first considered an unadjusted model and second, a model adjusted for sex, the two stratification variables, age (<29 or ≥29 years) and former hospitalization (yes/no). Each model, including the unadjusted model, included the baseline values of the outcome variable as a covariate, controlling for potential baseline differences between the two randomization groups. Interaction between group of randomization and visit number (time) on the outcome variable was tested in all analyses.

We considered 3 + 3 exploratory subgroup analyses using linear mixed models in relation to the primary outcomes in Table 2. Models 1A and 2A: models with depressive symptoms or manic symptoms, respectively, as the outcome measure excluding values with mixed depressive and manic symptoms, i.e. model 1A: HAMD-17 scores in which the HAMD-17 values were excluded at a given time point if the YMRS score was ≥7 at the given time point, and model 2A: YMRS scores in which the YMRS values were excluded at a given time point if the HAMD-17 score was ≥7 at the given time point. These values were chosen since we wanted to investigate the effect of the intervention in patients with remission of mixed symptoms. Models 1B and 2B: only scores on the HAMD-17 of >0 or the YMRS of >0 at any time point, respectively, were included. These values were chosen since we wanted to investigate the effect of the intervention if patients were presenting depressive or manic symptoms. Models 1C and 2C: only patients who presented with depressive or manic symptoms at baseline were included, i.e. only patients with a HAMD-17 score >7 or a YMRS score >7 at baseline. These values were chosen since we wanted to investigate the effect of the intervention if patients were presenting some level of depressive or manic symptoms at the beginning of the study.

For analyses on adherence to prescribed psychopharmacological treatment plasma concentration of medication was dichotomized to adherent/non-adherent. If the serum level was below the detection limit in patients for whom the drug was prescribed, the serum level was given the value 0, otherwise 1.

Data were entered using the data entry program Epidata® (The EpiData Association, Denmark) and the statistical software program Stata version 12.1 (StataCorp LP, USA) was used for the statistical analyses. The significance level of the p values in the statistical analyses was set at 0.05 (two-tailed).

Ethical considerations

The trial 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). Prior to trial commencement the trial was registered at ClinicalTrials.gov (NCT01446406). Electronic data collected from smartphones were stored at a secure server at Concern IT, Capital Region, Denmark (I-suite number RHP-2011-03). Smartphones were loaned to all patients free of charge during the trial and economic costs due to data traffic were refunded. Written and oral information was given to potential participants before informed consent was obtained, and participants were informed that they could withdraw from the trial at any time during the trial this without any consequence for their treatment. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Results

Patient flow and background characteristics

In the period from September 2011 to March 2013, 123 potential participants with bipolar disorder, aged 18–60 years and not pregnant were assessed for eligibility. Among these, 78 patients (63.4%) were included in the trial (Fig. 1). The last patient visit was in September 2013. Of the eligible patients three were excluded due to not meeting inclusion criteria (persistent HAMD-17 scores ≥17), two because exclusion criteria were met (lack of Danish-language skills), and 32 patients declined participation for various reasons (e.g. did not want to participate in the trial, expected participation to be too time consuming), four patients were unwilling to use the trial smartphone as their primary cell phone and four patients were inaccessible by both phone calls and letters. Among the 78 included patients who were randomized with 39 patients in each group, five patients in the intervention group and four patients in the control group did not attend the baseline assessment and randomization and therefore did not receive the allocated intervention; the researchers were unblinded to group of randomization of two patients during the trial, thus leaving a total of 67 patients for further intention-to-treat analyses. Two patients, one in the intervention group (due to travelling abroad) and one in the control group (due to not responding to phone calls and letters and not attending scheduled visits), had a discontinued intervention after 3 months follow-up. A total of 3.7% of patient visits were missing (3.6% in the intervention group and 3.8% in the control group) due to patients not attending.

During the 6-month trial period 93.03% (S.D. = 15.6) of patients randomized to the intervention group evaluated the subjective items in the MONARCA system on a daily basis. The MONARCA system was designed to allow for retrospective entries of self-monitoring for a maximum period of up to 2 days and 7.15% (S.D. = 7.79) of the subjective entries were done retrospectively. On average 85 text messages were sent and 11 phone calls were made to each patient, and all patients had contact with the study nurse during the trial period. In the first half part of the trial period more contacts between the study nurse and the patients were via phone calls (22.6%) than during the second part (11.1%). The study nurse experienced that in the beginning of the patients' trial period contact by phone calls was more feasible to get to know the patients' situation and needs. Later in the patients' trial period a text message was just as sufficient as a phone call and more feasible since it was not perceived as disruptive for the patients as a phone call.

Clinical and socio-demographic characteristics of the patients at baseline are presented in Table 1. The vast majority of patients had a new diagnosis of bipolar disorder. Overall, there were no significant differences in baseline characteristics between the intervention group and the control group. Regardless of randomization group, the mean age was 29.3 years (S.D. = 8.43), 67.1% ($n = 45$) were women, the mean number of years of education after primary school was 4.32 years (S.D. = 2.69) and 67.1% ($n = 45$) had a bipolar disorder type I diagnosis. The patients received flexible psychopharmacological treatment as per international guidelines during the trial period.

Table 1. *Clinical and socio-demographic characteristics of study participants at baseline (N = 67)*

	Intervention group ($n = 33$)	Control group ($n = 34$)
Socio-demographic data		
Age, years	29.1 (7.5)	29.5 (9.4)
Female sex, % (n)	65.7 (22)	68.6 (23)
In relationship, % (n)	54.3 (18)	42.9 (15)
Employed, % (n)	14.3 (5)	19.9 (7)
Student, % (n)	48.6 (16)	40.0 (14)
Children, number	0.40 (0.85)	0.46 (0.85)
Educational level		
Primary school or lower, % (n)	22.9 (8)	20.0 (7)
High school, % (n)	40.0 (13)	42.9 (15)
University undergraduate or more, % (n)	28.6 (9)	32.3 (11)
Years of education after primary school	4.2 (2.7)	4.5 (2.7)
Clinical history		
BP I diagnosis[a], % (n)	60.0 (20)	74.3 (25)
Admissions, number	1 [1–2]	1 [1–2]
Depressive episodes, number	4 [2–10]	4 [2–5]
Manic episodes, number	3 [2–6]	2 [1–5]
HAMD-17 baseline	9 [4–16]	8 [1–13]
YMRS baseline	2 [0–7]	2 [0–5]
First degree relative with affective disorder, % (n)	32.4 (11)	51.4 (18)
First degree relative with other mental illness, % (n)	5.9 (2)	2.9 (1)

HAMD-17, Hamilton Depression Rating Scale – 17 items; YMRS, Young Mania Rating Scale.

Data are mean (S.D.), median [IQR] or % (n) unless otherwise stated.

[a] Bipolar disorder type I disorder.

Primary outcomes

Differences between the intervention group and the control group in clinically rated depressive and manic symptoms based on HAMD-17 and YMRS scores are presented in the upper part of Table 2. The intervention group showed a trend towards more depressive symptoms throughout the trial period according to HAMD-17 scores compared to the control group [$B = 2.02$, 95% confidence interval (CI) −0.13 to 4.17, $p = 0.066$] (Fig. 2). Restricting the outcome measure to analysis on the 6-item HAMD (Lecrubier & Bech, 2007) only changed the results marginally (unadjusted model: $B = 1.08$, 95% CI −0.20 to 2.37, $p = 0.099$; adjusted model: $B = 1.08$, 95% CI −0.16 to 2.33, $p = 0.088$). The YMRS scores were logarithm-transformed to ensure model assumptions were met and estimated differences between the intervention and control groups were back-transformed as suggested by Laursen *et al.* (2014). No difference in manic symptoms according to the YMRS was observed between the intervention and control groups ($B = -0.34$, 95% CI −1.14 to 0.47, $p = 0.41$) (Fig. 3). Analyses on interactions between group of randomization and visit number (time) in the analyses on both the HAMD-17 and YMRS were non-significant, and are therefore not presented.

Exploratory analyses in relation to the primary outcomes

Results of these analyses are presented in Table 2. Excluding mixed depressive symptoms and mixed manic symptoms, patients in the intervention group experienced significantly more depressive symptoms compared to the control group in both the unadjusted ($B = 2.33$, 95% CI 0.10–4.56, $p = 0.040$) and the adjusted ($B = 2.57$, 95% CI 0.40–4.74, $p = 0.020$) models ($n = 50$)

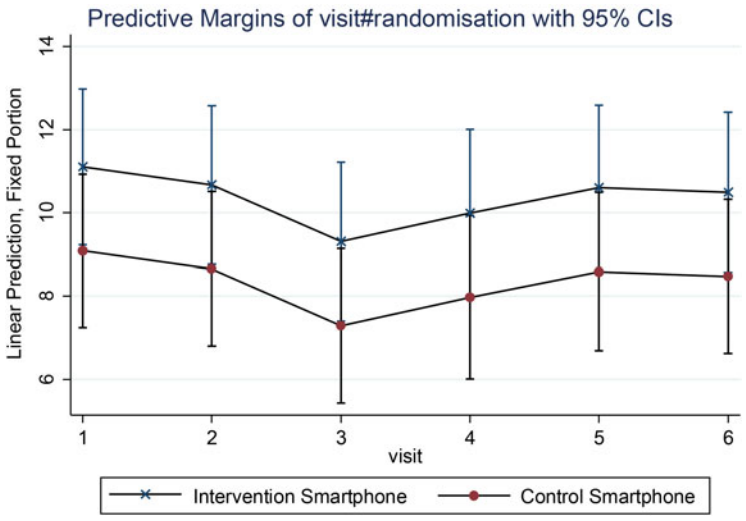

Fig. 2. Differences in depressive symptoms (HAMD-17 score) between the intervention group (–×–) and the control group (–●–) over 6 months (adjusted for HAMD-17 at baseline, previous hospitalization yes/no, age ≥29 or <29 years and sex), $n = 67$.

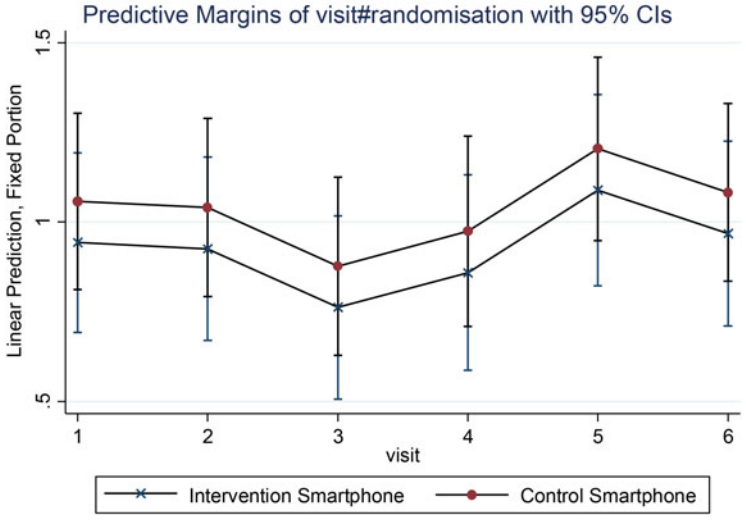

Fig. 3. Differences in manic symptoms (YMRS score) between the intervention group (–×–) and the control group (–●–) over 6 months (adjusted for YMRS at baseline, previous hospitalization yes/no, age ≥29 or <29 years and sex), $n = 67$.

(model 1A). When including only patients with a HAMD-17 score >7 at baseline, patients in the intervention group experienced significantly more depressive symptoms compared to the control group (adjusted model: $B = 2.69$, 95% CI 0.001–5.37, $p = 0.049$) ($n = 38$) (model 1C).

However, excluding mixed depressive symptoms and mixed manic symptom, patients in the intervention group experienced borderline significantly fewer manic symptoms compared to the control group in the adjusted model ($B = -1.07$, 95% CI −2.15 to 0.005, $p = 0.051$) ($n = 45$) (model 2A). When including only patients who presented with manic symptoms at least at one assessment time point during the trial period, patients in the intervention group experienced significantly fewer manic symptoms compared to the control group in both the unadjusted ($B = -0.88$, 95% CI −1.71 to −0.056, $p = 0.036$) and the adjusted ($B = -0.98$, 95% CI −1.80 to −0.16, $p = 0.019$) models ($n = 59$) (model 2B). When including only patients who presented with manic symptoms at baseline, patients in the intervention group experienced significantly fewer manic symptoms compared to the control group in both the unadjusted ($B = -4.20$, 95% CI −7.09 to −0.96, $p = 0.010$) and the adjusted ($B = -6.32$, 95% CI −9.21 to −3.34, $p < 0.001$) models ($n = 13$) (model 2C).

Table 2. *Estimated differences in outcomes between the intervention and control groups (control group serves as reference) (N = 67)*

	Unadjusted[a]			Adjusted[b]		
	Differences	95% CI	*p*	Differences	95% CI	*p*
HAMD-17	1.98	−0.24 to 4.19	0.080	2.02	−0.13 to 4.17	0.066
Sub-analyses on HAMD-17						
1A. HAMD-17 items, no mixed symptoms (*n* = 50)[c]	2.33	0.10 to 4.56	0.040	2.57	0.40 to 4.74	0.020
1B. HAMD-17 items >0 at any time point (*n* = 67)	1.55	−0.56 to 3.67	0.15	1.68	−0.37 to 3.74	0.11
1C. HAMD-17 items baseline >7 (*n* = 39)	2.45	−0.25 to 5.16	0.070	2.69	0.001 to 5.37	0.049
HAMD-17 items baseline £7 (*n* = 28)	0.75	−3.37 to 4.87	0.72	0.25	−3.65 to 4.15	0.90
YMRS	−0.24	−1.07 to 0.58	0.56	−0.34	−1.14 to 0.47	0.41
Sub-analyses on YMRS						
2A.YMRS, no mixed symptoms (*n* = 45)[d]	−0.96	−2.08 to 0.17	0.090	−1.07	−2.15 to 0.005	0.051
2B. YMRS>0 at any time point (*n* = 59)[e]	−0.88	−1.71 to −0.056	0.036	−0.98	−1.80 to −0.16	0.019
2C. YMRS baseline >7 (*n* = 14)	−4.02	−7.09 to −0.96	0.010	−6.32	−9.21 to −3.34	<0.001
YMRS baseline £7 (*n* = 53)	−0.58	−1.43 to 0.27	0.18	−0.78	−1.64 to 0.074	0.074
PSS	2.37	−0.39 to 5.13	0.092	2.40	−0.33 to 5.13	0.085
WHOQOL-BREF, quality of life	−1.35	−5.49 to 2.78	0.52	−1.24	−5.18 to 2.70	0.54
MDI	2.27	−1.72 to 6.27	0.27	2.28	−1.60 to 6.17	0.25
ASRM	−0.098	−0.81 to 0.61	0.79	−0.11	−0.78 to 0.55	0.74
FAST	1.20	−4.27 to 6.66	0.67	0.96	−4.36 to 6.28	0.72
MASS	0.38	−1.90 to 2.66	0.75	0.46	−1.80 to 2.71	0.69
CISS						
Task-oriented coping	−2.89	−6.6 to 0.88	0.13	−2.77	−6.58 to 1.04	0.15
Emotion-oriented coping	1.05	−2.58 to 4.69	0.57	1.14	−2.46 to 4.73	0.54
Avoidance-oriented coping	1.45	−0.80 to 3.69	0.21	1.72	−0.41 to 3.86	0.11
Distraction-oriented coping	1.75	0.20 to 3.20	0.018	1.77	0.31 to 3.22	0.017
Social diversion-oriented coping	0.020	−1.37 to 1.41	0.98	0.16	−1.13 to 1.45	0.81
SCIP-S	0.75	−4.83 to 6.32	0.79	0.53	−5.00 to 6.06	0.85

CI, Confidence interval; HAMD-17, 17-item Hamilton Depression Rating Scale; YMRS, Young Mania Rating Scale; PSS, Cohen's Perceived Stress Scale; WHOQOL-BREF, World Health Organization Quality of Life – short version; MDI, Major Depression Inventory; ASRM, Altman Self-Rating Mania Scale; FAST, Functioning Assessment Short test; MASS, Massachusetts General Hospital Cognitive and Physical Functioning Questionnaire; CISS, Coping Strategies in Stressful Situations; SCIP-S, Screen for Cognitive Impairment in Psychiatry.

[a] Adjusted for outcome variable at baseline

[b] Adjusted for outcome variable at baseline, previous hospitalization yes/no, age ≥29 or <29 years and sex unless otherwise specified

[c] Defined as YMRS ≥7.

[d] Defined as HAMD ≥7.

[e] Indicating analysis on patients with manic symptoms only (YMRS score >0).

Secondary and tertiary outcomes

Differences between the intervention and control groups on the secondary and tertiary outcomes are presented in the lower part of Table 2. In both the unadjusted and adjusted analyses there was a tendency towards a higher score of subjectively perceived stress (PSS) in the intervention group compared to the control group (adjusted model: $B = 2.40$, 95% CI −0.33 to 5.13, $p = 0.085$). Additional adjustment for the HAMD-17 resulted in a non-significant effect ($B = 0.72$, 95% CI −1.17 to 2.61, $p = 0.45$), suggesting that the depressive symptoms explained the higher amount of subjectively perceived stress in the intervention group. In both the unadjusted and the adjusted analyses there were no differences between the intervention and control groups in quality of life (WHOQOL-BREF), self-rated depressive symptoms (MDI), self-rated manic symptoms (ASRM), psychosocial functioning (FAST), self-rated cognitive function (MASS), or cognitive function (SCIP-S). Patients in the intervention group scored significantly higher on the self-reported measure of distraction-oriented coping (subscale on the CISS) compared to the control group in both the unadjusted ($B = 1.75$, 95% CI 0.20–3.20, $p = 0.018$) and the adjusted ($B = 1.77$, 95% CI 0.31–3.22, $p = 0.017$) models. Analysis on adherence

to prescribed psychopharmacological treatment did not reveal any differences in proportions of adherence/non-adherence between the intervention and control groups at baseline (lithium: $p = 0.49$, $n = 64$; lamotrigine: $p = 0.21$, $n = 56$; quetiapine: $p = 0.16$, $n = 48$) or at 6 months follow-up (lithium: $p = 0.22$, $n = 29$; lamotrigine: $p = 0.43$, $n = 26$; quetiapine: $p = 0.99$, $n = 17$). A total of 15 patients (eight in the intervention group, seven in the control group) at baseline and four (two in the intervention group, two in the control group) at 6-months follow-up received antidepressants.

Discussion

This is the first RCT investigating the effect of daily electronic self-monitoring using smartphones, including a two-level feedback loop, in patients with a psychiatric disorder. Patients with bipolar disorder were randomized to a smartphone for daily electronic self-monitoring including a two-level feedback loop (the MONARCA system) *v.* a placebo smartphone without the MONARCA system for a 6-month trial period. We found no significant effect of the intervention in relation to the primary outcomes that were defined as differences in depressive and manic symptoms according to scores on the HAMD-17 and YMRS, or in relation to perceived stress, quality of life, self-assessed depressive or manic symptoms or in relation to cognitive function (secondary and tertiary outcomes). However, exploratory analyses in relation to the primary outcomes revealed interesting findings suggesting that electronic daily self-monitoring including a two-level feedback loop may in fact sustain depressive symptoms but improve manic symptoms among patients who are not in remission or who present with mixed symptoms. Thus, including values with depressive symptoms only (i.e. excluding mixed depressive and manic symptoms) or including only patients with a HAMD-17 score of >7 at baseline, resulted in patients in the intervention group experiencing significantly more depressive symptoms during the trial period corresponding to around 2.6 points on the HAMD-17. By contrast, including only patients who presented with manic symptoms at least at one assessment time point during the trial period or patients who presented with manic symptoms at baseline resulted in patients in the intervention group experiencing fewer manic symptoms, corresponding to 6.3 scores on the YMRS in the model only including patients with a YMRS score of >7 at baseline. We find that these effect sizes in our exploratory analyses are of clinical relevance in the maintenance phase, in accordance with previous suggestions of 3 points on the HAMD-17 by the NICE Committee (2004) and correspondingly 3 points on the YMRS.

Since this is the first RCT investigating the effect of electronic self-monitoring using smartphones, including a two-level feedback loop, we believe that although the overall effect of the intervention was not significant, the findings from exploratory analyses provide important knowledge that can guide future studies and design of interventions within electronic monitoring in psychiatric research.

Interestingly, although studies on the effect of psychological interventions in affective phases in patients with bipolar disorder are very few, our findings are in accordance with the sparse literature revealing differential effectiveness of psychological interventions for manic and depressive phases (Scott & Colom, 2008). One study showed that teaching patients with bipolar disorder to identify early warning signs of relapse and to seek prompt treatment reduced manic relapses, but had no effect on depressive relapses (Perry *et al.* 1999). Similarly, in another randomized trial, a systematic multicomponent care management programme consisting of continuously monitoring and feedback showed an effect on manic frequency and severity but not on depressive symptoms (Simon *et al.* 2006). This is in line with a review discussing that the reasons for these differential effects of psychological interventions on manic and depressive symptoms are not clear (Scott & Colom, 2008). Thus, manic prodromes are more distinct and may be easier to detect and manic symptoms are treated more quickly and effectively with pharmacotherapy than depressive episodes (Perry *et al.* 1999). On the contrary, depressive symptoms are more difficult to differentiate from normal day-to-day problems and may have a more gradual onset and prolonged duration (Simon *et al.* 2006).

Additionally, it is possible that in the present trial daily self-monitoring of ongoing depressive symptoms may have sustained depressive symptoms due to negative processing bias induced by the daily confrontation with the depressive symptoms and, in turn, may have induced or increased fear of not recovering. Patients in the prodromal phase of depression may experience depressive ruminations with further risk of worsening and escalation of depressive symptoms (Perry *et al.* 1999; Morriss *et al.* 2007).

Several studies suggest that illness insight in bipolar disorder is state-dependent rather than trait-dependent, that illness insight is more impaired during mania than depression and that impaired illness insight is associated with lower adherence to psychopharmacological treatment (Ghaemi & Rosenquist, 2004; Yen *et al.* 2007; Látalová, 2012; De Assis da Silva *et al.* 2015). Psychoeducation aims at increasing the patients' awareness of their illness and understanding of their disorder and to teach patients methods to identify early warning signs, thereby increasing

adherence and preventing affective episodes (Colom *et al*. 2003). The MONARCA system shares some common properties to those of psychoeducation and aims at increasing the patients' awareness of prodromal depressive and manic symptoms. Thus, daily electronic self-monitoring of manic symptoms including a two-level feedback loop may help to teach patients' to correct unrealistic hypomanic/manic thoughts and behaviours and in this way increase insight into the hypomanic/manic state and increase acceptance and adherence to medication. Interventions focusing primarily on teaching individuals to recognize and intervene against early warning signs such as in our trial (e.g. behaviour change or increases in medication) may prevent isolated manic symptoms from cascading into a full-blown manic relapse, but may be less effective at identifying and intervening against a depressive prodrome. Thus, we believe that this study supports the hypothesis that it can be difficult for an intervention to have an effect on both depressive and manic symptoms given the complexity of bipolar disorder. It may be that the MONARCA system in the version investigated in this trial is effective in recognizing and allowing for intervention on early warning signs of hypomania/mania, but less effective in relation to early warning signs of depression. Emphasis on the differentiation of day-to-day problems and depressive symptoms should be a high priority and perhaps a positive reinforcing feedback mechanisms could help minimize the negative processing bias and thereby the sustained depressive symptoms.

Limitations

The overall finding from this trial is negative. A possible explanation for this could be that there were no differences concerning adherence to medication and patients randomized to the control group also received a well-defined intervention programme of combined evidence-based psychopharmacological treatment and supportive therapy when indicated, thus making it harder for any additional intervention to improve course of illness and treatment outcome any further. Another explanation may be that patients were included in a remitted or partially remitted phase with relatively few depressive and manic symptoms at baseline, thus making it harder for any intervention to have an effect on depressive and manic symptoms (Table 1). Nevertheless, we aimed to include patients who were able to manage the technical aspects of using a smartphone for detection of early warning signs of upcoming episodes and hypothesized that including patients with high levels of depressive and manic symptoms would complicate this process. It is possible that including patients with more depressive and/or manic symptoms may have resulted in other findings.

During the trial period the mean adherence to self-monitoring was >93%. As a part of the feedback loop the study nurse contacted the patients in case of non-adherence to the self-monitoring in the MONARCA system since this was interpreted as possible deterioration and presence of depressive or manic symptoms and thereby an indication for her to contact the patient. The high level of adherence is believed to reflect the high usability and low level of intrusiveness from the MONARCA system, a factor contributing to a high motivation from the patients. Moreover, interviews with the patients and the study nurse revealed that they found the MONARCA system easy to use with a high appeal and usability (Frost *et al*. in press).

Due to the type of intervention in this trial it was not possible to mask the patients, the clinicians or the study nurse to the allocated group of randomization, but the researchers performing the outcome assessments were blinded and thus the trial was single-blinded. All patients received standard treatment at The Copenhagen Clinic for Affective Disorder, Psychiatric Centre Copenhagen, Rigshospitalet, Denmark during the trial period, and, in addition, the patients in the intervention group received a combination of daily electronic self-monitoring and a feedback loop between the themselves and the clinic. Thus, the investigation of the effect is of a 'total MONARCA system' consisting of all the components in the intervention, and it is not possible from this trial to distinguish between the effect of these different aspects of the intervention.

In any non-pharmacological trial it is always difficult to define a proper control group. In this trial we decided to include a control group of patients who received a control smartphone for normal communicative purposes, but without the MONARCA system, avoiding the estimated effect that the intervention on depressive and manic symptoms was simply due to receiving a smartphone for cost-free communication.

Patients in the intervention group scored higher on a measure of distraction-oriented coping [a subscale on the CISS (Endler & Parker, 1998)] compared to the control group. Perhaps the intervention was able to change the patients coping style in some situations, but the follow-up period in the present trial was too short to teach the patients to distract themselves from their depressive symptoms or other stressful situations which probably would require a longer trial period teaching patients a more permanent change in coping mechanisms. Alternatively, this could be a chance finding.

Generalizability of the results

Although many different electronic self-monitoring instruments for patients with bipolar disorder (and

for a variety of other illnesses) have been developed, it has not previously been investigated whether daily electronic self-monitoring including a two-level feedback loop reduces symptoms of bipolar disorder (or any other psychiatric disorder). The feedback loop in the MONARCA system consisted of two levels of loops, with one being between the patients and the clinic and the other being within the patients themselves without the assistance of clinicians. Thus, the study investigated the effect of both levels of the feedback loop.

We realize that this trial was perfprmed in a tertiary, highly specialized mood disorder clinic, thus making applicability in a wide range of usual care settings more difficult. However, the trial had a pragmatic design with few exclusion criteria and few patients were excluded. Furthermore, the majority of patients entering the trial were in an early course of illness with a new diagnosis of bipolar disorder, and the MONARCA system was easy to use for both the patients and the clinicians with a high appeal and a very low dropout rate (3.7%) during the 6-month trial. Furthermore, all patients were offered the loan of a smartphone free of charge during the trial period thus eliminating any economic issues regarding selection of patients with access to a smartphone. Thus, we believe that the findings of this trial can be generalized to patients with bipolar disorder in general.

Implications

The present trial showed that both the intervention and control groups had a decrease in the severity of depressive and manic symptoms during the 6-month trial period. This unspecific effect of time was similar to findings from an observational study on electronic self-monitoring in which patients with bipolar disorder used weekly text messages on a regular cell phone as a monitoring instrument and improved over time (Bopp *et al.* 2010). The latter observational study did not include a control group, thereby making it impossible to identify the specific effect of using electronic self-monitoring.

Considering the potential harmful effects on depressive symptoms, findings from our trial highlight that electronic self-monitoring should not be uncritically used or implemented and that important aspects need further clarification. It is therefore recommended that future use of daily electronic self-monitoring needs to be further elucidated specifically focusing on how to use the electronic device to clinically detect and treat early warning signs of depression and mania and be subsequently investigated in larger RCTs, perhaps including patients with higher levels of depressive and manic symptoms at baseline and with a longer follow-up period.

Conclusion

Our hypothesis that daily electronic self-monitoring using smartphones including a two-level feedback loop reduces depressive and manic symptoms in patients with bipolar disorder could not be confirmed. Analyses showed a tendency towards more depressive symptoms in the intervention group using a smartphone for electronic self-monitoring including a feedback loop compared to the control group using a smartphone for normal communicative purposes. Subgroup analyses showed that in patients with no mixed symptoms, the patients in the intervention group had significantly more depressive symptoms than patients in the control group. Furthermore, sub-analysis also showed that, in patients with manic symptoms, the patients in the intervention group had significantly less manic symptoms than patients in the control group. These results highlight important questions in relation to electronic self-monitoring that need further clarification. Future clinical use of electronic self-monitoring should thus be monitored carefully in relevant research settings before it is implemented as a standard clinical tool.

Acknowledgements

The EU 7th Frame Programme funded the MONARCA I trial together with the Mental Health Services, Copenhagen, Denmark, Trygfonden, the Gert Einar Joergensens Foundation, and the A.P. Moeller og Hustru Chastine McKinney Moellers Foundation for general purposes. The funders had no role in the trial design, data collection, analyses and preparation of the manuscript. The authors thank the patients for participating in this trial and research nurse Hanne Steenberg Nikolajsen for running the feedback loop of the MONARCA system.

Declaration of Interest

M.F.J. has been a consultant for Eli Lilly. M.V. has been a consultant for Eli Lilly, Lundbeck, AstraZeneca and Servier. E.M.C. has been a consultant for Eli Lilly, AstraZeneca, Servier, Bristol-Myers-Squibb, Lundbeck and Medilink. R.L.M. has been a consultant for Lundbeck, AstraZeneca, Bristol-Myers-Squibb, Otsuka and Schering-Plough. U.K. has been a consultant for Servier and AstraZeneca. L.V.K. has within the past 3 years been a consultant for Lundbeck and AstraZeneca. M.F., C.R., A.S.J., and J.B. have no conflicts of interests.

References

Altman EG, Hedeker D, Peterson JL, Davis JM (1997). The Altman self-rating mania scale. *Biological Psychiatry* **42**, 948–955.

Baldessarini RJ, Salvatore P, Kaur Khalsa H-M, Gebre-Medhin P, Imaz H, González-Pinto A, Perez J, Cruz N, Maggini C, Tohen M (2010). Morbidity in 303 first-episode bipolar I disorder patients. *Bipolar Disorders* **12**, 264–270.

Bardram J, Frost M, Szanto K, Margu G (2012). The MONARCA self-assessment system: a persuasive personal monitoring system for bipolar patients. In *Proceeding IHI '12. Proceedings of the 2nd ACM SIGHIT*, ACM, New York, NY, USA, 2012, pp. 21–30.

Bardram JE, Frost M, Szánto K, Faurholt-Jepsen M, Vinberg M, Kessing LV (2013). Designing mobile health technology for bipolar disorder: a field trial of the Monarca system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2013, pp. 2627–2636.

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. *Archives of General Psychiatry* **48**, 807–812.

Bech P, Olsen LR (2001). Discovering depression. *Ugeskrift for laeger* **163**, 1980–1982.

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. *Journal of Affective Disorders* **66**, 159–164.

Bonsall MB, Wallace-Hadrill SMA, Geddes JR, Goodwin GM, Holmes EA (2012). Nonlinear time-series approaches in characterizing mood stability and mood instability in bipolar disorder. *Proceedings of the Royal Society of London, Series B: Biological Sciences* **279**, 916–924.

Bopp JM, Miklowitz DJ, Goodwin GM, Stevens W, Rendell JM, Geddes JR (2010). The longitudinal course of bipolar disorder as revealed through weekly text messaging: a feasibility study. *Bipolar Disorders* **12**, 327–334.

Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P (2008). Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Annals of Internal Medicine* **148**, 295–309.

Cohen S, Kamarck T, Mermelstein R (1983). A global measure of perceived stress. *Journal of Health and Social Behavior* **24**, 385–396.

Colom F, Vieta E, Martinez-Aran A, Reinares M, Goikolea JM, Benabarre A, Torrent C *et al.* (2003). A randomized trial on the efficacy of group psychoeducation in the prophylaxis of recurrences in bipolar patients whose disease is in remission. *Archives of General Psychiatry* **60**, 402–407.

De Assis da Silva R, Mograbi DC, Silva Silveira LA, Santos Nunes AL, Novis FD, Landeira-Fernandez J, Cheniaux E (2015). Insight across the different mood states of bipolar disorder. *Psychiatric Quarterly*. Published online: 18 January 2015. doi:10.1007/s11126-015-9340-z.

Depp CA, Kim DH, de Dios LV, 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 CA, Mausbach B, Granholm E, Cardenas V, Ben-Zeev D, Patterson TL, Lebowitz BD, Jeste DV (2010). Mobile interventions for severe mental illness: design and preliminary data from three approaches. *Journal of Nervous and Mental Disease* **198**, 715–721.

Endler NS, Parker JD (1998). *Coping Inventory for Stressful Situations (CISS)*, 2nd edn. Multi-health Systems. Manual.

Faurholt-Jepsen M, Frost M, Vinberg M, Christensen EM, Bardram JE, Kessing LV (2014). Smartphone data as objective measures of bipolar disorder symptoms. *Psychiatry Research* 217, 124-127.

Faurholt-Jepsen M, Vinberg M, Christensen EM, Frost M, Bardram J, Kessing LV (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. *British Medical Journal Open* **3**: e003353.

Fava M, Iosifescu DV, Pedrelli P, Baer L (2009). Reliability and validity of the Massachusetts general hospital cognitive and physical functioning questionnaire. *Psychotherapy and Psychosomatics* **78**, 91–97.

Frost M, Faurholt-Jepsen M, Doryab A, Kessing LV, Bardram JE (in press). Increasing awareness, insight and adherence in treatment of bipolar disorder through personal health technology: pilot study. *Journal of Medical Internet Research*.

Ghaemi SN, Rosenquist KJ (2004) Is insight in mania state-dependent?: a meta-analysis. *Journal of Nervous and Mental Disease* **192**, 771–775.

Guilera G, Pino O, Gómez-Benito J, Rojo JE, Vieta E, Tabarés-Seisdedos R, Segarra N *et al.* (2009). Clinical usefulness of the Screen for Cognitive Impairment in Psychiatry (SCIP-S) Scale in patients with type I bipolar disorder. *Health and Quality of Life Outcomes* **7**, 28.

Hamilton M (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology* **6**, 278–296.

Henry C, Van den Bulke D, Bellivier F, Roy I, Swendsen J, M'Baïlara K, Siever LJ, Leboyer M (2008). Affective lability and affect intensity as core dimensions of bipolar disorders during euthymic period. *Psychiatry Research* **159**, 1–6.

Husky MM, Gindre C, Mazure CM, 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.

Kessing LV, Hansen HV, Hvenegaard A, Christensen EM, 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.

Kessing LV, Hansen MG, Andersen PK (2004). Course of illness in depressive and bipolar disorders. naturalistic study, 1994–1999. *British Journal of Psychiatry* **185**, 372–377.

Kessing LV, Søndergård L, Kvist K, Andersen PK (2007). Adherence to lithium in naturalistic settings: results from a nationwide pharmacoepidemiological study. *Bipolar Disorders* **9**, 730–736.

Kreindler D, Levitt A, Woolridge N, Lumsden CJ (2003). Portable mood mapping: the validity and reliability of analog scale displays for mood assessment via hand-held computer. *Psychiatry Research* **120**, 165–177.

Látalová K (2012). Insight in bipolar disorder. *Psychiatric Quarterly* **83**, 293–310.

Laursen RP, Dalskov S-M, Damsgaard CT, Ritz C (2014). Back-transformation of treatment differences-an approximate method. *European Journal of Clinical Nutrition* **68**, 277–280.

Lecrubier Y, Bech P (2007). The Ham D(6) is more homogenous and as sensitive as the Ham D(17). *European Psychiatry* **22**, 252–255.

Leverich GS, Nolen WA, Rush AJ, McElroy SL, Keck PE, Denicoff KD, Suppes T, Althuler LL, Kupka R, Kramlinger KG, Post RM (2001). The Stanley Foundation bipolar treatment outcome network. I. Longitudinal methodology. *Journal of Affective Disorders* **67**, 33–44.

Lieberman DZ, Kelly TF, Douglas L, Goodwin FK (2010). A randomized comparison of online and paper mood charts for people with bipolar disorder. *Journal of Affective Disorders* **124**, 85–89.

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 Disorders* **5**, 349–355.

Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG (2010). CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *British Medical Journal (Clinical Research Edition)* **340**, c869.

Morriss RK, Faizal MA, Jones AP, Williamson PR, Bolton C, McCarthy JP (2007). Interventions for helping people recognise early signs of recurrence in bipolar disorder. *Cochrane Database of Systematic Reviews*. Art no. CD004854. doi:10.1002/14651858.CD004854.pub2.

National Institute for Clinical Excellence (2004). *Depression: Management of Depression in Primary and Secondary Care*. National Institute for Clinical Excellence: London, England.

Olsen LR, Jensen DV, Noerholm V, Martiny K, Bech P (2003). The internal and external validity of the major depression inventory in measuring severity of depressive states. *Psychological Medicine* **33**, 351–356.

Paykel ES, Abbott R, Morriss R, Hayhurst H, Scott J (2006). Sub-syndromal and syndromal symptoms in the longitudinal course of bipolar disorder. *British Journal of Psychiatry* **189**, 118–123.

Perry A, Tarrier N, Morriss R, McCarthy E, Limb K (1999). Randomised controlled trial of efficacy of teaching patients with bipolar disorder to identify early symptoms of relapse and obtain treatment. *British Medical Journal (Clinical Research Edition)* **318**, 149–153.

Rojo E, Pino O, Guilera G, Gómez-Benito J, Purdon SE, Crespo-Facorro B, Cuesta MJ *et al*. (2010). Neurocognitive diagnosis and cut-off scores of the Screen for Cognitive Impairment in Psychiatry (SCIP-S). *Schizophrenia Research* **116**, 243–251.

Rosa AR, Sánchez-Moreno J, Martínez-Aran A, Salamero M, Torrent C, Reinares M, Comes M *et al*. (2007). Validity and reliability of the Functioning Assessment Short Test (FAST) in bipolar disorder. *Clinical Practice and Epidemiology in Mental Health: CP & EMH* **3**, 5.

Sachs G (2014). Step-BD blank mood chart (http://www.manicdepressive.org/images/samplechart.gif).

Schärer LO, Hartweg V, Hoern M, Graesslin Y, Strobl N, Frey S, Biedermann C, Walser S, Walden J (2002). Electronic diary for bipolar patients. *Neuropsychobiology* **46** (Suppl. 1), 10–12.

Scott J, Colom F (2008). Gaps and limitations of psychological interventions for bipolar disorders. *Psychotherapy and Psychosomatics* **77**, 4–11.

Simon GE, Ludman EJ, Bauer MS, Unützer J, Operskalski B (2006). Long-term effectiveness and cost of a systematic care program for bipolar disorder. *Archives of General Psychiatry* **63**, 500–508.

Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR (2003). Patient compliance with paper and electronic diaries. *Controlled Clinical Trials* **24**, 182–199.

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 Psychiatrica Scandinavica* **128**, 194–202.

Totterdell P, Briner RB, Parkinson B, Reynolds S (1996). Fingerprinting time series: dynamic patterns in self-report and performance measures uncovered by a graphical non-linear method. *British Journal of Psychology* **87**, 43–60.

WHO (1998). Development of the World Health Organization WHOQOL-BREF Quality of Life Assessment. The WHOQOL Group. *Psychological Medicine* **28**, 551–558.

Whybrow PC, Grof P, Gyulai L, Rasgon N, Glenn T, Bauer M (2003). The electronic assessment of the longitudinal course of bipolar disorder: The ChronoRecord Software. *Pharmacopsychiatry* **36** (Suppl. 3), S244–S249.

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. *Archives of General Psychiatry* **47**, 589–593.

Yen C-F, Chen C-S, Ko C-H, Yen J-Y, Huang C-F (2007). Changes in insight among patients with bipolar I disorder: a 2-year prospective study. *Bipolar Disorders* **9**, 238–242.

Young RC, Biggs JT, Ziegler VE, Meyer DA (1978). A rating scale for mania: reliability, validity and sensitivity. *British Journal of Psychiatry* **133**, 429–435.