Psychological Medicine

cambridge.org/psm

The effect of smartphone-based monitoring on illness activity in bipolar disorder: the MONARCA II randomized controlled single-blinded trial

Maria Faurholt-Jepsen[1], Mads Frost[2], Ellen Margrethe Christensen[1], Jakob E. Bardram[3], Maj Vinberg[1] and Lars Vedel Kessing[1]

[1]Copenhagen Affective Disorder Research Center (CADIC), Psychiatric Center Copenhagen, Rigshospitalet, Copenhagen, Denmark; [2]IT University of Copenhagen, Rued Langgaards Vej 7, 2300 Copenhagen, Denmark and [3]Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark

Original Article

Cite this article: Faurholt-Jepsen M, Frost M, Christensen EM, Bardram JE, Vinberg M, Kessing LV (2020). The effect of smartphone-based monitoring on illness activity in bipolar disorder: the MONARCA II randomized controlled single-blinded trial. *Psychological Medicine* **50**, 838–848. https://doi.org/10.1017/S0033291719000710

Received: 28 June 2018
Revised: 14 March 2019
Accepted: 14 March 2019
First published online: 4 April 2019

Key words:
Bipolar disorder; depressive and manic symptoms; illness activity; MONARCA II; randomized controlled trial; smartphone; The Monsenso system

Author for correspondence:
Maria Faurholt-Jepsen,
E-mail: maria@faurholt-jepsen.dk

Abstract

Background. Recently, the MONARCA I randomized controlled trial (RCT) was the first to investigate the effect of smartphone-based monitoring in bipolar disorder (BD). Findings suggested that smartphone-based monitoring sustained depressive but reduced manic symptoms. The present RCT investigated the effect of a new smartphone-based system on the severity of depressive and manic symptoms in BD.

Methods. Randomized controlled single-blind parallel-group trial. Patients with BD, previously treated at The Copenhagen Clinic for Affective Disorder, Denmark and currently treated at community psychiatric centres, private psychiatrists or GPs were randomized to the use of a smartphone-based system or to standard treatment for 9 months. Primary outcomes: differences in depressive and manic symptoms between the groups.

Results. A total of 129 patients with BD (ICD-10) were included. Intention-to-treat analyses showed no statistically significant effect of smartphone-based monitoring on depressive ($B = 0.61$, 95% CI −0.77 to 2.00, $p = 0.38$) and manic ($B = -0.25$, 95% CI −1.1 to 0.59, $p = 0.56$) symptoms. The intervention group reported higher quality of life and lower perceived stress compared with the control group. In sub-analyses, the intervention group had higher risk of depressive episodes, but lower risk of manic episodes compared with the control group.

Conclusions. There was no effect of smartphone-based monitoring. In patient-reported outcomes, patients in the intervention group reported improved quality of life and reduced perceived stress. Patients in the intervention group had higher risk of depressive episodes and reduced risk of manic episodes. Despite the widespread use and excitement of electronic monitoring, few studies have investigated possible effects. Further studies are needed.

Introduction

Bipolar disorder (BD) is characterized by recurrent episodes of depression and (hypo)mania with intervening periods of euthymia, but mood deviances between episodes can also present in the form of subsyndromal mood changes. BD is one of the most important causes of disability worldwide (Pini *et al.*, 2005; Harrison *et al.*, 2016). Naturalistic follow-up studies suggest that the progressive development of BD is not prevented with the present treatment options (Angst and Sellaro, 2000; Kessing *et al.*, 2004). Given the limited resources available, a recent report by the World Health Organization stated that 'the use of mobile and wireless technologies (mhealth) to support the achievement of health objectives has the potential to transform the face of health service delivery across the globe' (WHO, 2011). During recent years, there has been a rapid increase in mhealth technologies for self-monitoring within mental health (Smith, 2012; Donker *et al.*, 2013; Faurholt-Jepsen *et al.*, 2016*b*; Wang *et al.*, 2018). Within BD research, a number of studies have reported on electronic self-monitoring of depressive and manic symptoms using regular cell phones (Bopp *et al.*, 2010), personal digital assistants (Depp *et al.*, 2010; Mutschler *et al.*, 2012), computers (Whybrow *et al.*, 2003; Bauer *et al.*, 2004, 2005, 2006, 2007, 2008; Chinman *et al.*, 2004; Adli *et al.*, 2005; Lieberman *et al.*, 2011) and smartphones (Depp *et al.*, 2012; Hidalgo-Mazzei *et al.*, 2015; Bilderbeck *et al.*, 2016) as the electronic self-monitoring tools. However, none of the studies has included data on objective measures of illness activity, and the effect of smartphone-based monitoring has only been investigated in a single randomized controlled trial (RCT), the MONARCA I trial, conducted by the authors (Faurholt-Jepsen *et al.*, 2015*a*). The MONARCA I trial investigated the effect of smartphone-based monitoring including a clinical feedback loop during 6 months compared with using a placebo smartphone in patients with BD. Overall, no

CAMBRIDGE UNIVERSITY PRESS

differences between the intervention group and the control group were found, but importantly sub-analyses suggested that smartphone-based monitoring may sustain depressive symptoms and reduce manic symptoms (Faurholt-Jepsen *et al.*, 2015*a*).

Social activity and communication (Weinstock and Miller, 2008), as well as mobility (Kupfer *et al.*, 1974; Kuhs and Reschke, 1992; Faurholt-Jepsen *et al.*, 2012) represent central aspects reflecting illness activity in BD. It is likely that the ability of smartphone-based self-monitored measures may not be sufficient to detect prodromal depressive and manic symptoms compared with automatically generated objective smartphone data (referred to as objective smartphone data) on measures of illness activity such as phone usage, social activity and mobility which we have shown in previous studies correlate with the level of depressive and manic symptoms, discriminates between states and may represent a potential diagnostic marker (Faurholt-Jepsen *et al.*, 2014*a*, 2015*b*, 2016*c*, 2018). It has never been tested in an RCT whether smartphone-based monitoring including a clinical feedback loop integrating subjective as well as objective smartphone data in patients with BD improves illness outcome.

Following the MONARCA I study, adjustments to the smartphone-based monitoring system were made and a new integrated clinical feedback loop, based on real-time prediction models including both subjective measures and objective smartphone data, was established and implemented in the Monsenso system.

We therefore conducted the MONARCA II RCT and hypothesized that smartphone-based monitoring and mood prediction including a clinical feedback loop in patients with BD would reduce the level of depressive and manic symptoms more than standard treatment.

Methods

The trial is reported according to the CONsolidated Standards Of Reporting Trials (CONSORT) statement and Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) (Boutron *et al.*, 2008; Moher *et al.*, 2010; Chan *et al.*, 2013).

Details concerning the design and methods of the trial have previously been published (Faurholt-Jepsen *et al.*, 2014*b*).

Design, settings and participants

The trial was a randomized controlled single-blind parallel-group design with an unbalanced allocation ratio (2:1) of adult patients with BD with a 9-month follow-up period. All patients with a diagnosis of BD who previously (but not currently) had been treated at The Copenhagen Clinic for Affective Disorder Copenhagen, Denmark in the period 2004 to January 2016 but currently treated elsewhere (e.g. at community psychiatric centres, private psychiatrists and general practitioners) were invited to participate in the trial, corresponding to approximately 735 potential participants (Fig. 1). The Copenhagen Clinic for Affective Disorder is a specialized outpatient clinic covering a recruitment area of the Capital Region, Denmark, corresponding to 1.4 million people. Patients with BD were initially referred to the Copenhagen Clinic for Affective Disorder from secondary healthcare when a diagnosis of a single mania or BD was made for the first time or in the case of occurrence of treatment resistance (persistent depressive or manic symptoms or recurrence despite treatment with standard care). Treatment at the Copenhagen Clinic for Affective Disorder comprised combined psychopharmacological treatment as according to international guidelines and supporting therapy for a 2-year period. All the patients included in the present trial had finalized their course of treatment at the Copenhagen Clinic for Affective Disorder before they were invited to participate in the present trial.

Inclusion criteria: BD diagnosis according to the International Classification of Diseases (ICD-10) (World Health Organisation, 2011) using Schedules for Clinical Assessments in Neuropsychiatry (SCAN) (Wing *et al.*, 1990) and previous treatment at the Copenhagen Clinic for Affective Disorder, Denmark.

Exclusion criteria: Schizophrenia, schizotypal or delusional disorders; previous use of the MONARCA system; pregnancy and a lack of Danish language skills.

Randomization, allocation and implementation

In addition to standard treatment, patients were randomized with an unbalanced allocation of 2:1 to either (1) the intervention group: active use of the smartphone-based Monsenso system on either (i) a smartphone capable of collecting objective smartphone data on measures of illness activity (Android smartphones) or (ii) a smartphone not capable of collecting objective smartphone data (iPhone) or to (2) the control group: use of a smartphone for usual communication. A computer-generated list of randomization allocation numbers was obtained by an independent researcher (KM) using http://randomization.com. The unbalanced allocation ratio was chosen due to differences in the level of available data collected for the clinical feedback loop in the intervention group. The type of smartphone used during the trial (Android or iPhone) was chosen by the patients themselves before randomization and was estimated to be approximately 50% on each of these types of smartphones – Android (capable of collecting objective smartphone data) *v.* iPhone (not capable of collecting objective smartphone data). A non-stratified randomization with random block sizes was used. The allocation sequence was concealed in numbered opaque envelopes in a locked cabinet of unknown location to the blinded researchers.

Blinding

The MONARCA II trial was a single-blinded trial. Owing to the type of intervention, the patients, the patients' health care provider and the study nurse were aware of the allocated randomization group. The researchers conducting outcome assessments, data entry, data analysis, interpretation and writing of the present paper were kept blinded to allocation.

Intervention

The intervention group

The Monsenso system was available for smartphones capable of collecting objective smartphone data (Android) or smartphones not capable of collecting objective smartphone data (iPhone). Regardless of the type of smartphone used, the patients randomized to the intervention group had the Monsenso system installed on their own or a borrowed smartphone and instructed to use the system for daily self-monitoring (Fig. 2). Self-monitored measures included mood, sleep duration, activity level, etc. [details provided in the online Supplementary data e1 and in Faurholt-Jepsen *et al.* (2014b)].

Objective smartphone data including data on phone usage, social activity and mobility were included in the clinical feedback loop (details provided in the online Supplementary data). All data

https://doi.org/10.1017/S0033291719000710 Published online by Cambridge University Press

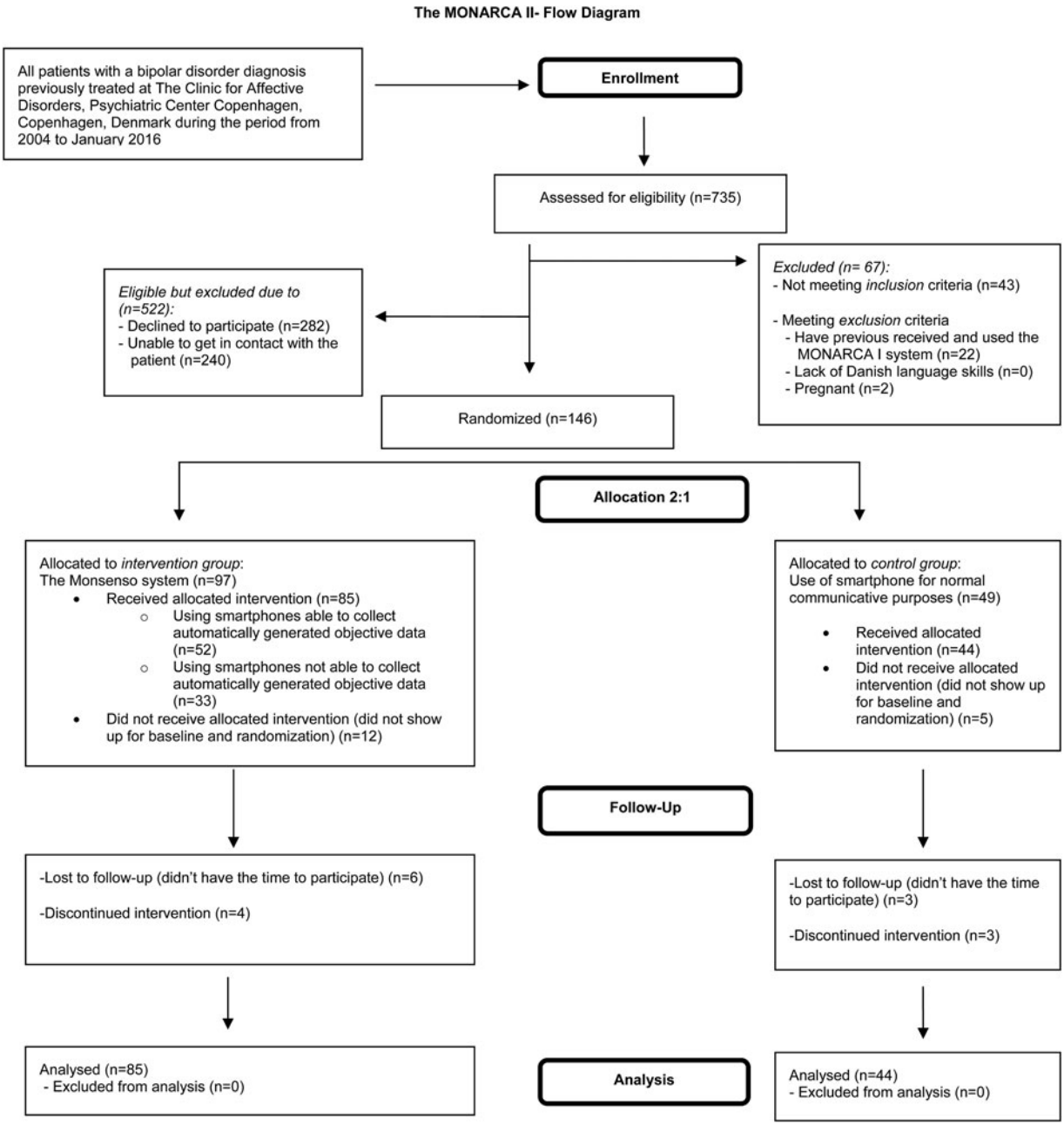

Fig. 1. The MONARCA II – flow diagram.

were used in a real-time forecasting algorithm to predict future mood. The clinical feedback loop between patients and clinicians comprised a study nurse who examined the collected data on a web-page, including the mood prediction (details provided in the online Supplementary data). If mood prediction was exceeding a predefined threshold, a message would be sent both to the patients and the study nurse (details provided in the online Supplementary data).

The control group

The patients who were randomized to the control group were offered to borrow a smartphone or used their smartphones for usual communication.

All patients continued with standard treatment.

Outcomes and assessments

Outcome measures were defined *a priori* (Faurholt-Jepsen *et al.*, 2014*b*). Primary outcome measures were differences in clinically rated depressive and manic symptoms measured using the Hamilton Depression Rating Scale 17-items (HDRS-17) and the Young Mania Rating Scale (YMRS), respectively (Hamilton, 1967; Young *et al.*, 1978) between the intervention group and the control group.

Sub-analyses in relation to the primary outcomes: Since the effect of smartphone-based monitoring in BD has been sparingly investigated and based on interesting findings from the MONARCA I trial, sub-analyses in relation to the primary outcome were defined *a priori* (Faurholt-Jepsen *et al.*, 2014*b*). Differences in depressive and manic symptoms measured using

https://doi.org/10.1017/S0033291719000710 Published online by Cambridge University Press

Fig. 2. The Monsenso system for smartphones.

HDRS-17 item and YMRS, respectively, between the intervention group and the control group in patients with (a) no mixed symptoms (defined as YMRS⩾7 or HDRS⩾7), (b) presence of depressive and manic symptoms (defined as HDRS-17>0 or YMRS>0) *at any given visit* and (c) with depressive or manic symptoms at baseline (defined as HDRS⩾7 or YMRS⩾7) were investigated. Also, differences in rates of relapse of depressive and manic episodes during the trial period between the intervention group and the control group were investigated.

Secondary outcomes were defined as differences in functioning according to the Functional Assessment Short Test (FAST) (Rosa *et al.*, 2007) between the intervention group and the control group.

Tertiary outcomes were defined as differences between the intervention group and the control group in perceived stress according to Cohen's Perceived Stress Scale (PSS) (Cohen *et al.*, 1983), quality of life according to the WHO Quality of Life-BREF (WHOQoL-BREF) (WHO, 1998), self-rated depressive symptoms according to Beck's Depressive Inventory (BDI) (Bech *et al.*, 2001), self-rated manic symptoms according to the Altman Self Rating scale for Mania (ASRM) (Altman *et al.*, 1997), self-rated recovery according to the Recovery Assessment Scale (Corrigan *et al.*, 2004), self-rated medicine adherence according to the Medicine Adherence Rating Scale (MARS) (Thompson *et al.*, 2000) and empowerment according to Roger's Empowerment Scale (Rogers *et al.*, 1997).

All assessments were carried out rater-blinded by one researcher (MFJ) at baseline, 4 weeks, 3, 6 and 9 months.

Statistical methods

Statistical power

A priori, we defined that the minimal clinically relevant difference in the level of depressive and manic symptoms between the intervention group and the control group was three points on HDRS-17 item and YMRS, respectively. Further, the standard deviation (S.D.) on the HDRS-17 item and the YMRS in the overall population included was 4.7 points. The statistical power was 80% with $\alpha = 0.05$ for a two-sample comparison of means when including a minimum of 39 patients in each of the intervention groups (39×2: $n = 78$) and 39 patients in the control group ($N = 117$).

Statistical analyses

The statistical analyses were conducted using an intention-to-treat (ITT) analysis. Two-sample *t* tests and χ^2 tests were used to assess the differences in means and proportions of background characteristics (Table 1). For each outcome, a two-level linear mixed-effect model, which accommodates both intra-individual and inter-individual variation of the specified outcome variable, including a fixed effect of visit number and a patients-specific random effect allowing for individual intercept and slope for each participant was defined (Table 2 and 3). Level 1 represented repeated measures of symptoms (e.g. HDRS, YMRS) and level 2 represented inter-individual variation. Differences between the intervention group and the control were investigated first in unadjusted models (except for differences in baseline values of the outcome variable) and second in models adjusted for age and sex. Model assumptions were checked. Potential interactions between randomization group and visit number on the outcome of interest were investigated. Differences in rates of relapse of episodes during the study were investigated using survival analysis [hazard rate (HR)] with reasons for censoring being death or date of drop out (Table 2). Analyses on probability of individual psychopharmacological prescriptions were investigated using logistic regression analyses [odds ratio (OR)] (Table 3). The threshold for statistical significance was $p \leqslant 0.05$ (two-tailed). Analyses were done using STATA 14 (StataCorp LP, College Station, TX, USA).

Ethical considerations

Ethical permission was obtained from the Regional Ethics Committee in The Capital Region of Denmark and The Danish Data Protection Agency (H-2-2014-059). The law on handling of personal data was respected. The trial was registered at ClinicalTrials.gov as NCT02221336. Electronic data collected from smartphones were stored at a secure server at CIMT, Capital Region, Denmark (I-suite number RHP-2011-03). All patients were offered to borrow a smartphone free of charge during the trial period, and economic costs due to data traffic from the Monsenso system were refunded. No economic compensation for participating was provided. The trial complied with the Declaration of Helsinki of 1975, as revised in 2008.

https://doi.org/10.1017/S0033291719000710 Published online by Cambridge University Press

Table 1. Clinical and socio-demographic characteristics of study participants at baseline, $N = 129$

Randomization group	Intervention group	Control group
n	85	44
Socio-demographic data		
Age, years	43.0 (12.4)	43.2 (12.4)
Female sex, % (*n*)	61.2 (52)	54.5 (24)
In relationship, % (*n*)	55.3 (47)	58.1 (25)
Full-time employed, % (*n*)	17.7 (15)	16.7 (7)
Part-time employed, % (*n*)	14.1 (12)	26.2 (12)
Student, % (*n*)	15.3 (13)	11.9 (5)
Educational level		
Primary school or lower, % (*n*)	4.9 (4)	0
High school, % (*n*)	19.5 (17)	28.6 (12)
University undergraduate or more, % (*n*)	64.7 (55)	54.7 (23)
Years of education after primary school	5.8 (2.7)	5.8 (2.8)
Clinical history		
Bipolar disorder type I diagnosis, % (*n*)	63.5 (54)	48.8 (21)
Admissions, number	2 [1–3]	2 [0–3]
Depressive episodes, number	4 [2–10]	5 [3–10]
Manic episodes, number	3 [2–7]	5 [2–10]
Psychopharmacological treatment, baseline, % (*n*)		
Antipsychotics	49.4 (42)	42.9 (36)
Anticonvulsants	56.5 (48)	50 (22)
Antidepressants	22.6 (19)	15.9 (7)
Lithium	63.5 (54)	31.8 (14)
Hamilton Depression Rating Scale 17 items score, baseline	6 [4–11]	7 [1–14]
Hamilton Depression Rating Scale 17 items score, 9 months	6 [2–10]	5 [3–8]
Young Mania Rating Scale score, baseline	2 [0–5]	1[0–4]
Young Mania Rating Scale score, 9 months	1 [0–3]	1 [0–3]
First degree relative with affective disorder, % (*n*)	53.0 (45)	47.6 (20)

Data are mean (s.d.), median [IQR] or % (*n*) unless otherwise stated.

Results

Background characteristics

Figure 1 presents the flow of patients in the MONARCA II trial. A total of 735 patients previously treated at the Copenhagen Clinic for Affective Disorder, Denmark from 2004 to January 2016, but currently treated at community psychiatric centres, private psychiatrists and general practitioners, were assessed for eligibility. Of these 544 patients were not included due to; unable to get in contact with the patient ($n = 240$), declined to participate (main reasons: did not have the time, did not want to participate in a research study or had moved too far away making transportation a problem) ($n = 282$ patients), or were excluded due to previous use of the MONARCA system ($n = 22$). The last patient visit was in January 2018. As 17 patients did not show up for baseline assessments and randomization, a total of 129 patients were randomized 2:1 to either the intervention group [active use of the smartphone-based Monsenso system on either (i) a smartphone collecting objective smartphone data or (ii) a smartphone not collecting objective smartphone data] or to the control group (use of a smartphone for usual communication) and left for ITT analyses. Of these, nine patients expressed that they wished to drop out during the trial (intervention group: six; control group: three) due to not having the time to participate (one patient wished to drop out due to self-monitoring). A total of seven patients (5.4%) had a discontinued intervention (intervention group: four; control group: three). Thus, there was only missing outcome data to a small extent and therefore not accounted for in the statistical analyses. Clinical and socio-demographic characteristics are presented in Table 1. Details regarding results on the adherence to the self-monitoring in the intervention group, technical aspects of the system and the clinical feedback loop are provided in the online Supplementary data (e2).

Primary outcomes

Differences between the intervention group and the control in depressive and manic symptoms according to HDRS and YMRS, respectively, are presented in Table 2. There were no differences between the intervention group and the control group in level of depressive [*adjusted model: coefficient B* (B) = 0.61, 95% CI −0.77 to 2.00, $p = 0.38$] and manic symptoms (*adjusted model:* $B = -0.25$, −1.1 to 0.59, $p = 0.56$). Analyses on interactions between randomization group and time (visit number) were statistically non-significant, and therefore not reported.

Exploratory analyses in relation to primary outcomes

Sub-analyses are presented in Table 2. Excluding mixed symptoms, patients in the intervention group experienced statistically significantly lower levels of manic symptoms compared with the control group (*adjusted model:* $B = -1.11$, 95% CI −2.22 to −0.01, $p = 0.050$) ($n = 88$). When including patients with a HDRS score ⩾7 at baseline, there was no difference between the intervention group and the control group (*adjusted model:* $B =$ 0.74, 95% CI −1.30 to 2.79, $p = 0.48$) ($n = 64$). However, patients in the intervention group presenting with an YMRS score ⩾7 at baseline experienced higher levels of manic symptoms compared with the control group (*adjusted model:* $B = 4.21$, 95% CI 0.45–7.97, $p = 0.028$) ($n = 19$).

In separate independent exploratory stratified analyses on (a) ⩾3 or <3 previous affective episodes, (b) BD type I or type II, (c) ⩾3 or <3 previous psychiatric hospitalization, (d) index affective episode depression or mania, there were no differences in the level of depressive and manic symptoms between the intervention group and the control group (all $p > 0.05$).

In survival analyses on differences in the rates of relapse in depressive and manic episodes based on patients without being in a current affective episode at baseline, there were no differences in the number of episodes, when not considering the polarity of episodes (depression or mania), between the intervention group and the control group ($N = 120$, *HR* 0.76, 95% CI 0.41–1.42, $p = 0.39$). However, there was a statistically significantly higher

https://doi.org/10.1017/S0033291719000710 Published online by Cambridge University Press

Table 2. Estimated differences in primary outcomes between the intervention group and the control group (control group serve as reference), $N = 129$

	Unadjusted[a]			Adjusted[b]		
	Difference between groups	95% CI	*p*	Difference between groups	95% CI	*p*
HDRS-17[c]	**0.67**	**−0.70 to 2.1**	**0.33**	**0.61**	**−0.77 to 2.00**	**0.38**
HDRS-6[d]	0.13	−0.50 to 0.77	0.68	0.083	−0.55 to 0.77	0.80
Sub-analyses on HDRS 17-items						
1A. HDRS-17 items, no mixed symptoms ($n = 114$)[e]	0.73	−0.77 to 2.24	0.34	0.62	−0.88 to 2.11	0.42
1B. HDRS-17 items>0 at any time point ($n = 121$)	0.84	−0.54 to 2.21	0.23	0.78	−0.59 to 2.14	0.27
1C. HDRS-17 items baseline ⩾7 ($n = 64$)	0.93	−1.21 to 3.07	0.39	0.74	−1.30 to 2.79	0.48
1D. HDRS-17 items baseline <7 ($n = 57$)	0.35	−1.55 to 2.24	0.72	0.37	−1.52 to 2.27	0.70
YMRS[f]	**−0.24**	**−1.1 to 0.60**	**0.57**	**−0.25**	**−1.1 to 0.59**	**0.56**
Sub-analyses on YMRS						
2A. YMRS, no mixed symptoms ($n = 88$)[g]	−1.10	−2.22 to −0.02	0.050	−1.11	−2.22 to −0.01	0.050
2B. YMRS>0 at any time point ($n = 105$)	0.01	−1.15 to 1.16	0.99	0.00	−1.16 to 1.15	0.39
2C. YMRS baseline ⩾7 ($n = 19$)	4.04	0.33–7.75	0.033	4.21	0.45–7.97	0.028
2D. YMRS baseline <7 ($n = 102$)	−0.38	−1.24 to 0.47	0.38	−0.38	−1.24 to 0.47	0.38
	Hazard rate	95% CI	*p*	Hazard rate	95% CI	*p*
Depressive episodes[h] ($n = 97$)	1.65	0.65–4.19	0.29	2.89	1.02–8.23	0.047
Manic episodes[i] ($n = 110$)	0.40	0.12–1.32	0.13	0.17	0.037–0.78	0.022
Affective episodes regardless polarity ($n = 120$)	0.95	0.52–1.73	0.86	0.76	0.41–1.42	0.39

The use of bold indicate the primary main analyses.
[a]Adjusted for outcome variable at baseline.
[b]Adjusted for outcome variable at baseline, age and gender unless otherwise specified.
[c]The Hamilton Depression Rating Scale 17-items.
[d]The Hamilton Depression Rating Scale six-items.
[e]Defined as YMRS ⩾7.
[f]The Young Mania Rating Scale.
[g]Defined as HDRS 17-items ⩾7.
[h]Defined as HDRS-17 ⩾13.
[i]Defined as YMRS ⩾13.

rate of depressive episodes between the intervention group and the control group ($N = 97$, *HR* 2.89, 95% CI 1.02–8.23, $p = 0.047$), but a lower rate of manic episodes in the intervention group ($N = 116$, *HR* 0.17, 95% CI 0.037–0.78, $p = 0.022$).

Secondary and tertiary outcomes

Analyses on secondary and tertiary outcomes are presented in Table 3. There were significantly lower perceived stress and higher quality of life in the intervention group compared with the control group (perceived stress, adjusted model: $B = -2.08$, 95% CI −4.15 to −0.01, $p = 0.049$; quality of life, adjusted model: $B = 4.00$, 95% CI 0.04–7.97, $p = 0.048$). There were no differences between the intervention group and the control group on other secondary and tertiary outcome measures.

In logistic regression models on probability of prescription of individual groups of psychopharmacological treatments, the intervention group had higher probability of being prescribed lithium than the control group (OR 2.78, 95% CI 1.33–6.20, $p = 0.007$). In survival analyses on differences in the rate of prescription of lithium among patients not in lithium treatment at baseline, there was no difference between the intervention group and the control group ($N = 57$, *HR* 0.30, 95% CI 0.032–2.93, $p = 0.30$). There was no difference in the rate of discontinuation of lithium among patients in lithium treatment at baseline between the intervention group and the control group ($N = 66$, *HR* 0.36, 95% CI 0.08–1.61, $p = 0.18$).

Notably, further adjustment of the analyses on the primary outcomes (HDRS and YMRS) for the use of lithium did not alter the estimates.

There were no differences in the probability of prescription of other psychopharmacological treatments (antipsychotics, antiepileptics and antidepressants) between the intervention group and the control group (all p values > 0.29). In addition, there were no differences in the rate of prescription or discontinuation of other psychopharmacological treatments (antipsychotics, antiepileptics and antidepressants) between the intervention group and the control group (all p values > 0.52). Further adjustment of the analyses on the primary outcomes (HDRS and YMRS) for the use of other psychopharmacological treatments (antipsychotics, antiepileptics and antidepressants) did not alter the estimates.

The two intervention groups were lumped together for the reported statistical analyses. Analyses according to each of the three randomization groups did not affect the estimates reported in Tables 2 and 3. There were no statistically significant differences in any of the outcome measures or in any of the statistical analyses (including sub-analyses) between the two groups

https://doi.org/10.1017/S0033291719000710 Published online by Cambridge University Press

Table 3. Estimated differences in secondary and tertiary outcomes between the intervention group and the control group (control group serve as reference), $N = 129$

	Unadjusted[a]			Model 1[b]			Model 2[c]		
	Difference between groups	95% CI	*p*	Difference between groups	95% CI	*p*	Difference between groups	95% CI	*p*
FAST	0.54	−2.36 to 3.43	0.72	0.36	−2.49 to 3.20	0.81	−0.92	−3.01 to 1.17	0.39
PSS	−1.72	−4.74 to 1.29	0.26	−1.85	−4.82 to 1.12	0.22	−2.08	−4.15 to −0.01	0.049
WHOQoL-bref	3.63	−1.26 to 8.51	0.15	3.77	−1.13 to 8.61	0.13	4.00	0.04–7.97	0.048
BDI	−0.61	−3.29 to 2.06	0.65	−0.67	−3.96 to 2.01	0.62	–	–	–
ASRM	0.08	−0.54 to 0.69	0.81	0.035	−0.56 to 0.63	0.91	–	–	–
Recovery	−0.12	−6.25 to 6.01	0.97	1.20	−4.51 to 6.91	0.68	2.13	−2.98 to 7.34	0.41
Empowerment	−0.23	−2.60 to 2.15	0.85	−0.19	−2.54 to 2.15	0.87	0.10	−2.00 to 2.19	0.93
Adherence to medication	0.04	−0.53 to 0.61	0.90	0.05	−0.40 to 0.49	0.84	0.05	−0.37 to 0.47	0.81
	Odds ratio	95% CI	*p*	Odds ratio	95% CI	*p*			
Antipsychotics prescribed	1.48	0.71–3.08	0.29	1.38	0.66–2.91	0.40	–	–	–
Anticonvulsants prescribed	1.59	0.76–3.33	0.22	1.50	0.71–3.18	0.29	–	–	–
Antidepressants prescribed	1.20	0.49–2.91	0.69	1.05	0.42–2.62	0.92	–	–	–
Lithium prescribed	2.79	1.32–5.90	0.007	2.87	1.33–6.20	0.007	–	–	–

FAST, The Functional Assessment Short Test; PSS, Cohen's Perceived Stress Scale; WHOQoL-bref, World Health Organization Quality of Life – short version; BDI, Beck's Depressive Inventory; ASRM, The Altman Self Rating scale for Mania; Recovery, The Recovery Assessment Scale; Empowerment, Roger's Empowerment Scale; Adherence to medication, The Medicine Adherence Rating Scale.

[a]Adjusted for outcome variable at baseline.

[b]Adjusted for outcome variable at baseline, age and gender unless otherwise specified.

[c]Adjusted for variable at baseline, age, gender, HDRS score and YMRS score.

receiving the active smartphone-based intervention (all p value > 0.30). In addition, there were no statistically significant differences in any of the baseline characteristics between any of the three groups (including the two groups receiving the active smartphone-based intervention) (p value > 0.78).

Discussion

The present paper presents results from the first RCT investigating the effect of smartphone-based patient-reported and objective monitoring including a mood prediction system used with a clinical feedback in patients with a psychiatric disorder. In this single-blinded trial, patients with BD were randomized to a smartphone-based monitoring system including a feedback loop or to the use of a smartphone for usual communication for 9 months. Overall, there was no effect on clinically rated depressive and manic symptoms. However, patients in the intervention group reported significantly higher quality of life and lower perceived stress compared with the control group. Since the effect of smartphone-based monitoring in BD has been sparingly investigated and based on interesting findings from the previous MONARCA I trial, sub-analyses in relation to the primary outcome were defined *a priori* and employed (Faurholt-Jepsen *et al.*, 2014*b*).

Exploratory analyses suggested that smartphone-based monitoring may reduce the risk of relapse of manic episodes but increase the risk of relapse of depressive episodes. The exploratory sub-analyses were based on small number of patients, and may be chance findings, but interestingly some of the *a priori* defined analyses resulted in findings in line with findings from our previous trial (Faurholt-Jepsen *et al.*, 2015*a*).

Currently, approximately 1/3 of the world's adult population owns and uses a smartphone, and it is estimated that by 2018, this proportion will increase to 50% (FierceWireless, 2011). During recent years, there have been a rapidly increasing international interest in and use of mhealth technologies within mental health (Smith, 2012; Anthes, 2016; Faurholt-Jepsen *et al.*, 2016*b*; Insel, 2017; Wang *et al.*, 2018). However, the scientific evidence base is small as there are yet few RCTs investigating the effect of smartphone-based monitoring in patients with a psychiatric disorder [e.g. psychosis (Krzystanek *et al.*, 2018), anxiety (Stolz *et al.*, 2018), depression (Ben-Zeev *et al.*, 2018; Furukawa *et al.*, 2018) and BD (Depp *et al.*, 2015; Faurholt-Jepsen *et al.*, 2015*a*)], and these are limited by methodological issues such as small samples, short follow-up periods and including unblinded patient-reported outcome measures. Within BD, RCTs investigating the effect of mhealth interventions are strikingly few, with most trials investigating the effect of web-based interventions with varying contents (Lieberman *et al.*, 2010; Todd *et al.*, 2014; Barnes *et al.*, 2015; Lauder *et al.*, 2015), and only two trials used smartphone as the monitoring tool (Depp *et al.*, 2015; Faurholt-Jepsen *et al.*, 2015*a*). Of these, only one investigated the effect of smartphone-based monitoring and clinical feedback compared with standard treatment (Faurholt-Jepsen *et al.*, 2015*a*), whereas the other investigated the effect of smartphone-based delivered interactive intervention linking patient-reported mood states with personalized self-management strategies without feedback compared with paper-and-pencil mood monitoring (Depp *et al.*, 2015). The MONARCA I RCT was the first to investigate the effect of smartphone-based monitoring including a feedback loop between patients and clinicians compared with standard treatment on the level of depressive and manic symptoms in patients with BD (Faurholt-Jepsen *et al.*, 2015*a*). Findings from sub-analyses suggested that patients using the MONARCA system had more sustained depressive symptoms than the control group but had fewer manic symptoms. Interestingly, the findings from the present MONARCA II trial were in line with these findings. It may be that identifying and treating manic prodromes may be easier than for depressive prodromes (Perry *et al.*, 1999; Simon *et al.*, 2006), and that continuous monitoring of depressive symptoms may have a detrimental effect. Although the idea of electronic self-monitoring and mood prediction seems appealing, studies using rigorous methodology carefully investigating beneficial as well as possible harmful effects of electronic monitoring are needed.

Given the limited access to appropriate treatment facilities, smartphone-based monitoring represents a flexible and adjustable system that could potentially be of great support for both patients and health care providers (Musiat *et al.*, 2014). Continuous availability and support with tailored content of intervention may engage patients who would not seek help through traditional routes. Continuous monitoring could possibly increase the patients' illness insight and awareness of sub-syndromal symptoms, which have been associated with poor prognostic factors including impaired functioning and high risk of relapse (Judd *et al.*, 2003; Strejilevich *et al.*, 2013; Patel *et al.*, 2015; O'Donnell *et al.*, 2018), and alertness of early warning signs of upcoming depressive or manic episodes. This will provide new opportunities for early intervention in outpatient settings which has been shown to reduce the risk of rehospitalization (Kessing *et al.*, 2013). In this way, it is possible that outpatient treatment can be optimized, and that the frequency of necessary health care provider/clinician and other clinical staff visits can be reduced and be more flexible according to the needs of the patients.

Limitations

The overall findings from the present trial are negative. It could be that there is an effect of smartphone-based monitoring on other outcomes (e.g. readmissions, mood instability, illness insight, rumination, activity) which we did not include, or which the clinical rating scales used in the present trial (the HDRS and the YMRS) did not capture comprehensively enough. However, the HDRS and the YMRS are currently the international golden standard to evaluate the severity of depressive or manic symptoms in BD and will therefore give the opportunity to compare results across trials and include in potential future meta-analyses. Ongoing studies investigating the effect of smartphone-based monitoring on rate of readmission, mood instability and rumination in patients with BD will hopefully provide more insight (Hidalgo-Mazzei *et al.*, 2015; Saunders *et al.*, 2016; Faurholt-Jepsen *et al.*, 2017; Mühlbauer *et al.*, 2018).

There may be different effects of the different components of the monitoring system. The MONARCA II trial was designed to investigate the effect of the total Monsenso system. Thus, it was not possible to address the possible positive or negative effects of the individual components of the intervention. Future trials including groups with different levels of smartphone-based monitoring could be able to investigate the effects of individual components. Since the objective of the present trial was to investigate the effect of smartphone-based monitoring compared with standard treatment, we did not conduct *within* groups investigating differences in outcome measures from baseline to the end of the trial in the individual groups.

https://doi.org/10.1017/S0033291719000710 Published online by Cambridge University Press

The feedback loop was based on prediction models based on subjective measures and on smartphones capable of collecting these data also on objective smartphone data. Previous studies by the authors have shown that, even in patients with rather severe levels of depressive and manic symptoms, the self-monitored data as well as the objective smartphone data reflect clinically rated depressive and manic symptoms (Faurholt-Jepsen *et al.*, 2015*b*, 2016*a*, 2016*c*). The prediction models in the feedback loop were adjusted during the trial (a learning model) allowing for optimal titration of the intervention and to achieve the best prediction possible (Insel, 2017). However, during the trial there were periods with technical problems regarding the prediction models, and thus feedback based on these data was functioning less optimal.

The patients presented with a rather low level of depressive and manic symptoms making it harder to demonstrate a potential effect on depressive and manic symptoms (Table 1). Including patients presenting with more severe symptoms could have resulted in other findings, and will be investigated further by the authors (Faurholt-Jepsen *et al.*, 2017). A previous study by the authors found that patients were able to validly evaluate the level of depressive and manic symptoms during more severe episodes than in the present study and further, that patients tolerated and adhered to the self-monitoring during these states (Faurholt-Jepsen *et al.*, 2016*c*).

Due to the type of intervention, it was not possible to blind the patients, the clinicians or the study nurse to randomization group. However, the researchers collecting clinically evaluated outcome data (HDRS, YMRS and FAST) were blinded to randomization group. The reported tertiary outcomes (e.g. perceived stress and quality of life) were due to the design of the trial collected unblinded to randomization and therefore prone to detection bias. Thus, given that patients in the intervention group knew that they were being monitored, this might have given an additional subject sensation of control over their symptoms and support (less perceived stress and higher quality of life).

In any non-pharmacological treatment trial, it is always a challenge to define a control group. The trial was designed to eliminate the effect of simply using a smartphone by including a control group of patients using a smartphone for usual communication.

Generalizability

An RCT represents a study design with high interval validity, with a possible cost of lower external validity and lower generalizability of the results. The present trial used a single-blinded design, with data on outcome measures collected by a researcher not aware of randomization group and therefore data were not affected by bias. The trial included patients treated in community centres, private psychiatrists or general practice at the time of the trial. Further, the trial had a pragmatic design with few exclusion criteria. The follow-up period was quite long. Thus, the results from the present trial reflect the use of smartphone-based monitoring during clinical settings. In addition, the Monsenso system is easy to use, and user-friendly with a high usability (Bardram *et al.*, 2012, 2013; Faurholt-Jepsen *et al.*, 2013, 2014). The trial had a high retention rate, which could be due to the pragmatic design and few requirements from the participants during the trial. Also, patients expressed that they wished to contribute to the lack of evidence regarding smartphone-based monitoring in BD. Overall, the findings of the present trial are found to be generalizable to patients with BD in general.

Conclusion

Smartphone-based monitoring and real-time mood prediction, including a clinical feedback loop, did not reduce the severity of depressive and manic symptoms in the present trial. Exploratory sub-group analyses showed that the intervention group had increased risk of relapse of depressive episodes compared with the control group. In exploratory analysis of patients without mixed symptoms, the intervention group had statistically significantly fewer manic symptoms and lower risk of manic episodes. Findings from sub-analyses were based on small numbers, could represent chance findings and should be interpreted with caution. Patient-evaluated unblinded data showed that the intervention group experienced higher quality of life and lower levels of perceived stress compared with the control group. Potential negative effects of smartphone-based monitoring and mood prediction should be investigated carefully before implemented as a standard clinical tool.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291719000710.

Author ORCIDs. Maria Faurholt-Jepsen, 0000-0002-0462-6444.

Acknowledgements. The authors would like to thank the patients for participating in the trial and nurse Bente Støyer for running the integrated clinical feedback loop. The trial received financial support from The Mental Health Services in the Capital Region of Denmark including the Centre for Telemedicine in the Capital Region of Denmark and EU 7th Frame Programme, IT University of Copenhagen and the Lundbeck Foundation. The research team has no financial connection to any of the contributors.

Author contributions. MFJ and LVK conceived and conducted the trial, analysed data and authored the first draft of the manuscript. MV, MF and JB have been revising and optimising the manuscript. All authors contributed to and approved the final version of the manuscript.

Financial support. The funding organizations had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; and decision to submit the manuscript for publication.

Conflict of interest. MFJ has no competing interests. MF and JEB are co-founders and shareholders in Monsenso ApS. EMC has been a consultant for Eli Lilly, Astra Zeneca, Servier, Bristol-Myers Squibb, Lundbeck and Medilink. MV has been a consultant for Lundbeck within the past 3 years. LVK has been a consultant for Sunovion.

References

Adli M, Whybrow PC, Grof P, Rasgon N, Gyulai L, Baethge C, Glenn T and Bauer M (2005) Use of polypharmacy and self-reported mood in outpatients with bipolar disorder. *International Journal of Psychiatry in Clinical Practice* **4**, 251–256.

Altman EG, Hedeker D, Peterson JL and Davis JM (1997) The Altman Self-Rating Mania Scale. *Biological Psychiatry* **42**, 948–955.

Angst J and Sellaro R (2000) Historical perspectives and natural history of bipolar disorder. *Biological Psychiatry* **48**, 445–457.

Anthes E (2016) Mental health: there's an app for that. *Nature* **532**, 20–23.

Bardram J, Frost M, Szanto K and 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*). *ACM, New York, NY, USA, 21-30.* vol 2012, pp 21–30. ACM New York, NY, USA.

Bardram JE, Frost M, Szánto K, Faurholt-Jepsen M, Vinberg M and 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, pp2627-2636. New York, NY, USA, 2013.

Barnes CW, Hadzi-Pavlovic D, Wilhelm K and Mitchell PB (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 and Whybrow PC (2004) Using technology to improve longitudinal studies: self-reporting with ChronoRecord in bipolar disorder. *Bipolar Disorders* **6**, 67–74.

Bauer M, Rasgon N, Grof P, Gyulai L, Glenn T and Whybrow PC (2005) Does the use of an automated tool for self-reporting mood by patients with bipolar disorder bias the collected data? *MedGenMed: Medscape General Medicine* **7**, 21.

Bauer M, Grof P, Rasgon N, Glenn T, Alda M, Priebe S, Ricken R and Whybrow PC (2006) Mood charting and technology: new approach to monitoring patients with mood disorders. *Current Psychiatry Reviews* **2**, 423–429.

Bauer M, Glenn T, Grof P, Pfennig A, Rasgon NL, Marsh W, Munoz RA, Sagduyu K, Alda M, Quiroz D, Sasse J and Whybrow PC (2007) Self-reported data from patients with bipolar disorder: frequency of brief depression. *Journal of Affective Disorders* **101**, 227–233.

Bauer M, Wilson T, Neuhaus K, Sasse J, Pfennig A, Lewitzka U, Grof P, Glenn T, Rasgon N, Bschor T and Whybrow PC (2008) Self-reporting software for bipolar disorder: validation of ChronoRecord by patients with mania. *Psychiatry Research* **159**, 359–366.

Bech P, Rasmussen NA, Olsen LR, Noerholm V and 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.

Ben-Zeev D, Brian RM, Jonathan G, Razzano L, Pashka N, Carpenter-Song E, Drake RE and Scherer EA (2018) Mobile health (mHealth) versus clinic-based group intervention for people with serious mental illness: a randomized controlled trial. *Psychiatric Services* **69**, 978–985.

Bilderbeck AC, Atkinson LZ, McMahon HC, Voysey M, Simon J, Price J, Rendell J, Hinds C, Geddes JR, Holmes E, Miklowitz DJ and Goodwin GM (2016) Psychoeducation and online mood tracking for patients with bipolar disorder: a randomised controlled trial. *Journal of Affective Disorders* **205**, 245–251.

Bopp JM, Miklowitz DJ, Goodwin GM, Stevens W, Rendell JM and 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 and Ravaud P (2008) Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Annals of Internal Medicine* **148**, 295–309.

Chan A-W, Tetzlaff JM, Altman DG, Dickersin K and Moher D (2013) SPIRIT 2013: new guidance for content of clinical trial protocols. *Lancet* **381**, 91–92.

Chinman M, Young AS, Schell T, Hassell J and Mintz J (2004) Computer-assisted self-assessment in persons with severe mental illness. *The Journal of Clinical Psychiatry* **65**, 1343–1351.

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

Corrigan PW, Salzer M, Ralph RO, Sangster Y and Keck L (2004) Examining the factor structure of the recovery assessment scale. *Schizophrenia Bulletin* **30**, 1035–1041.

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

Depp CA, Kim DH, de Dios LV, Wang V and 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, Ceglowski J, Wang VC, Yaghouti F, Mausbach BT, Thompson WK and Granholm EL (2015) Augmenting psychoeducation with a mobile intervention for bipolar disorder: a randomized controlled trial. *Journal of Affective Disorders* **174**, 23–30.

Donker T, Petrie K, Proudfoot J, Clarke J, Birch M-R and Christensen H (2013) Smartphones for smarter delivery of mental health programs: a systematic review. *Journal of Medical Internet Research* **15**, e247.

Faurholt-Jepsen M, Brage S, Vinberg M, Christensen EM, Knorr U, Jensen HM and Kessing LV (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 EM, Frost M, Bardram J and 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. *BMJ Open* **3**, 7.

Faurholt-Jepsen M, Frost M, Vinberg M, Christensen EM, Bardram JE and Kessing LV (2014*a*) Smartphone data as objective measures of bipolar disorder symptoms. *Psychiatry Research* **217**, 1–2.

Faurholt-Jepsen M, Vinberg M, Frost M, Christensen EM, Bardram J and Kessing LV (2014*b*) Daily electronic monitoring of subjective and objective measures of illness activity in bipolar disorder using smartphones – the MONARCA II trial protocol: a randomized controlled single-blind parallel-group trial. *BMC Psychiatry* **14**, 309.

Faurholt-Jepsen M, Frost M, Ritz C, Christensen EM, Jacoby AS, Mikkelsen RL, Knorr U, Bardram JE, Vinberg M and Kessing LV (2015*a*) Daily electronic self-monitoring in bipolar disorder using smartphones – the MONARCA I trial: a randomized, placebo-controlled, single-blind, parallel group trial. *Psychological Medicine* **45**, 2691–2704.

Faurholt-Jepsen M, Vinberg M, Frost M, Christensen EM, Bardram JE and Kessing LV (2015*b*) Smartphone data as an electronic biomarker of illness activity in bipolar disorder. *Bipolar Disorders* **17**, 715–728.

Faurholt-Jepsen M, Busk J, Frost M, Vinberg M, Christensen EM, Winther O, Bardram JE and Kessing LV (2016*a*) Voice analysis as an objective state marker in bipolar disorder. *Translational Psychiatry* **6**, e856.

Faurholt-Jepsen M, Munkholm K, Frost M, Bardram JE and Kessing LV (2016*b*) Electronic self-monitoring of mood using IT platforms in adult patients with bipolar disorder: a systematic review of the validity and evidence. *BMC Psychiatry* **16**, 7.

Faurholt-Jepsen M, Vinberg M, Frost M, Debel S, Margrethe Christensen E, Bardram JE and Kessing LV (2016*c*) Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder. *International Journal of Methods in Psychiatric Research* **25**, 309–323.

Faurholt-Jepsen M, Frost M, Martiny K, Tuxen N, Rosenberg N, Busk J, Winther O, Bardram JE and Kessing LV (2017) Reducing the rate and duration of Re-ADMISsions among patients with unipolar disorder and bipolar disorder using smartphone-based monitoring and treatment – the RADMIS trials: study protocol for two randomized controlled trials. *Trials* **18**, 277.

Faurholt-Jepsen M, Busk J, Þórarinsdóttir H, Frost M, Bardram JE, Vinberg M and Kessing LV (2018) Objective smartphone data as a potential diagnostic marker of bipolar disorder. *Australian & New Zealand Journal of Psychiatry* **53**, 119–128.

FierceWireless (2011) *Report: Global smartphone penetration to jump 25% in 2014, led by Asia-Pacific*. FierceWireless.

Furukawa TA, Horikoshi M, Fujita H, Tsujino N, Jinnin R, Kako Y, Ogawa S, Sato H, Kitagawa N, Shinagawa Y, Ikeda Y, Imai H, Tajika A, Ogawa Y, Akechi T, Yamada M, Shimodera S, Watanabe N, Inagaki M and Hasegawa A (2018) Cognitive and behavioral skills exercises completed by patients with major depression during smartphone cognitive behavioral therapy: secondary analysis of a randomized controlled trial. *JMIR Mental Health* **5**, e4.

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

Harrison PJ, Cipriani A, Harmer CJ, Nobre AC, Saunders K, Goodwin GM and Geddes JR (2016) Innovative approaches to bipolar disorder and its treatment. *Annals of the New York Academy of Sciences* **1366**, 76–89.

Hidalgo-Mazzei D, Mateu A, Reinares M, Undurraga J, Bonnín Cdel M, Sánchez-Moreno J, Vieta E and Colom F (2015) Self-monitoring and psychoeducation in bipolar patients with a smart-phone application (SIMPLe) project: design, development and studies protocols. *BMC Psychiatry* **15**, 52.

https://doi.org/10.1017/S0033291719000710 Published online by Cambridge University Press

Insel TR (2017) Digital phenotyping: technology for a new science of behavior. *JAMA* **318**, 1215–1216.

Judd LL, Schettler PJ, Akiskal HS, Maser J, Coryell W, Solomon D, Endicott J and Keller M (2003) Long-term symptomatic status of bipolar I vs. bipolar II disorders. *The International Journal of Neuropsychopharmacology* **6**, 127–137.

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

Kessing LV, Hansen HV, Hvenegaard A, Christensen EM, Dam H, Gluud C and 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. *The British Journal of Psychiatry* **202**, 212–219.

Krzystanek M, Borkowski M, Skałacka K and Krysta K (2018) A telemedicine platform to improve clinical parameters in paranoid schizophrenia patients: results of a one-year randomized study. *Schizophrenia Research* **204**, 389–396.

Kuhs H and Reschke D (1992) Psychomotor activity in unipolar and bipolar depressive patients. *Psychopathology* **25**, 109–116.

Kupfer DJ, Weiss BL, Foster G, Detre TP and 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 DW and Berk M (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 DZ, Kelly TF, Douglas L and Goodwin FK (2010) A randomized comparison of online and paper mood charts for people with bipolar disorder. *Journal of Affective Disorders* **124**, 85–89.

Lieberman DZ, Swayze S and Goodwin FK (2011) An automated Internet application to help patients with bipolar disorder track social rhythm stabilization. *Psychiatric Services* **62**, 1267–1269.

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

Mühlbauer E, Bauer M, Ebner-Priemer U, Ritter P, Hill H, Beier F, Kleindienst N and Severus E (2018) Effectiveness of smartphone-based ambulatory assessment (SBAA-BD) including a predicting system for upcoming episodes in the long-term treatment of patients with bipolar disorders: study protocol for a randomized controlled single-blind trial. *BMC Psychiatry* **18**, 349.

Musiat P, Goldstone P and Tarrier N (2014) Understanding the acceptability of e-mental health – attitudes and expectations towards computerised self-help treatments for mental health problems. *BMC Psychiatry* **14**, 109.

Mutschler J, von Zitzewitz F, Rossler W and Grosshans M (2012) Application of electronic diaries in patients with schizophrenia and bipolar disorders. *Psychiatria Danubina* **24**, 206–210.

O'Donnell LA, Ellis AJ, de Loo MMV, Stange JP, Axelson DA, Kowatch RA, Schneck CD and Miklowitz DJ (2018) Mood instability as a predictor of clinical and functional outcomes in adolescents with bipolar I and bipolar II disorder. *Journal of Affective Disorders* **236**, 199–2206.

Patel R, Lloyd T, Jackson R, Ball M, Shetty H, Broadbent M, Geddes JR, Stewart R, McGuire P and Taylor M (2015) Mood instability is a common feature of mental health disorders and is associated with poor clinical outcomes. *BMJ Open* **5**, e007504.

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

Pini S, de Queiroz V, Pagnin D, Pezawas L, Angst J, Cassano GB and Wittchen H-U (2005) Prevalence and burden of bipolar disorders in European countries. *European Neuropsychopharmacology* **15**, 425–434.

Rogers ES, Chamberlin J, Ellison ML and Crean T (1997) A consumer-constructed scale to measure empowerment among users of mental health services. *Psychiatric Services* **48**, 1042–1047.

Rosa AR, Sánchez-Moreno J, Martínez-Aran A, Salamero M, Torrent C, Reinares M, Comes M, Colom F, Van Riel W, Ayuso-Mateos JL, Kapczinski F and Vieta E (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.

Saunders KEA, Cipriani A, Rendell J, Attenburrow M-J, Nelissen N, Bilderbeck AC, Vasudevan SR, Churchill G, Goodwin GM, Nobre AC, Harmer CJ, Harrison PJ and Geddes JR (2016) Oxford lithium Trial (OxLith) of the early affective, cognitive, neural and biochemical effects of lithium carbonate in bipolar disorder: study protocol for a randomised controlled trial. *Trials* **17**.

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

Smith A (2012) *Cell internet use* 2012. *A project of the Pew Research Center*. Available at pew.internet.org/Reports/2012/Cell-Internet-Use-2012/Key-Findings.aspxpew.internet.org/Reports/2012/Cell-Internet-Use-2012/Key-Findings.aspx.

Stolz T, Schulz A, Krieger T, Vincent A, Urech A, Moser C, Westermann S and Berger T (2018) A mobile app for social anxiety disorder: a three-arm randomized controlled trial comparing mobile and PC-based guided self-help interventions. *Journal of Consulting and Clinical Psychology* **86**, 493–504.

Strejilevich SA, Martino DJ, Murru A, Teitelbaum J, Fassi G, Marengo E, Igoa A and Colom F (2013) Mood instability and functional recovery in bipolar disorders. *Acta Psychiatrica Scandinavica* **128**, 194–202.

Thompson K, Kulkarni J and Sergejew AA (2000) Reliability and validity of a new Medication Adherence Rating Scale (MARS) for the psychoses. *Schizophrenia Research* **42**, 241–247.

Todd NJ, Jones SH, Hart A and Lobban FA (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.

Wang K, Varma DS and Prosperi M (2018) A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. *Journal of Psychiatric Research* **107**, 73–78.

Weinstock LM and Miller IW (2008) Functional impairment as a predictor of short-term symptom course in bipolar I disorder. *Bipolar Disorders* **10**, 437–442.

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

WHO (2011) *mHealth: New Horizons for Health Through Mobile Technologies*. Geneva: WHO. Available at http://www.who.int/goe/publications/goe_mhealth_web.pdf.

Whybrow PC, Grof P, Gyulai L, Rasgon N, Glenn T and 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 and Sartorius N (1990) SCAN. Schedules for Clinical Assessment in Neuropsychiatry. *Archives of General Psychiatry* **47**, 589–593.

World Health Organisation W (2011) *ICD-10: International Statistical Classification of Diseases and Related Health Problems*, vol **2011**. Geneva: World Health Organization.

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