Original Paper

Smartphone-Based Monitoring of Objective and Subjective Data in Affective Disorders: Where Are We and Where Are We Going? Systematic Review

Ezgi Dogan[1], MD; Christian Sander[1,2], PhD; Xenija Wagner[1], MD; Ulrich Hegerl[1,2], MD; Elisabeth Kohls[1], PhD

[1]Medical Faculty, Department of Psychiatry and Psychotherapy, University Leipzig, Leipzig, Germany

[2]Depression Research Centre, German Depression Foundation, Leipzig, Germany

Corresponding Author:
Christian Sander, PhD
Medical Faculty
Department of Psychiatry and Psychotherapy
University Leipzig
Semmelweisstrasse 10, Haus 13
Leipzig, 04103
Germany
Phone: 49 341 9724558
Fax: 49 341 9724539
Email: Christian.Sander@medizin.uni-leipzig.de

Abstract

Background: Electronic mental health interventions for mood disorders have increased rapidly over the past decade, most recently in the form of various systems and apps that are delivered via smartphones.

Objective: We aim to provide an overview of studies on smartphone-based systems that combine subjective ratings with objectively measured data for longitudinal monitoring of patients with affective disorders. Specifically, we aim to examine current knowledge on: (1) the feasibility of, and adherence to, such systems; (2) the association of monitored data with mood status; and (3) the effects of monitoring on clinical outcomes.

Methods: We systematically searched PubMed, Web of Science, PsycINFO, and the Cochrane Central Register of Controlled Trials for relevant articles published in the last ten years (2007-2017) by applying Boolean search operators with an iterative combination of search terms, which was conducted in February 2017. Additional articles were identified via pearling, author correspondence, selected reference lists, and trial protocols.

Results: A total of 3463 unique records were identified. Twenty-nine studies met the inclusion criteria and were included in the review. The majority of articles represented feasibility studies (n=27); two articles reported results from one randomized controlled trial (RCT). In total, six different self-monitoring systems for affective disorders that used subjective mood ratings and objective measurements were included. These objective parameters included physiological data (heart rate variability), behavioral data (phone usage, physical activity, voice features), and context/environmental information (light exposure and location). The included articles contained results regarding feasibility of such systems in affective disorders, showed reasonable accuracy in predicting mood status and mood fluctuations based on the objectively monitored data, and reported observations about the impact of monitoring on clinical state and adherence of patients to the system usage.

Conclusions: The included observational studies and RCT substantiate the value of smartphone-based approaches for gathering long-term objective data (aside from self-ratings to monitor clinical symptoms) to predict changes in clinical states, and to investigate causal inferences about state changes in patients with affective disorders. Although promising, a much larger evidence-base is necessary to fully assess the potential and the risks of these approaches. Methodological limitations of the available studies (eg, small sample sizes, variations in the number of observations or monitoring duration, lack of RCT, and heterogeneity of methods) restrict the interpretability of the results. However, a number of study protocols stated ambitions to expand and intensify research in this emerging and promising field.

XSL•FO
RenderX

KEYWORDS

review; mood disorders; smartphone; ecological momentary assessment

Introduction

Unipolar depressive disorder is one of the leading chronic medical conditions in Europe, accounting for 13.7% of the disability burden [1], and is projected by the World Health Organization to become the second highest condition leading to burden of disease by 2030 [2]. International lifetime prevalence rates of depression range from 6.3-10.3% [3], and comorbidity [3] and mortality [4,5] are common; these disorders also incur a considerable economic burden [6]. Unipolar depression often takes a chronic recurrent course, with more than 70% of patients suffering several episodes throughout their lifetimes [7,8]. Bipolar disorder affects 2-2.5% of the adult population [9,10] and is considered a chronic disorder [11]. Both unipolar and bipolar affective disorders have severe consequences on functioning and quality of life [12]. Although effective treatment is available [13], it is estimated that at least 50% of patients with depression receive no treatment at all [14,15]. Furthermore, most patients do not adhere to medication regimens and referrals [16,17]. Barriers for help-seeking and treatment adherence include stigmatization, fear of discrimination, and an unavailability of services [18,19].

Electronic mental health (e-mental health) interventions represent a promising means of overcoming such barriers and increasing the capacity for patients' self-management of depression [20]. Using the Internet to deliver treatment for affective disorders has been shown to be an effective option for reaching patients who were not able to receive face-to-face treatment due to geographical or other situational barriers [21], or to augment face-to-face therapy [22]. Internet availability in developed countries has expanded rapidly and now reaches individuals in nearly all social, age, ethnic, and education-level groups [23]. It is of promise for e-mental health interventions that people suffering from mental disorders actively use the Internet just as often as those in the general population who do not have mental illnesses (80%) [24]. A newer technological development is the rapid spread and technological improvement of smartphones. In the first quarter of 2016, there were 7.4 billion mobile phone subscriptions worldwide, of which 3.4 billion were for smartphones [25]. These numbers are estimated to reach 9 billion mobile phone and 6.3 billion smartphone subscriptions by the year 2021 [25]. Over 80% of the general population in Western countries already use mobile phones [23], thereby passively collecting an extensive amount of data that is not yet sufficiently used for medical and scientific purposes.

Capitalizing on these technological advances, a novel approach in e-mental health has emerged with the development and testing of several technology-enhanced systems for affective disorders that rely on combinations of subjective (ie, self-reported) and objective data, measured by sensors embedded within (or connected to) smartphones. When considering subjective self-ratings, smartphone apps result in better adherence and less missing data compared with paper-pencil or personal computer-based questionnaires or diaries [26]. However, the strongest asset of smartphone-based systems may be their high

suitability for collecting objective data from behavioral and physiological monitoring. Differences in subjective (ie, psychological) and objective (ie, behavioral or physiological) data are a common finding (eg, pronounced differences in self-ratings on sleep quality compared to objectively recorded sleep quality are often seen in subjects with sleep disorders, such as insomniac patients who typically overestimate their actual wake times during the night). This issue has sparked a discussion about whether or not both approaches are alternatives or complement each other by capturing different aspects of the respective construct (a detailed discussion of this topic can be found in Tams et al [27]). Using both self-reported and objectively measured data derived from the smartphone-based systems may help to better monitor disease progression and optimize self-management approaches. Modern smartphones are able to incorporate a large number of various sensors (eg, global positioning system [GPS], Bluetooth, accelerometers, microphones), are carried mostly close to the body, have long lasting batteries, and contain large memory storage and powerful central processing units for recording and processing information [28,29]. Mobile technology has the vast advantage that it can be easily integrated into patients' lives as a comfortable, simple, time-unconstrained, user-friendly, economical, and noninvasive method of registering and monitoring relevant signs and symptoms. Such technologies can be used to provide continuous self-management and psycho-educational content tailored to the specific needs of each individual on the basis of data registered on their own smartphones [23].

Using objective data for symptom monitoring is especially relevant for affective disorders, in which reliable diagnostic and predictive biomarkers are still lacking and self-reported symptoms are not only influenced by recall bias, but also by a mood-state associated bias. Therefore, it is an important question whether long-term continuous monitoring of subjective and objective parameters combined with a time-series analysis could provide earlier and more reliable predictors and indicators of mood changes. Such findings could be helpful for the individual patient to improve their self-management and prevent relapses [23]. Although there is already an ever-growing number of systems (usually smartphone apps) focusing on monitoring depressive symptoms, as well as collection of objective data from smartphone or external sensors, only a small portion of these have been evaluated regarding feasibility, efficacy, and safety [23]. Considering the importance of this emerging field, the speed of innovations and new developments, it was our aim to review the state-of-the-art knowledge about smartphone-based monitoring systems for affective disorders and their benefits for treatment.

Objective

The aim of this systematic review is to examine the literature on smartphone-based systems that combine subjective ratings with objectively measured data for the self-monitoring of depressive symptoms. The following aspects will be covered: (1) feasibility and adherence, (2) the association of monitoring data with mood status, and (3) effects on clinical outcome.

Methods

Overview

Literature was systematically reviewed in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines [30]. This systematic review has been registered in the PROSPERO international prospective register of systematic reviews (registration number CRD42017058539). A PRISMA checklist is presented in Multimedia Appendix 1.

Information Source

In February 2017, the search for relevant articles published in the last ten years (2007-2017) was conducted in the following bibliographic databases: PubMed, Web of Science, PsycINFO, and Cochrane Central Register of Controlled Trials. Additional articles were identified via pearling, author correspondence, selected reference lists, and trial protocols.

Search Strategy

The search strategy was developed by three of the authors and included combinations of the search terms, such as *affective disorder*, *ecological momentary assessment*, *smartphone*, *e-mental health*, and *self-monitoring*. The search took place in six content-related categories and the application of iterative combinations of these categories by employing the Boolean search operators (see Multimedia Appendix 2 for the full search strategy).

Eligibility Criteria

Titles and abstracts were screened for eligibility and were included if: (1) abstracts were published in English or German, (2) any form of smartphone-assessed self-monitoring of depressive symptoms was mentioned, and (3) adult study participants with depressive symptoms (but not postpartum/postnatal depression or pregnant women) were investigated. In cases of insufficient information to determine eligibility, papers were nonetheless subjected to further screening. Considering the novelty of this research area, the review was planned to be as inclusive as possible. Therefore, experimental and observational studies were included.

Selection of Studies

We obtained full reports of all potentially relevant papers. Articles were discarded if they met at least one of the following exclusion criteria: (1) review papers only; (2) study protocols or descriptive planned studies or future study intents; (3) descriptive studies that did not include participants diagnosed with affective disorders (apart from postpartum/postnatal depression); (4) studies that did not monitor objective data involving a smartphone/mobile device (eg, computerized/Internet-delivered or Internet-based Cognitive Behavioral Therapy interventions or Ecological Momentary Assessment [EMA] studies in which only subjective ratings were obtained); or (5) other reasons (eg, comments, letters, poster abstracts, articles not written in English). Additional information from study authors was obtained when necessary to resolve questions about eligibility. Four authors (ED, EK, CS, and XW) screened the full text reports and decided whether

these papers met the inclusion criteria. Records were originally split among the four reviewers, each of whom made independent decisions regarding exclusion. Uncertainties and disagreement were resolved through discussion (ED, EK, CS, and UH). Reasons for excluding articles were recorded (including multiple assignments to exclusion criteria, if applicable). Finally, all exclusion assignments were reviewed and each record was allocated to one exclusion criteria (CS, XW, and EK).

Assessment of Methodological Quality

The Downs and Black Instrument [31], which is recommended for both randomized and nonrandomized studies, was used for assessing the quality of included studies. This 27-item checklist rates studies on the following subscales: reporting, external validity, internal validity-bias, internal validity-confounding (selection bias), and power. The total maximum score that can be achieved with this instrument is 32. High internal consistency, as well as good inter-rater variability and test-retest-reliability, were reported [32]. For the purpose of this study, the last question was modified by changing the score from 0-5 to a dichotomous score (0/1) equivalent to the procedure outlined by Hootman et al [32]. If a sample size or power calculation was provided, the final item was scored as 1. The item was scored 0 if sample size and power calculations (or explanations about the appropriateness of the sample size) were missing. Thus, the total maximum score of this modified checklist is 28.

Data Synthesis

Due to the novelty of the approach/research field and the heterogeneous nature of the included studies, a quantitative data synthesis (meta-analysis) has not been performed, since there was an insufficient amount of data to calculate effect sizes. Standardized forms were developed for data collection purposes and two authors (ED and EK) were trained to use and apply them. Piloting of the extraction forms with a subsample of randomly selected included studies also took place. To ensure consistency between reviewers, calibration exercises took place before starting the review. Data extraction was performed in duplicate by two authors (ED and EK), disagreement was resolved by discussion, and another author (CS) verified the information to reduce bias and errors in data extraction. We contacted study authors to resolve any uncertainties or for clarification, if needed. A systematic narrative synthesis will be provided with information presented in the text and tables to summarize and explain the characteristics and findings of the included studies.

Results

After the description of the selection process and characteristics of included systems and studies are detailed, the results section is structured according to the objectives of the review. The *Association of Monitoring Data With Mood Status* section lists the different parameters recorded by the smartphone-based systems included in the review and, where applicable, is extended by further evidence.

The electronic database search resulted in a total of 4122 records (PubMed, n=650; PsycINFO, n=378; Web of Science, n=1534; Cochrane Central Register of Controlled Trial, n=1491; and

n=69 identifications by cross-referencing), of which 3463 were unique articles. In the first selection step (abstract screening), 3173 records were excluded after it was determined that they did not meet inclusion criteria based on information in the abstract. The majority of these studies were excluded because they only contained descriptions of online/computerized Cognitive Behavioral Therapy systems, descriptions of pure online interventions, or self-monitoring or screening tools (using only subjective assessment and self-monitoring or smartphone-based ecological momentary assessments). A total of 290 full text papers were retrieved for further consideration. After reviewing these full texts, an additional 261 records were excluded. Figure 1 depicts a flowchart of the screening process (see Multimedia Appendix 3 for the list of excluded full text articles).

Figure 1. PRISMA Flow chart.
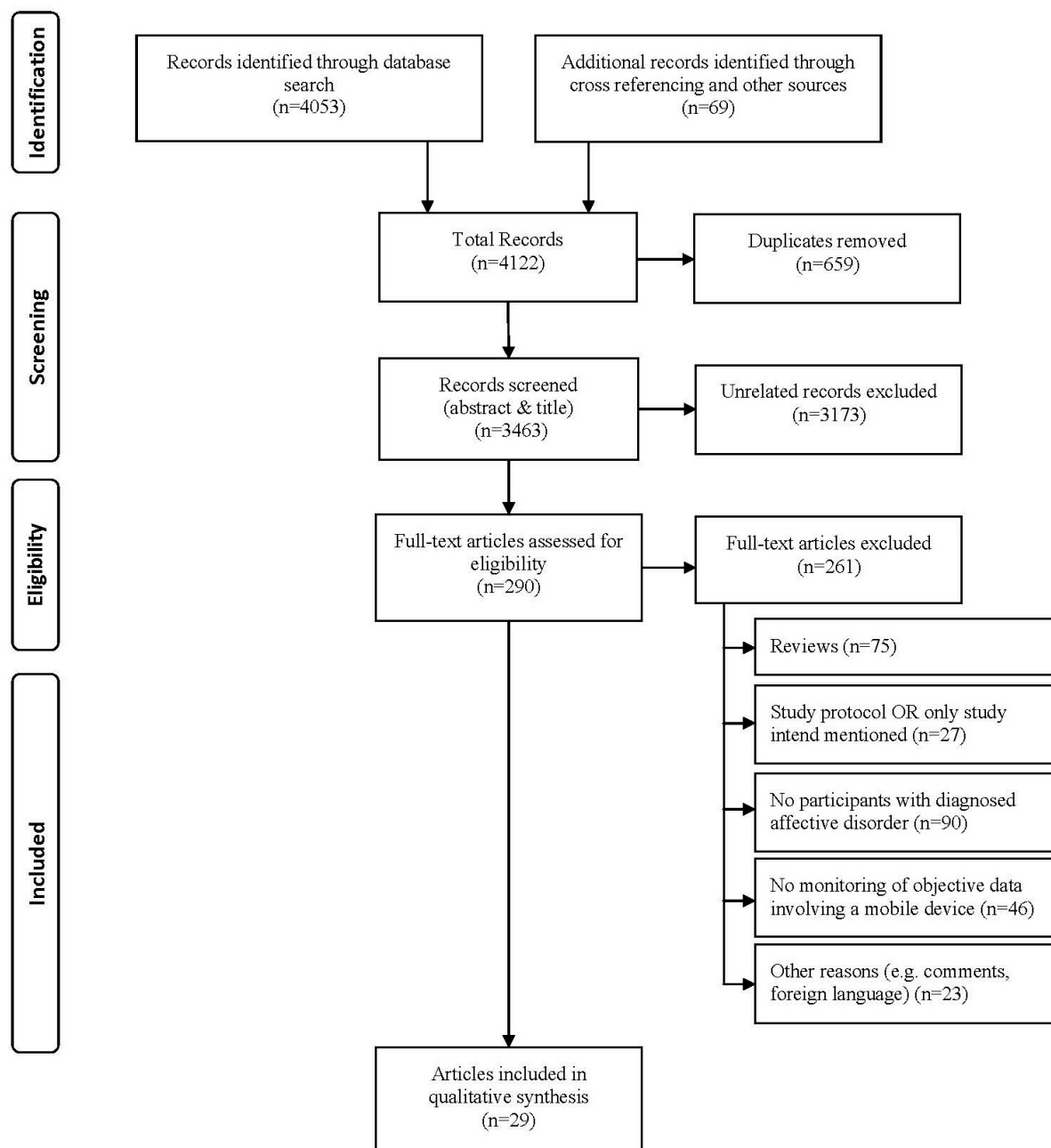
Characteristics of Included Systems and Studies

Twenty-nine studies were included for qualitative synthesis in this review. Although the search range spanned from 2007 to 2017, the included studies were all published after 2010, attesting to the novelty of the research field. Table 1 provides an overview of all systems that were described within the included articles. Most systems incorporated different data sources: in all cases, subjective ratings were required by the users (eg, ratings on mood, estimates on sleep, or information

on medication). These data were complemented by different bio-parameters (eg, electrocardiogram [ECG], respiration, skin conductance), which were either recorded by the smartphone itself (eg, accelerometer, microphone) or by additional wearables. Smartphone usage was monitored by many of the included systems, although different parameters were used (eg, screen state, number of incoming and outgoing calls). Smartphone sensors also provided information about context and environmental conditions that were taken into account (eg, light sensor, GPS).

A detailed description of the included studies is provided in Multimedia Appendix 4. Studies were conducted in the following countries: United States/England (n=10),

Austria/Germany (n=8), Denmark (n=8), and Italy/France (n=3). The majority of included studies used an observational design (n=27), as most were pilot or feasibility studies focusing on feasibility of, and adherence to, the respective monitoring systems, or were aimed at investigating the association of monitored data with clinical mood state. Only three studies [33-35] provided data on effects of monitoring on clinical outcomes; of those, only one used a randomized controlled trial (RCT) design [35]. Accordingly, the sample sizes ranged from 1 to 78. While 4 studies included patients with diagnosed major depression, 24 studies were conducted with patients with bipolar affective disorder. Two studies [36,37] also included schizophrenia patients in their study population with affective disorders.

Table 1. Overview of smartphone-based systems that combine subjective mood ratings and objective assessment of physiological and behavioral parameters for the monitoring of affective symptoms.

	Self-Report Data	Physiological Data	Behavioral Data			Context and Environmental Data		Type of clinical trials
			Phone Usage	Physical Activity	Voice features	Light Exposure	Location	
Empath [38]	+	+	-	+	+	-	-	Case Study
FINE [39]	+	-	+	+	-	-	+	Feasibility Study
MedLink [33][a]	+	-	-	-	-	-	-	Pilot Study
Mobilyze! [34]	+	-	+	+	-	+	+	Feasibility Study
MONARCA [26,35,40-44]	+	-	+	+	+	+	+	Feasibility Study, Pilot Study, RCT
MoodRhythm [45]	+	-	+	+	-	+	-	Pilot Study
PSYCHE [46-48]	+	+	-	+	+	-	-	Pilot Study
SIMBA [49]	+	-	+	+	-	-	+	Pilot Study
Unnamed System [50][b]	+	+	+	+	-	-	-	Feasibility Study
Unnamed System [36]	+	+[c]	-	+	-	-	-	Pilot Study

[a]Objective monitoring of medication adherence using a *Wisepill bottle*.

[b]System was utilized in an Automated Monitoring of Symptom Severity (AMoSS) study.

[c]An ingestion sensor was used for gathering the physiological data.

Results on Methodological Quality Assessment

According the Downs and Black instrument, the methodological quality scores of the studies ranged from 5 [44] to 22 [35]. Methodological bias of the included studies is presented in Multimedia Appendix 5. In most cases the included studies had low scores, as they were mostly conducted as single-arm observational trials, feasibility studies, or pilot studies, and were not randomized trials. In addition, studies reporting only qualitative data [37,39,50,51] also scored lower. Except for one study [36], no adverse effects were reported in the selected studies.

Results on Feasibility and Adherence

Among the included studies, large variations in drop-out rates were reported, ranging from 0% [26,33,39,44,46,48,50,52-56] to 22% [45]. Technical problems, including inappropriate operating systems [34,40,45,57], were reported to be the most common reason for discontinuation. Several articles reported on feasibility and discussed aspects of improving adherence. Prociow et al [51] reported large initial interest when approaching self-support groups and charities, but only three patients were willing to participate in their feasibility trial, of which two withdrew before a second visit. Compared to their healthy participants who were described as being "accustomed with mobile technology", the patient (although being euthymic during the study duration) adhered only poorly to the monitoring regimen due to self-reported difficulties with the management

of the unfamiliar technology. The system in this study was comprised of environmental as well as wearable sensors, with the latter causing a higher level of discomfort.

In a field trial of MONARCA 1.0 [40] the adherence rates for phone-based and paper-based self-assessments were found to be comparable. Moreover, high compliance scores of the four mandatory self-assessment parameters by phone-based self-assessment were reported. The patients in this field trial also rated the system usability and usefulness on a Likert scale from 1 (strongly agree) to 7 (strongly disagree). Average rates for these scores were: overall satisfaction=2.6, system usefulness=1.93, information quality=3.32, interface quality=2.71, usefulness of the system for disease management=3.16, and perceived usefulness for disease management=2.16. Taking into account the days on which the system was working, an adherence rate of 87% was reported for this field trial, with an increased adherence (91%) being reported for MONARCA 2.0 [44].

Beiwinkel et al [49] reported a compliance rate of 55.7% for filling in self-reporting data, with no decline over time, during an average study duration of 12 months. In a pilot study (n=28) assessing feasibility of a digital health feedback system using an ingestion sensor for physiologic assessments and confirmation of tablet ingestions, a medication adherence rate of 74% was documented [36]. Among the participants who completed the study (n=27), the concept was found easy to understand by 70%, 89% reported that the system might be useful for them, and 78% indicated that they would prefer to receive reminders on their smartphones in case of forgetting to take the medications [36]. In the pilot study (n=8) of the Medlink system [33], medication adherence was assessed through the *Wisepill bottle* which produces a signal when it is opened, thus providing real-time information about patient pill-taking activity. According to data gathered from the pillbox, 72.5-100% of doses were taken by study participants [33]. On a scale from 1 (strongly disagree) to 7 (strongly agree) the average rating for satisfaction was found to be 4.8 (standard deviation [SD]=0.8). Ease of use was rated 5.7 (SD=1.1), while learnability scored 6.1 (SD=1.5) and an average rate of 4.6 (SD=1.0) was found for usefulness.

The overall satisfaction of users with smartphone-based systems is reported to be rather positive (eg, the satisfaction with the Mobilyze! intervention was reported to be 5.71 [SD=1.38] on a scale ranging from 1 [strong disagreement] to 7 [strong agreement]) [34]. However, the usage pattern in this study showed a constant decline over time with participants using their phone 15.3 times (SD=8.3) during the first week of the intervention, 9.0 (SD=6.5) during the fourth week, but only 4.8 times (SD=4.6) during the eighth week [34]. In a post-study usability survey of MoodRhythm, the average System Usability Scale rating among participants was found to be 85.94 (SD=10.43) out of 100 [58].

Patient compliance was reported as a limitation in several studies [39,41-43,45,49,51,52,54,55,59], as it could not completely be controlled if participants carried their smartphone with them all the time, did not share their phone with others, or did not turn it off occasionally during the monitoring phase. In the study by

Muaremi et al [60], 6 of 12 included participants were excluded for mood state classification analyses due to either not experiencing mood state changes, or data incompleteness caused by not using the study smartphones for phone-calls or switching it off for several days. Maxhuni et al [53] also used information from only 5 of 10 participants due to the lack of data gathered from phone calls. As a precaution, Faurholt-Jepsen et al [26] did not include participants who were not willing to use the study smartphone regularly.

Reporting experiences from a first field trial of their FINE app prototype, Dang et al [39] stated that patients provided information about using the study smartphone only irregularly. However, the patients shared that they would be interested in using the device in addition to their conventional treatment if transparency concerning the recorded data and its further use was given [39]. Participants also criticized not being allowed to disable the background data logging, specifically mentioning GPS tracking [39]. Saunders et al [50] interviewed 21 patients that had participated in a self-monitoring study and found that patients considered it very important that assessment devices look normal to avoid drawing attention in public. Wrist-worn actigraphy devices and a daily mood rating were considered tolerable; a higher frequency of assessments (in the study, up to 10 mood samples during a day were collected) was considered inconvenient [50]. All but one patient was willing to share their data with others, especially since most considered this to be beneficial for the decision making of their psychiatrists; however, patients also stated that they would be interested in giving context explanations to their data to avoid misinterpretations [50]. The size of wrist-worn actigraphy devices was also rated favorably by the participants of the Naslund et al study [37]. Wearing the device was considered uncomplicated, yet some participants with limited prior experience with mobile technology experienced operating with the unfamiliar smartphone technology as distracting and discouraging [37].

Results on Association of Monitored Data With Mood Status

At present, most empirical data exist on the association between objectively recorded data and clinical state or symptom changes (self-reported or rated by clinicians). Therefore, the following section summarizes results obtained with the different (bio-) parameters. If applicable and relevant, additional evidence from nonincluded publications is given to provide a comprehensive overview. Based on our results, no cumulative statements about the direction and/or magnitude of effect sizes can be drawn, since all studies differed concerning selected features, statistical approaches, and handling of ground truth (classification of affective states) for mood state prediction. Overall, the prediction accuracy depended largely on the methodological approach and the respective affective states as outcome measures (eg, depression versus euthymic state compared to depression versus [hypo]mania).

Many of the included studies reported results on a small number of subjects. However, when individual time-series are to be analyzed, a sufficient recording duration for the individual patients seems to be more desirable than a larger number of

subjects. Twenty-five studies monitored data for several months or weeks, three of which still had to exclude subjects from the final analysis, as they did not experience any changes in mood states, and thus associations between mood states and objective parameters could not be analyzed [54,55,60].

Self-Assessment/Self-Report

Self-assessment or subjective rating of mood status via different assessment methods is an integral part of nearly all included studies and systems currently being tested. The superiority of EMAs in mental health versus standard paper-pencil questionnaires has been proved by many studies [61]. Among the selected studies, 20 [26,33-36,38-45,49,50,52,53,55,57,60] provided a self-assessment option. Mood agenda, sleep agenda, and diary have also been included in the PSYCHE platform [46], but no findings were reported in that context.

The content of self-monitored items varies from system to system. Within the MONARCA system, self-assessment of mood, sleep length, medication adherence, irritability, activity, mixed mood, cognitive problems, stress, alcohol consumption, menstruation for women, and individualized early warning signs are requested [62]. In the system used by Grünerbl et al [57], information about the activities of daily living as well as psychological state, physical state, and amount of activity were assessed daily. Within the SIMBA system, there were only two questions about mood and energy [49]. Several systems use established questionnaires. In the MoodRhythm system the Social Rhythm Metric (SRM-5) is used [45]; the FINE app monitors mood through administration of the Patient Health Questionnaire (PHQ-9) [39], and within the MedLink system, mood was assessed with PHQ-9 [33]. The Center for Epidemiologic Studies Depression Scale is administered weekly in the Empath system [38], and in the system used for the AMoSS study *Mood Zoom* was used for daily mood monitoring while the *True Colours* system was selected for weekly mood measurements [50]. In the pilot study of a digital health feedback system using an ingestion sensor for confirmation of medication ingestion and measurement of physical metrics, a self-assessed sleep quality questionnaire was used [36]. Additionally, in some systems there are extra self-assessment opportunities, such as notes or diary sections that are incorporated [45]. However, in the field trial of MONARCA 1.0 [40] it was reported that participants did not use the optional text entry on the phone.

Self-assessments may serve different purposes: they provide a continuous support to individuals (eg, supportive monitoring) and also allow for a better estimation of individuals' need for support, and may serve as an underlying information system to guide and tailor care accordingly (eg, by detecting symptom deterioration or early warning signs for change in moods states) [63]. This longitudinal technology-enhanced assessment of symptoms and behaviors (eg, outcome monitoring and EMA) can help to improve our understanding of illness development and recovery, and can further optimize care by integrating self-monitoring systems into specific e-mental health interventions and apps. In the study of Mobilyze! [34], self-reported items were used for individualizing prediction models. Frost et al [44] showed in the field trial of MONARCA 2.0 that self-reported activity, stress, and sleep were among the

three ranked parameters according to the correlation between Impact Factor and the mood score. Faurholt-Jepsen et al [26] showed that there was a significant correlation between Hamilton Depression Rating Scale, 17 item version (HDRS-17) scores and self-rated mood in adjusted (for age and sex) and unadjusted models. Another study by Faurholt-Jepsen et al [43] showed significant positive correlations between self-monitored mood rate and HDRS-17 in adjusted and unadjusted models (n=29). In relation to manic symptoms, the Young Mania Rating Scale (YMRS) correlated significantly and positively with self-monitored data on mood in adjusted and unadjusted models. Maxhuni et al [53] reported that adding information from questionnaires to the information from accelerometer and audio features improved the overall accuracy results of classifying bipolar disorder episodes; however, they emphasized that precision and recall results were very similar with and without questionnaires.

Physiological Parameters

According to Valenza et al [64], studying mood swings over time shows that changes in mental status have intrinsic dynamics, which suggests a link between autonomous nervous system dynamics (a division of peripheral nervous system) and bipolar disorders. These changes can be shown in physiological parameters such as respiration activity, heart rate variability (HRV), and electrodermal activity. Physiological parameters were assessed in four of the included studies [36,46,48,50]. Since two of the studies did not report any association with the mood states, only the other two will be discussed.

Using the PSYCHE platform, ECG/HRV and respiration activity were gathered using a sensorized t-shirt [46]. Additionally, an accelerometer for measuring physical activity was included in the wearable system. The collected raw data was preprocessed and feature extraction was performed to determine data series for the HRV and respiratory rates. Studies using the PSYCHE platform to identify and predict mood states indicated that HRV series can be employed for classification of mood states: after analyzing HRV series Lanata et al [46] suggested that sample entropy of these series can be regarded as a marker for clinical severity in bipolar disorders. In a study analyzing long-term HRV dynamics in eight patients with bipolar disorder, Gentili et al [48] found that HRV features that were normalized using information from future and past mood states provided a significantly higher classification accuracy (mean=99.52) compared with other HRV-features, which were normalized using other procedures or were not normalized at all. The authors claimed that the results of this study show the high accuracy of HRV features in labeling of mood states [48].

Phone Usage

During a depressive phase, the desire and ability for social interaction is mostly reduced, whereas during the (hypo)manic phase it is increased [59]. The included studies measured social interaction through duration and number of incoming and outgoing phone calls and number of incoming and outgoing short messaging system (SMS) texts per day as a behavioral marker. Nine of the included studies [26,34,42,43,49,52,57,59,60] reported the association between social interactions with mood states.

According to Grunerbl et al [57] the length and number of phone calls increase in mild depressive phases compared to severe depression or a normal mood state. In another study by the same group [59], it was reported that behavioral data only from phone call features did not provide high recognition rate. Muaremi et al [60] examined phone call features in three groups: phone call statistics (eg, number of calls/day), social signal processing (eg, average speaking length), and acoustic emotion recognition (eg, pitch frequency). The average mood state prediction performance of phone call statistics was lower than other features [60].

Within the MONARCA project, significant positive correlations between the number and duration of incoming and outgoing calls per day and scores on YMRS, and between scores on the HDRS-17 and duration of incoming and outgoing calls per day, were found; additionally, a significant positive correlation between scores on the YMRS and the number of outgoing text messages per day was shown [42]. In another study from this group [26] with bipolar patients (n=17) there was no significant correlation between affective symptoms assessed with HDRS-17 and YMRS and the number of outgoing text messages. In another study with 29 participants [43] HDRS-17 scores were found to be significantly and negatively correlated with numbers of incoming text messages per day as well as numbers of outgoing calls per day, whereas positive correlations were shown between numbers of incoming calls and missed calls per day and HDRS-17 scores. In relation to manic symptoms, YMRS scores were found to be significantly and positively correlated with numbers of outgoing text messages per day and daily duration of phone calls in adjusted models [43]. Furthermore, daily duration of outgoing calls, as well as number of characters in incoming text messages, were significantly and negatively correlated with YMRS scores [43].

Beiwinkel et al [49] found a negative association between number of outgoing SMS texts and HDRS-17 scores in within-patient analyses and between-patient analyses. As duration of calls did not show any significant relationships with YMRS and HDRS, the number of calls made was positively associated with manic symptoms, assessed via YMRS in between-patient analyses [49].

Phone usage patterns including screen state and usage of apps might provide information about patients' current mood state. Alvarez-Lorenzo et al [52] reported that self-reported mood level showed a negative correlation with the amount of time the phone screen is on, the percentage of social apps, the percentage of entertainment apps, as well as the amount of time patients interact with the smartphone, respectively (n=18). Conversely, positive correlations between average number of apps and self-reported mood level, as well as between browser apps (eg, Chrome, Firefox) and self-reported mood level, were reported [52]. For 17 bipolar patients, no significant correlation between the amount of time the screen was on and affective symptoms (assessed with HDRS-17 and YMRS) was reported by Faurholt-Jepsen et al [26]. However, in another study by this group [43], duration of screen-on per day was found to be positively and significantly correlated with HDRS-17 scores in unadjusted as well as adjusted models.

Physical Activity

Patients with affective disorder tend to have lower acceleration and activity energy expenditure compared to healthy individuals [65]. Furthermore, a lower activity energy expenditure per day and acceleration have been demonstrated in bipolar patients compared to patients with unipolar depression [65]. By gathering the physical activity data continuously, it is possible to observe the effects of physical activity on mental health more comprehensively. Among the included studies, physical activity was measured mainly with accelerometers incorporated in the smartphones. In the PSYCHE project [46], the accelerometer was integrated into a wearable system.

Eight of the selected studies reported on the relationship between affective states and physical activity [38,45,49,53-55,57,59]. In a study with 12 bipolar patients, the average precision/recall values (for state change detection) were approximately 60% for acceleration [55]. Conversely, in a study with 10 bipolar patients an average recognition accuracy of 70% for movement features (measured with accelerometers) was reported by Grünerbl et al [59]. This finding was supported by results from an older study by this group [57], in which an increase in motion ratio occurred with improved mood state. Furthermore, the authors pointed out that in manic patients the trend was reversed: with the progression from manic to normal state, a decrease of 33.7% in the amount of movement was detected [57]. Abdullah et al [45] pointed out that nonsedentary duration was weakly correlated with mood pattern. Beiwinkel et al [49] reported that device activity (measured by accelerometers) did not have a significant relationship with clinical symptoms in between-patient analyses. Osmani et al [54] reported a correlation between the overall physical activity level and the patients' state. Furthermore, this study found a much stronger correlation between the physical activity divided in daily intervals and patient's state than there was for the overall physical activity levels [54]. Maxhuni et al [53] reported that using accelerometer features recorded during telephone conversations have accuracy results over 80% for classifying bipolar disorder episodes. According to these analyses, the frequency domain feature showed better performance in comparison to the time domain feature [53]. However, Dickerson et al [38] could not find a significant relationship between mood levels and a movement factor.

Location

Changes in travel patterns (eg, distance travelled per day, location changes), which can be gathered through GPS, Bluetooth (via detection of other wireless devices), cell tower identification (ID) and WiFi, may provide information about patients' mood states. For example, it is expected that people in a depressive phase travel and stay outside less often [55]. Eight studies [26,34,43,45,49,55,57,59] investigated the association between travel patterns and mood states. Promising accuracy rates for state recognition by using travel patterns were shown in two studies [55,59]. Grünerbl et al [57] found an increase of 200% in the average time spent outside from a depressive state to normal state. Beiwinkel et al [49] reported that the amount of movement between cell towers (indicating changes of location) did not have a significant relationship with

clinical symptoms but distance travelled showed significant negative association with clinical manic symptoms in between-patient analyses. Additionally, in within-patient analyses a negative relationship between the increase in cell tower movements and clinical manic (as well as depressive) symptoms was found [49]. Conversely, changes in distance travelled were not significantly related to symptom change [49]. Faurholt-Jepsen et al [26] reported a significant correlation between the higher score on HDRS-17 and lowered number of changes in cell tower ID per day in their unadjusted model; however, in their adjusted model the correlation was borderline significant. In another study using MONARCA system [43], Faurholt-Jepsen et al reported significant positive correlations between cell tower ID changes per day and both rating scale scores (YMRS and HDRS-17). In the study by Abdullah et al [45], a significant positive correlation between location clusters and trend of self-assessed energy scores was found, while location clusters had the second highest ranking among distance travelled, nonsedentary duration, and conversation frequency in stable and unstable status classification.

Light Exposure

Light density in the environment can be used to provide information about patients' sleep-wake cycles, which is disturbed in unipolar depression as well as in bipolar disorder [66,67]. Differentiation between natural and artificial light can also be used to approximate time spent outdoors, which has also been shown to have a negative correlation with depressive symptoms [68]. Only three of the included studies [34,45,51] used ambient light sensors. Prociow et al [51] added a light detector as an environmental sensor in their system. The Mobilyze! app also used ambient light data to approximate environmental context, but technical problems were reported (implausible results above the maximum meaningful value were obtained by the ambient light sensor) [34]. The MoodRhythm system included light sensor data in its evaluation of sleep-wake cycles [58], while the study protocol of the MONARCA 2.0 trial categorized ambient light under phone usage information [62]. The monitoring of ambient light is also planned in the further development of the PSYCHE platform [69].

Voice Features

Analysis of speech and voice features appears to be a useful approach in predicting mood state in unipolar depression and in bipolar disorder [70,71]. Furthermore, using microphones to capture ambient sound can allow for the gathering of the number of conversations as a measure of social isolation, which is a behavioral pattern common in depression [45,59].

Eight articles [38,41,45,47,53,56,60,72] reported on relationships between voice features and affective states. The study by Gideon et al [72] dealt with a typical challenge of real-life monitoring of speech patterns: patients use different smartphones causing acoustic variations during the voice recordings. Using various approaches for preprocessing, feature extraction, and data modeling, the authors were able to increase the discriminative power of their mood state prediction algorithms [72]. However, only structured interview calls were analyzed in this study. Karam et al [56] also demonstrated that mood states of bipolar individuals can be separated based on

voice recordings from clinical interactions and that hypomania can also be detected from unstructured calls, whereas their system struggled to detect depression from voice recordings during nonclinical interactions. Abdullah et al [45] reported that conversation frequency was weakly correlated with mood pattern. Within PSYCHE project, Guidi et al [47] concluded from a case report that changes in voice pitch features can be extracted via smartphones and correlate significantly with changes between euthymic, depressed, and hypomanic states in a bipolar patient.

Analyses of voice and speech features are also performed within the frame of the MONARCA project. Muaremi et al [60] extracted three types of features from phone call conversations and reported that using acoustic features (eg, kurtosis energy, pitch frequency) resulted in the best state recognition in comparison to phone call statistics (eg, number of phone calls during the day, average duration of phone calls) and social signals (eg, average speaking turn duration, average number of speaker turns). Maxhuni et al [53] reported that using audio features (emotional and spectral) resulted in accuracy results >80% for classifying the mood state of patients with bipolar affective disorder. Testing emotional and spectral features isolated and together resulted in similar accuracy [53]. Faurholt-Jepsen et al [41] reported an accuracy ranging from 0.61 to 0.74 in the classification of mood states based on voice feature. Dickerson et al [38] found speech factors and sleep factors more indicative of depression than other features measured in this case study.

Aggregation of the Data Collected From Smartphone Sensors

Most of the reviewed papers described a combination of self-reported data with behavioral and/or physiological parameters. Grünerbl et al [55,59] reported that the combination of features (eg, social behavior, travel patterns, and movement) optimized the state detection and provided better results than single sensors or fusion of only a few sensor modalities. Abdullah et al [45] also reported that combining self-reported data with data from several smartphone sensors (light sensor, accelerometer, microphone) and communication patterns resulted in reliable prediction of SRM-5 scores (75.89% of correctly classified labels had a probability >.7). Ranking the used features according to their importance for the classification of stable and unstable status (according to SRM-5 scores) showed the following order: distance travelled, location cluster, nonsedentary duration, and finally conversation frequency. Maxhuni et al [53] found only a small improvement in average classification accuracy for emotional states of their participants when audio features derived from phone conversations were combined with accelerometer features recorded during phone conversations. Faurholt-Jepsen et al [41] stated that increases in the sensitivity, specificity, and accuracy were found after combining voice features, self-monitored data, and automatically generated objective data. The Empath system described by Dickerson et al [38] also aims at combining several behavioral markers into a depression index, or a global depression risk index. However, only a single patient case study was reported, mainly dealing with feasibility issues [38]. Prociow et al [51] also introduced a system that combines environmental (motion

detectors, light detector, microphone, door and bed sensors, remote control monitor) and wearable sensors (accelerometer, light detector, microphone, GPS).

Results on Effects on Clinical Outcome

Most of the reviewed studies did not report any results regarding the effects of smartphone-based monitoring on depressive symptoms or clinical outcomes. However, Burns et al [34] reported that during the Mobilyze! intervention program, depressive symptoms (self-reported on the PHQ-9) decreased significantly over time ($t_{13}=7.02$, $beta_{week}=−.82$, $P<.001$) according to intent-to-treat analyses. Moreover, this study found that evaluator-rated depressive symptoms (rated on the Quick Inventory of Depressive Symptomatology) improved ($t_{13}=8.22$, $beta_{week}=−.81$, $P<.001$) and it became less likely for participants to fulfill diagnostic criteria for Major Depressive Disorder ($Z=2.15$, $beta_{week}=−.65$, $P=.03$) [34]. Additionally, anxiety symptoms decreased ($t_{13}=4.59$, $beta_{week}=−.71$, $P<.001$) according to Generalized Anxiety Disorder questionnaire scores [34].

Saunders et al [50] performed qualitative interviews with a subset of participants (n=21) of the AMoSS study, which had monitored their moods using a smartphone combined with movement sensing devices. The authors reported that most patients (n=16) experienced the self-monitoring as enhancing their illness insight and that half of their patients stated that they had improved their mood due to being able to better identify their feelings [50]. Half of the patients also reported a change in behavior, in most cases related to increased exercise levels, but four patients also pointed to potential negative effects due to becoming too occupied with the monitoring [50]. Naslund et al [37] provided 11 patients with activity tracking devices combined with smartphones for a 6-month lifestyle intervention aimed at weight reduction, and then performed qualitative interviews. Participants considered the devices to be motivating and to increase awareness about their level of physical activity, thus helping them to increase their overall activity [37]. In the pilot trial of MedLink system [33] which assessed medication adherence through a *Wisepill bottle*, a decrease of depressive symptoms assessed with PHQ-9 was observed ($P=.008$). The authors suggested that lag time between monitoring and the initiation of medication can be reduced via the MedLink system [33].

Results from only one RCT were available. Faurholt-Jepsen et al [35] reported that in the MONARCA 1.0 trial patients in the intervention group showed more depressive symptoms during the trial according to HDRS-17 scores compared with the control group (B=2.02, 95% CI −.13 to 4.17, $P=.07$). Additionally, in further analyses (excluding mixed depressive and manic symptoms in both adjusted and unadjusted models) significantly more depressive symptoms were reported in the intervention group [35]. No difference in manic symptoms (based on the YMRS) was observed between the intervention and control groups (B=−.34, 95% CI −1.14 to .47, $P=.41$). However, excluding mixed depressive and manic symptoms, a trend for fewer manic symptoms in the intervention group was reported compared to the control group in the adjusted model (B=−1.07, 95% CI −2.15 to 0.005, $P=.051$) [35]. Furthermore, a trend towards a higher score of subjectively perceived stress was

reported in the intervention group compared to the participants in the control group (adjusted model: B 2.40, 95% CI -.33 to 5.13, $P=.09$). No differences in psychosocial functioning, quality of life, self-rated depressive or manic symptoms, or self-rated or screened cognitive function were found between the intervention and control groups [35].

Discussion

To our knowledge, this is the first systematic review of the current scientific literature on smartphone-based systems for the subjective and objective monitoring of mood disorders and studies testing these systems. A considerable number of projects have started developing smartphone-based tools for the management and monitoring of mood disorders and have conducted promising tests of feasibility. Nonetheless, it became obvious during the review process that only a small number of systems that are currently available (or in preparation) have been described in scientific papers, let alone been subjected to empirical studies.

The results showed great potential in monitoring parameters relevant to the course of affective disorders, although most of the reviewed studies were feasibility tests or pilot trials without control groups, conducted with only a small number of participants and with short assessment periods, and did not investigating any medium- or long-term effects or possible risks of the respective applications. It is important to highlight that there are several systems described in publications that could not be included in this review due to: not meeting the inclusion criteria (eg, no monitoring of objective data [73,74]), not involving a smartphone for the objective monitoring [64,69,75-77], not having studied participants with an affective disorder diagnosis [78-81], or not having published results from trials being conducted at the time of our search [82].

Given the lack of reliable biomarkers in affective disorders, the combination of different physiological and behavioral parameters seems to be a highly promising approach [83,84]. This is especially true if time-series analyses are applied to continuously monitored behavioral and/or physiological data from individual patients to predict changes in their clinical states. It can be noted that significant heterogeneity exists with respect to how the data was analyzed across studies. Many groups developed their own data analytic approaches (eg, machine learning algorithms) in parallel to the set-up of their systems, and the chosen approaches have considerable impacts on the results of the respective studies. Those studies using such approaches could show good prediction values and probability estimations for changes in mood states or clinical outcomes [45,55,59].

The acceptance of using smartphones for continuous monitoring (of subjective and objective data) seems encouraging based on the results of this review and reflects patients' interests. However, regarding the decrease in adherence rates in several studies [34,78], some caution might be taken to ensure the continuity of usage. For instance, a promising concept for user engagement is the employment of gamification (game design elements such as infographics, feedback, activity, and progress bars [85]). Future trials should use reminders and/or other

evidence-based patient engagement methods (eg, financial incentives, gamification, and peer mentoring) to minimize the problem of missing data that stems from nonadherence of carrying or using the smartphone for data collection.

Technical issues that might occur in this context (ie, energy management) could serve as a major obstacle. Using external sensors that do not rely on the smartphone's power source was discussed as one possible solution, but Prociow et al [51] reported that any additional devices were perceived as uncomfortable by the patients and tends to increase reluctance to use the system. It is also important to enable data collection from as many operating system as possible, since using another smartphone system was reported to be a reason to decline the participation of a study [41,43].

Only five of the reviewed articles reported on clinical outcomes [33-35,37,50]. In the pilot study of the Mobilyze! intervention program [34] the depressive symptoms were decreased during the trial. Conversely in the first RCT of the MONARCA system [35] the exploratory analyses in relation to primary outcomes revealed that electronic daily self-monitoring (including feedback to the users as well as their health care practitioners) may improve manic symptoms but also sustain depressive symptoms among some, but not all, patients. Following the authors' explanation, it can be speculated that self-monitoring increases illness insight and self-awareness, which are often lacking in phases of (hypo)mania. Therefore, patients within this mood state might benefit more than depressed patients, where a daily confrontation with their perceived shortcomings might sustain or even worsen their depression. However, the potential risks and adverse effects associated with continuous self-monitoring should be considered and carefully evaluated. Surprisingly, these aspects are highly underrepresented or even neglected in the recent literature body. Conceivably, a fragmentation of daily life, increased somatization due to continuous monitoring of biological parameters, negative effects on attention, decreased mindfulness, and social and behavioral aspects could be potential negative effects [86,87]. These considerations indicate the importance of conducting more experimental and randomized trials with larger samples sizes and longer assessment durations.

Combining smartphone-based self-monitoring with digital intervention strategies has a high potential for improving the current treatment and management of affective disorders. By investigating the efficacy of smartphone-based interventions, the notion of the "digital placebo" effect [88] should also be taken into account. High levels of expectation, trust, and personalization of smartphones influence the outcome of app-based interventions, as simply interacting with and receiving feedback from a smartphone app may already increase subjective well-being [88,89]. For example, one of the first studies on mental health apps by Kauer et al [90] showed that depressive symptom recording and self-monitoring via smartphone significantly reduced symptoms even without any direct therapeutic intervention. Not only the placebo effects, but also the adverse effects and safety issues of using digital mental health apps, are very rarely taken into account. A number of adverse effects that are unique to the use of mobile apps and the Internet have been discussed (eg, increased levels of

inactivity, sleep pattern disturbances, eye strain, and reduced social skills [87]), but to date there is no information available on occurrence or frequency of risks, or whether certain mental health apps have a higher or lower risk potential than others.

Additionally, there are several ethical issues arising from continuous monitoring and the use of objective smartphone data. While some users/patients may find that remote monitoring allows them to feel more control over their illness, also opening new possibilities for patient-empowerment and self-management, others may feel overwhelmed by the added responsibility of self-monitoring, or perceive it as a constant reminder that they are ill [91]. The feeling of being watched or being under surveillance may also cause unease or even increase anxiety. Furthermore, there is a certain risk for patients and health care providers alike of placing too much trust in smartphone-based monitoring systems, especially concerning the (limited) possibilities of an adequate response in emergency situations (eg, acute suicidal ideation). The increasing usage and application of these technologies in mental health care and research poses the question of responsibility, and new ethical dilemmas will undoubtedly arise [92].

Sensitive data concerning mental health require an especially high standard of protection from security breaches due to a high risk of stigmatization and discrimination in the case of data disclosure. Patients' concerns about the security of the data monitored and recorded by the system have significant impacts on the acceptance of the apps and the readiness to use them for continuous monitoring [93]. Processing data locally on the smartphone has been suggested as one possible solution [47]. The classification of mental health apps as medical products is also of capital importance for patient safety and acceptance. When a product is considered a medical product, the respective mental-health apps (especially apps focusing on specific disorders and symptom reduction) would need to meet certain quality standards, and would have to undergo efficacy and safety trials (comparable to the approval process of drugs or physical therapies). For example, this would ensure that patients would not falsely rely on a completely ineffective app or even use an app giving them potentially harmful recommendations [89]. However, not all of the available mental health apps (especially those focusing on self-management and lifestyle) are considered medical products, and therefore do not require these defined standards. To summarize, further attention should be given to questions of data security and confidentiality, especially regarding the huge amounts of mental health apps being openly available (eg, Apple App Store, Google Play Store) that show major flaws in security settings [94].

Future Directions/Perspectives

Smartphone-based systems for managing and monitoring mood disorders present a highly promising field of innovation in health care. Simply using and carrying a smartphone on a day-to-day basis generates a larger amount of data than is typically collected in questionnaire-based studies or online interventions. Analyses of covariations of recorded bio-parameters within individual time-series and identification of individual patterns that predict state changes must be employed, in addition to group statistics. However, to identify such individual patterns, long monitoring

durations with a certain amount of state changes are required. Most of the available studies used short or moderately sized recording durations and three of them [54,55,60] had to exclude participants from the final analyses, as they did not exhibit recognizable changes in mood state.

Apart from the continuous assessment of bio-parameters themselves, smartphone-based monitoring also allows researchers to gather information on context and environment, which could prove valuable for the interpretation of the monitored biomedical data (eg, information about weather conditions) and allow for a better interpretation of changes in body temperature or skin conductance levels (as well as locomotor activity) that might be restricted on rainy versus sunny days. Furthermore, these large amounts of data allow for the employment of data mining techniques to potentially generate and explore new hypotheses about causes and predictors of affective episodes, define individual disorder phenotypes, or analyze the consequences of changes in medication or treatment regimens, or potentially even health policy changes [94].

Continuous measurement of physiological and behavioral parameters is a valuable addition to self-reported symptoms to help complete the clinical picture, especially since there is often significant variation between measured parameters and subjective perspective of the self-reporting patient [95]. When considering disorders with impaired insight into illness, such as bipolar disorder [96], objectively measured data may provide a valuable method of tracking and predicting mood state changes. A highly valuable aspect of continuous physiological and behavioral monitoring approaches is the versatility of the usage of the resulting data: it could be used by the patients themselves to enhance self-monitoring practices and improve self-management strategies, and help clinicians adjust treatment by providing a more detailed picture of symptom trajectory. However, as of now, there is only sparse empirical evidence backing up these expectations, as most systems are currently in development and/or have only been studied within feasibility or pilot studies.

Smartphone-based monitoring systems can also be combined with electronically delivered psychotherapeutic interventions. By linking behavioral monitoring and intervention feedback loops, interventions can be delivered right-on-time at the most useful moment, in contrast to face-to-face psychotherapy mainly based on retrospective data and information [23]. However, fully relying on mental-health apps for disorder management and therapy would be placing false confidence and trust in a young technology. A broader empirical database is needed regarding effectiveness and potential adverse effects of continuous monitoring of behavioral and physiological data using smartphones. Meanwhile, an integrated treatment approach (face-to-face interaction between patients and clinicians and smartphone-based disorder management) should be preferred.

Limitations

In this review, we have presented the current state of research on smartphone-based mental-health technologies for monitoring and managing mood disorders. A limitation of this publication is that it only covers those smartphone-based systems in which studies have been conducted, and results of these studies have been published in scientific literature. We also focused on articles listed in scientific literature databases; thus, grey literature (eg, conference proceedings and abstracts) was not included. There are a great number of apps for the monitoring and management of mood disorders available to the public that have either not been subjected to a scientific study, or the results have not been published (yet). Furthermore, some publications had to be excluded due to not reporting results on patients with affective disorders, although they were also introducing smartphone-based systems for affective symptoms. Another limitation is that a number of the included studies were feasibility studies, allowing conclusions to be drawn about the practicality of a certain monitoring technology, yet not about its clinical efficacy or effectiveness. However, the results of this systematic review on the current state of smartphone-based monitoring of depressive symptoms demonstrates that using smartphone apps for EMA and management of mood disorders is a promising field, in which more and larger efficacy studies and longitudinal research regarding placebo effects, possible risks, and adverse effects are required.

Conclusion

Smartphone-based monitoring of objective and subjective data in mood disorders is a rapidly growing approach and research field. The current body of literature consists mainly of observational studies and substantiates the value of smartphone-based approaches for gathering long-term objective data (aside from self-ratings) to monitor clinical symptoms, predict changes in clinical states, and investigate causal inferences about state changes in patients with affective disorders. Although promising, a much larger evidence-base is necessary to fully assess the potential, as well as the risks, of these approaches. Fortunately, a number of study protocols stated ambitions to expand and intensify research in this emerging and promising field in the coming years.

Authors' Contributions

Authors UH, ED, and EK developed the concept of this paper and the search strategy. Authors ED, EK, and CS completed the data extraction of relevant articles. Authors ED, XW, CS, and EK wrote the first draft of the manuscript. After the critical revision

of the manuscript by UH, the final draft of the manuscript was written by the authors, and all authors gave final approval of the submitted version of the manuscript.

Conflicts of Interest

The author UH declares several potential conflicts of interest with respect to the research, authorship, and/or publication of this article. Within the last three years, UH was: an advisory board member for Lundbeck, Takeda Pharmaceuticals, Servier, and Otsuka Pharma; a consultant for Bayer Pharma; and a speaker for Medice Arzneimittel, Novartis, and Roche Pharma. ED, XW, CS and EK declare no competing interests.

Multimedia Appendix 1

PRISMA checklist.

[PDF File (Adobe PDF File), 83KB-Multimedia Appendix 1]

Multimedia Appendix 2

Search strategy.

[PDF File (Adobe PDF File), 64KB-Multimedia Appendix 2]

Multimedia Appendix 3

Reference list of excluded full text articles with reason for exclusion.

[PDF File (Adobe PDF File), 260KB-Multimedia Appendix 3]

Multimedia Appendix 4

Overview of selected studies.

[PDF File (Adobe PDF File), 203KB-Multimedia Appendix 4]

Multimedia Appendix 5

Methodological Quality Scores.

[PDF File (Adobe PDF File), 168KB-Multimedia Appendix 5]

References

1. World Health Organization. The European Health Report 2009, Health and health systems. Copenhagen: WHO Regional Office for Europe; 2009. URL: http://www.euro.who.int/__data/assets/pdf_file/0009/82386/E93103.pdf [accessed 2017-06-30] [WebCite Cache ID 6rbLntH0B]
2. Eaton WW, Shao H, Nestadt G, Lee HB, Lee BH, Bienvenu OJ, et al. Population-based study of first onset and chronicity in major depressive disorder. Arch Gen Psychiatry 2008 May;65(5):513-520 [FREE Full text] [doi: 10.1001/archpsyc.65.5.513] [Medline: 18458203]
3. Baumeister H, Härter M. Prevalence of mental disorders based on general population surveys. Soc Psychiatry Psychiatr Epidemiol 2007 Jul;42(7):537-546. [doi: 10.1007/s00127-007-0204-1] [Medline: 17516013]
4. Thomson W. Lifting the shroud on depression and premature mortality: a 49-year follow-up study. J Affect Disord 2011 Apr;130(1-2):60-65. [doi: 10.1016/j.jad.2010.09.028] [Medline: 20952074]
5. Ustün TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murray CJ. Global burden of depressive disorders in the year 2000. Br J Psychiatry 2004 May;184:386-392 [FREE Full text] [Medline: 15123501]
6. Chisholm D, Sweeny K, Sheehan P, Rasmussen B, Smit F, Cuijpers P, et al. Scaling-up treatment of depression and anxiety: a global return on investment analysis. Lancet Psychiatry 2016 May;3(5):415-424 [FREE Full text] [doi: 10.1016/S2215-0366(16)30024-4] [Medline: 27083119]
7. Hardeveld F, Spijker J, De GR, Nolen WA, Beekman AT. Recurrence of major depressive disorder and its predictors in the general population: results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). Psychol Med 2013 Jan;43(1):39-48. [doi: 10.1017/S0033291712002395] [Medline: 23111147]
8. Klein DN, Kotov R. Course of depression in a 10-year prospective study: evidence for qualitatively distinct subgroups. J Abnorm Psychol 2016 Apr;125(3):337-348 [FREE Full text] [doi: 10.1037/abn0000147] [Medline: 26845258]
9. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):593-602. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]

XSL•FO

RenderX

10. Merikangas KR, Jin R, He J, Kessler RC, Lee S, Sampson NA, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. Arch Gen Psychiatry 2011 Mar;68(3):241-251 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2011.12] [Medline: 21383262]

11. Grande I, Berk M, Birmaher B, Vieta E. Bipolar disorder. Lancet 2016 Apr 09;387(10027):1561-1572. [doi: 10.1016/S0140-6736(15)00241-X] [Medline: 26388529]

12. Bonnín CM, Torrent C, Goikolea JM, Reinares M, Solé B, Valentí M, et al. The impact of repeated manic episodes and executive dysfunction on work adjustment in bipolar disorder. Eur Arch Psychiatry Clin Neurosci 2014 Apr;264(3):247-254. [doi: 10.1007/s00406-013-0431-2] [Medline: 23912643]

13. Malhi GS, Bassett D, Boyce P, Bryant R, Fitzgerald PB, Fritz K, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. Aust N Z J Psychiatry 2015 Dec;49(12):1087-1206. [doi: 10.1177/0004867415617657] [Medline: 26643054]

14. Fernández A, Haro JM, Martinez-Alonso M, Demyttenaere K, Brugha TS, Autonell J, et al. Treatment adequacy for anxiety and depressive disorders in six European countries. Br J Psychiatry 2007 Feb;190:172-173 [FREE Full text] [doi: 10.1192/bjp.bp.106.023507] [Medline: 17267936]

15. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. Bull World Health Organ 2004 Nov;82(11):858-866 [FREE Full text] [Medline: 15640922]

16. Trivedi MH, Lin EH, Katon WJ. Consensus recommendations for improving adherence, self-management, and outcomes in patients with depression. CNS Spectr 2007 Aug;12(8 Suppl 13):1-27. [Medline: 17986951]

17. Vuorilehto MS, Melartin TK, Riihimäki K, Isometsä ET. Pharmacological and psychosocial treatment of depression in primary care: low intensity and poor adherence and continuity. J Affect Disord 2016 Sep 15;202:145-152. [doi: 10.1016/j.jad.2016.05.035] [Medline: 27262636]

18. Clement S, Schauman O, Graham T, Maggioni F, Evans-Lacko S, Bezborodovs N, et al. What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. Psychol Med 2015 Jan;45(1):11-27. [doi: 10.1017/S0033291714000129] [Medline: 24569086]

19. Mohr DC, Ho J, Duffecy J, Baron KG, Lehman KA, Jin L, et al. Perceived barriers to psychological treatments and their relationship to depression. J Clin Psychol 2010 Apr;66(4):394-409 [FREE Full text] [doi: 10.1002/jclp.20659] [Medline: 20127795]

20. Arensman E, Koburger N, Larkin C, Karwig G, Coffey C, Maxwell M, et al. Depression awareness and self-management through the Internet: protocol for an internationally standardized approach. JMIR Res Protoc 2015 Aug 06;4(3):e99 [FREE Full text] [doi: 10.2196/resprot.4358] [Medline: 26251104]

21. Vallury KD, Jones M, Oosterbroek C. Computerized cognitive behavior therapy for anxiety and depression in rural areas: a systematic review. J Med Internet Res 2015;17(6):e139 [FREE Full text] [doi: 10.2196/jmir.4145] [Medline: 26048193]

22. Høifødt RS, Lillevoll KR, Griffiths KM, Wilsgaard T, Eisemann M, Waterloo K, et al. The clinical effectiveness of web-based cognitive behavioral therapy with face-to-face therapist support for depressed primary care patients: randomized controlled trial. J Med Internet Res 2013;15(8):e153 [FREE Full text] [doi: 10.2196/jmir.2714] [Medline: 23916965]

23. Hidalgo-Mazzei D, Mateu A, Reinares M, Matic A, Vieta E, Colom F. Internet-based psychological interventions for bipolar disorder: Review of the present and insights into the future. J Affect Disord 2015 Aug 28;188:1-13. [doi: 10.1016/j.jad.2015.08.005] [Medline: 26342885]

24. Trefflich F, Kalckreuth S, Mergl R, Rummel-Kluge C. Psychiatric patients' internet use corresponds to the Internet use of the general public. Psychiatry Res 2015 Mar 30;226(1):136-141. [doi: 10.1016/j.psychres.2014.12.037] [Medline: 25623020]

25. Ericsson Mobility Report.: Rima Qureshi; 2016. On the pulse of the networked society URL: https://www.ericsson.com/res/docs/2016/ericsson-mobility-report-2016.pdf [accessed 2017-06-30] [WebCite Cache ID 6rbZt17c8]

26. Faurholt-Jepsen M, Frost M, Vinberg M, Christensen EM, Bardram JE, Kessing LV. Smartphone data as objective measures of bipolar disorder symptoms. Psychiatry Res 2014 Jun 30;217(1-2):124-127. [doi: 10.1016/j.psychres.2014.03.009] [Medline: 24679993]

27. Tams S, Hill K, Ortiz de Guinea A, Thatcher J, Grover V. NeuroIS-alternative or complement to existing methods? Illustrating the holistic effects of neuroscience and self-reported data in the context of technostress research. J Assoc Inf Syst 2014;15(10):e1 [FREE Full text]

28. Dey A, Wac K, Ferreira D, Tassini K, Hong J, Ramos J. Getting closer: an empirical investigation of the proximity of user to their smart phones. In: Proceedings of the 13th International Conference on Ubiquitous Computing. USA: ACM; 2011 Presented at: 13th International Conference on Ubiquitous Computing; 2011; New York p. A.

29. Gravenhorst F, Muaremi A, Bardram J, Grünerbl A, Mayora O, Wurzer G, et al. Mobile phones as medical devices in mental disorder treatment: an overview. Pers Ubiquit Comput 2014 Sep 21;19(2):335-353. [doi: 10.1007/s00779-014-0829-5]

30. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009 Oct;62(10):1006-1012. [doi: 10.1016/j.jclinepi.2009.06.005] [Medline: 19631508]

31. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health 1998 Jun;52(6):377-384 [FREE Full text] [Medline: 9764259]

XSL•FO
RenderX

32. Hootman JM, Driban JB, Sitler MR, Harris KP, Cattano NM. Reliability and validity of three quality rating instruments for systematic reviews of observational studies. Res Synth Methods 2011 Jun;2(2):110-118. [doi: 10.1002/jrsm.41] [Medline: 26061679]

33. Mohr DC, Stiles-Shields C, Brenner C, Palac H, Montague E, Kaiser SM, et al. MedLink: a mobile intervention to address failure points in the treatment of depression in general medicine. Int Conf Pervasive Comput Technol Healthc 2015 May;2015:100-107 [FREE Full text] [Medline: 26640740]

34. Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, et al. Harnessing context sensing to develop a mobile intervention for depression. J Med Internet Res 2011;13(3):e55 [FREE Full text] [doi: 10.2196/jmir.1838] [Medline: 21840837]

35. Faurholt-Jepsen M, Frost M, Ritz C, Christensen EM, Jacoby AS, Mikkelsen RL, et al. Daily electronic self-monitoring in bipolar disorder using smartphones - the MONARCA I trial: a randomized, placebo-controlled, single-blind, parallel group trial. Psychol Med 2015 Oct;45(13):2691-2704. [doi: 10.1017/S0033291715000410] [Medline: 26220802]

36. Kane JM, Perlis RH, DiCarlo LA, Au-Yeung K, Duong J, Petrides G. First experience with a wireless system incorporating physiologic assessments and direct confirmation of digital tablet ingestions in ambulatory patients with schizophrenia or bipolar disorder. J Clin Psychiatry 2013 Jun;74(6):e533-e540. [doi: 10.4088/JCP.12m08222] [Medline: 23842023]

37. Naslund JA, Aschbrenner KA, Bartels SJ. Wearable devices and smartphones for activity tracking among people with serious mental illness. Ment Health Phys Act 2016 Mar;10:10-17 [FREE Full text] [doi: 10.1016/j.mhpa.2016.02.001] [Medline: 27134654]

38. Dickerson R, Gorlin E, Stankovic J. Empath: A continuous remote emotional health monitoring system for depressive illness. In: Proceedings of the 2nd Conference on Wireless Health 2011. USA: ACM; 2011 Presented at: 2nd Conference on Wireless Health; 2011; New York p. A.

39. Dang M, Mielke C, Diehl A, Haux R. Accompanying depression with FINE - a smartphone-based approach. Stud Health Technol Inform 2016;228:195-199. [Medline: 27577370]

40. Bardram J, Frost M, Szántó K, Faurholt-Jepsen M, Vinberg M, Kessing L. 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. USA: ACM; 2013 Presented at: SIGCHI Conference on Human Factors in Computing Systems; 2013; New York p. A.

41. Faurholt-Jepsen M, Busk J, Frost M, Vinberg M, Christensen EM, Winther O, et al. Voice analysis as an objective state marker in bipolar disorder. Transl Psychiatry 2016 Jul 19;6:e856. [doi: 10.1038/tp.2016.123] [Medline: 27434490]

42. Faurholt-Jepsen M, Vinberg M, Frost M, Christensen EM, Bardram JE, Kessing LV. Smartphone data as an electronic biomarker of illness activity in bipolar disorder. Bipolar Disord 2015 Nov;17(7):715-728. [doi: 10.1111/bdi.12332] [Medline: 26395972]

43. Faurholt-Jepsen M, Vinberg M, Frost M, Debel S, Margrethe CE, Bardram JE, et al. Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder. Int J Methods Psychiatr Res 2016 Dec;25(4):309-323. [doi: 10.1002/mpr.1502] [Medline: 27038019]

44. Frost M, Doryab A, Faurholt-Jepsen M, Kessing L, Bardram J. Supporting disease insight through data analysis: refinements of the MONARCA self-assessment system. In: Proceedings of the 2013 ACM international joint conference on Pervasiveubiquitous computing. USA: ACM; 2013 Presented at: ACM international joint conference on Pervasive and ubiquitous computing; 2013; New York p. A-142.

45. Abdullah S, Matthews M, Frank E, Doherty G, Gay G, Choudhury T. Automatic detection of social rhythms in bipolar disorder. J Am Med Inform Assoc 2016 May;23(3):538-543. [doi: 10.1093/jamia/ocv200] [Medline: 26977102]

46. Lanata A, Valenza G, Nardelli M, Gentili C, Scilingo EP. Complexity index from a personalized wearable monitoring system for assessing remission in mental health. IEEE J Biomed Health Inform 2015 Jan;19(1):132-139. [doi: 10.1109/JBHI.2014.2360711] [Medline: 25291802]

47. Guidi A, Salvi S, Ottaviano M, Gentili C, Bertschy G, de RD, et al. Smartphone application for the analysis of prosodic features in running speech with a focus on bipolar disorders: system performance evaluation and case study. Sensors (Basel) 2015 Nov 06;15(11):28070-28087 [FREE Full text] [doi: 10.3390/s151128070] [Medline: 26561811]

48. Gentili C, Valenza G, Nardelli M, Lanatà A, Bertschy G, Weiner L, et al. Longitudinal monitoring of heartbeat dynamics predicts mood changes in bipolar patients: a pilot study. J Affect Disord 2017 Feb;209:30-38. [doi: 10.1016/j.jad.2016.11.008] [Medline: 27870943]

49. Beiwinkel T, Kindermann S, Maier A, Kerl C, Moock J, Barbian G, et al. Using smartphones to monitor bipolar disorder symptoms: a pilot study. JMIR Ment Health 2016;3(1):e2 [FREE Full text] [doi: 10.2196/mental.4560] [Medline: 26740354]

50. Saunders KE, Bilderbeck AC, Panchal P, Atkinson LZ, Geddes JR, Goodwin GM. Experiences of remote mood and activity monitoring in bipolar disorder: a qualitative study. Eur Psychiatry 2017 Mar;41:115-121. [doi: 10.1016/j.eurpsy.2016.11.005] [Medline: 28135594]

51. Prociow P, Wac K, Crowe J. Mobile psychiatry: towards improving the care for bipolar disorder. Int J Ment Health Syst 2012 May 29;6(1):5 [FREE Full text] [doi: 10.1186/1752-4458-6-5] [Medline: 22643029]

52. Alvarez-Lozano J, Osmani V, Mayora O, Frost M, Bardram J, Faurholt-Jepsen M, et al. Tell me your apps and I will tell you your mood: correlation of apps usage with bipolar disorder state. In: Proceedings of the 7th International Conference

on Pervasive Technologies Related to Assistive Environments. USA: ACM; 2014 Presented at: 7th International Conference on Pervasive Technologies and Related to Assistive Environments; 2014; New York p. A.

53. Maxhuni A, Muñoz-Meléndez A, Osmani V, Perez H, Mayora O, Morales EF. Classification of bipolar disorder episodes based on analysis of voice and motor activity of patients. Pervasive Mob Comput 2016 Sep;31:50-66. [doi: 10.1016/j.pmcj.2016.01.008]

54. Osmani V, Maxhuni A, Grünerbl A, Lukowicz P, Haring C, Mayora O. Monitoring activity of patients with bipolar disorder using smart phones. In: Proceedings of International Conference on Advances in Mobile Computing & Multimedia. USA: ACM; 2013 Presented at: Proceedings of International Conference on Advances in Mobile Computing & Multimedia; 2013; New York p. A.

55. Gruenerbl A, Osmani V, Bahle G, Carrasco J, Oehler S, Mayora O, et al. Using smart phone mobility traces for the diagnosis of depressive and manic episodes in bipolar patients. In: Proceedings of the 5th Augmented Human International Conference. USA: ACM; 2014 Presented at: 5th Augmented Human International Conference; 2014; New York p. A.

56. Karam ZN, Provost EM, Singh S, Montgomery J, Archer C, Harrington G, et al. Ecologicaly valid long-term mood monitoring of individuals with bipolar disorder using speech. Proc IEEE Int Conf Acoust Speech Signal Process 2014 May;2014:4858-4862 [FREE Full text] [doi: 10.1109/ICASSP.2014.6854525] [Medline: 27630535]

57. Grünerbl A, Oleksy P, Bahle G, Haring C, Weppner J, Lukowicz P. Towards smart phone based monitoring of bipolar disorder. In: Proceedings of the Second ACM Workshop on Mobile Systems, Applications, and Services for HealthCare. USA: ACM; 2012 Presented at: Second ACM Workshop on Mobile Systems, Applications, and Services for HealthCare; 2012; New York p. A.

58. Matthews M, Abdullah S, Murnane E, Voida S, Choudhury T, Gay G, et al. Development and evaluation of a smartphone-based measure of social rhythms for bipolar disorder. Assessment 2016 Jun 29:472-483. [doi: 10.1177/1073191116656794] [Medline: 27358214]

59. Grünerbl A, Muaremi A, Osmani V, Bahle G, Ohler S, Tröster G, et al. Smartphone-based recognition of states and state changes in bipolar disorder patients. IEEE J Biomed Health Inform 2015 Jan;19(1):140-148. [doi: 10.1109/JBHI.2014.2343154] [Medline: 25073181]

60. Muaremi A, Gravenhorst F, Grünerbl A, Arnrich B, Tröster G. Assessing bipolar episodes using speech cues derived from phone calls. In: Pervasive Computing Paradigms for Mental Health. Cham: Springer; 2014 Presented at: 4th International Symposium, MindCare; 2014; Tokyo p. 978-973. [doi: 10.1007/978-3-319-11564-1]

61. Moore RC, Depp CA, Wetherell JL, Lenze EJ. Ecological momentary assessment versus standard assessment instruments for measuring mindfulness, depressed mood, and anxiety among older adults. J Psychiatr Res 2016 Apr;75:116-123. [doi: 10.1016/j.jpsychires.2016.01.011] [Medline: 26851494]

62. Faurholt-Jepsen M, Vinberg M, Frost M, Christensen EM, Bardram J, Kessing LV. 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 2014;14:309 [FREE Full text] [doi: 10.1186/s12888-014-0309-5] [Medline: 25420431]

63. Bauer S, Moessner M. Technology-enhanced monitoring in psychotherapy and e-mental health. J Ment Health 2012 Aug;21(4):355-363. [doi: 10.3109/09638237.2012.667886] [Medline: 22548363]

64. Valenza G, Lanatà A, Paradiso R, Scilingo EP. Advanced technology meets mental health: how smartphones, textile electronics, and signal processing can serve mental health monitoring, diagnosis, and treatment. IEEE Pulse 2014;5(3):56-59. [doi: 10.1109/MPUL.2014.2309582] [Medline: 24838213]

65. Faurholt-Jepsen M, Brage S, Vinberg M, Christensen EM, Knorr U, Jensen HM, et al. Differences in psychomotor activity in patients suffering from unipolar and bipolar affective disorder in the remitted or mild/moderate depressive state. J Affect Disord 2012 Dec 10;141(2-3):457-463. [doi: 10.1016/j.jad.2012.02.020] [Medline: 22391514]

66. Luik AI, Zuurbier LA, Direk N, Hofman A, Van Someren EJ, Tiemeier H. 24-hour activity rhythm and sleep disturbances in depression and anxiety: a population-based study of middle-aged and older persons. Depress Anxiety 2015 Sep;32(9):684-692. [doi: 10.1002/da.22355] [Medline: 25693731]

67. Gonzalez R. The relationship between bipolar disorder and biological rhythms. J Clin Psychiatry 2014 Apr;75(4):e323-e331. [doi: 10.4088/JCP.13r08507] [Medline: 24500332]

68. Beyer KM, Szabo A, Nattinger AB. Time spent outdoors, depressive symptoms, and variation by race and ethnicity. Am J Prev Med 2016 Sep;51(3):281-290. [doi: 10.1016/j.amepre.2016.05.004] [Medline: 27320702]

69. Valenza G, Gentili C, Lanatà A, Scilingo EP. Mood recognition in bipolar patients through the PSYCHE platform: preliminary evaluations and perspectives. Artif Intell Med 2013 Jan;57(1):49-58. [doi: 10.1016/j.artmed.2012.12.001] [Medline: 23332576]

70. Cummins N, Scherer S, Krajewski J, Schnieder S, Epps J, Quatieri TF. A review of depression and suicide risk assessment using speech analysis. Speech Communication 2015 Jul;71:10-49. [doi: 10.1016/j.specom.2015.03.004]

71. Vanello N, Guidi A, Gentili C, Werner S, Bertschy G, Valenza G, et al. Speech analysis for mood state characterization in bipolar patients. Conf Proc IEEE Eng Med Biol Soc 2012;2012:2104-2107. [doi: 10.1109/EMBC.2012.6346375] [Medline: 23366336]

XSL·FO

RenderX

72. Gideon J, Provost EM, McInnis M. Mood state prediction from speech of varying acoustic quality for individuals with bipolar disorder. Proc IEEE Int Conf Acoust Speech Signal Process 2016 Mar;2016:2359-2363 [FREE Full text] [doi: 10.1109/ICASSP.2016.7472099] [Medline: 27570493]

73. Hidalgo-Mazzei D, Mateu A, Reinares M, Murru A, Del Mar Bonnín C, Varo C, et al. Psychoeducation in bipolar disorder with a SIMPLe smartphone application: feasibility, acceptability and satisfaction. J Affect Disord 2016 Aug;200:58-66. [doi: 10.1016/j.jad.2016.04.042] [Medline: 27128358]

74. Faurholt-Jepsen M, Ritz C, Frost M, Mikkelsen RL, Margrethe CE, Bardram J, et al. Mood instability in bipolar disorder type I versus type II-continuous daily electronic self-monitoring of illness activity using smartphones. J Affect Disord 2015 Nov 01;186:342-349. [doi: 10.1016/j.jad.2015.06.026] [Medline: 26277270]

75. Valenza G, Nardelli M, Lanata A, Gentili C, Bertschy G, Kosel M, et al. Predicting mood changes in bipolar disorder through heartbeat nonlinear dynamics. IEEE J Biomed Health Inform 2016 Apr 20:1. [doi: 10.1109/JBHI.2016.2554546] [Medline: 28113920]

76. Valenza G, Nardelli M, Lanatà A, Gentili C, Bertschy G, Paradiso R, et al. Wearable monitoring for mood recognition in bipolar disorder based on history-dependent long-term heart rate variability analysis. IEEE J Biomed Health Inform 2014 Sep;18(5):1625-1635. [doi: 10.1109/JBHI.2013.2290382] [Medline: 24240031]

77. Burton C, Szentagotai TA, McKinstry B, Matheson C, Matu S, Moldovan R, Help4Mood Consortium. Pilot randomised controlled trial of Help4Mood, an embodied virtual agent-based system to support treatment of depression. J Telemed Telecare 2016 Sep;22(6):348-355. [doi: 10.1177/1357633X15609793] [Medline: 26453910]

78. Anguera JA, Jordan JT, Castaneda D, Gazzaley A, Areán PA. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. BMJ Innov 2016 Jan;2(1):14-21 [FREE Full text] [doi: 10.1136/bmjinnov-2015-000098] [Medline: 27019745]

79. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile sensing and support for people with depression: a pilot trial in the wild. JMIR Mhealth Uhealth 2016 Sep 21;4(3):e111 [FREE Full text] [doi: 10.2196/mhealth.5960] [Medline: 27655245]

80. Arean PA, Hallgren KA, Jordan JT, Gazzaley A, Atkins DC, Heagerty PJ, et al. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. J Med Internet Res 2016 Dec 20;18(12):e330 [FREE Full text] [doi: 10.2196/jmir.6482] [Medline: 27998876]

81. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. J Med Internet Res 2015;17(7):e175 [FREE Full text] [doi: 10.2196/jmir.4273] [Medline: 26180009]

82. Warmerdam L, Riper H, Klein M, van den Ven P, Rocha A, Ricardo HM, et al. Innovative ICT solutions to improve treatment outcomes for depression: the ICT4Depression project. Stud Health Technol Inform 2012;181:339-343. [Medline: 22954884]

83. Olbrich S, van DR, Arns M. Personalized medicine: review and perspectives of promising baseline EEG biomarkers in major depressive disorder and attention deficit hyperactivity disorder. Neuropsychobiology 2015;72(3-4):229-240. [doi: 10.1159/000437435] [Medline: 26901357]

84. Young JJ, Silber T, Bruno D, Galatzer-Levy IR, Pomara N, Marmar CR. Is there progress? An overview of selecting biomarker candidates for major depressive disorder. Front Psychiatry 2016;7:72 [FREE Full text] [doi: 10.3389/fpsyt.2016.00072] [Medline: 27199779]

85. Comello ML, Qian X, Deal AM, Ribisl KM, Linnan LA, Tate DF. Impact of game-inspired infographics on user engagement and information processing in an eHealth program. J Med Internet Res 2016 Sep 22;18(9):e237 [FREE Full text] [doi: 10.2196/jmir.5976] [Medline: 27658469]

86. Markowetz A, Błaszkiewicz K, Montag C, Switala C, Schlaepfer TE. Psycho-informatics: big data shaping modern psychometrics. Med Hypotheses 2014 Apr;82(4):405-411. [doi: 10.1016/j.mehy.2013.11.030] [Medline: 24529915]

87. Naeem F, Gire N, Xiang S, Yang M, Syed Y, Shokraneh F, et al. Reporting and understanding the safety and adverse effect profile of mobile apps for psychosocial interventions: an update. World J Psychiatry 2016 Jun 22;6(2):187-191 [FREE Full text] [doi: 10.5498/wjp.v6.i2.187] [Medline: 27354959]

88. Torous J, Firth J. The digital placebo effect: mobile mental health meets clinical psychiatry. Lancet Psychiatry 2016 Feb;3(2):100-102. [doi: 10.1016/S2215-0366(15)00565-9] [Medline: 26851322]

89. Anthes E. Mental health: there's an app for that. Nature 2016 Apr 7;532(7597):20-23. [doi: 10.1038/532020a] [Medline: 27078548]

90. Kauer SD, Reid SC, Crooke AH, Khor A, Hearps SJ, Jorm AF, et al. Self-monitoring using mobile phones in the early stages of adolescent depression: randomized controlled trial. J Med Internet Res 2012;14(3):e67 [FREE Full text] [doi: 10.2196/jmir.1858] [Medline: 22732135]

91. Glenn T, Monteith S. New measures of mental state and behavior based on data collected from sensors, smartphones, and the Internet. Curr Psychiatry Rep 2014 Dec;16(12):523. [doi: 10.1007/s11920-014-0523-3] [Medline: 25308392]

92. Wenze SJ, Miller IW. Use of ecological momentary assessment in mood disorders research. Clin Psychol Rev 2010 Aug;30(6):794-804. [doi: 10.1016/j.cpr.2010.06.007] [Medline: 20619520]

XSL•FO
RenderX

93. Proudfoot J, Parker G, Hadzi PD, Manicavasagar V, Adler E, Whitton A. Community attitudes to the appropriation of mobile phones for monitoring and managing depression, anxiety, and stress. J Med Internet Res 2010;12(5):e64 [FREE Full text] [doi: 10.2196/jmir.1475] [Medline: 21169174]

94. Monteith S, Glenn T, Geddes J, Whybrow PC, Bauer M. Big data for bipolar disorder. Int J Bipolar Disord 2016 Dec;4(1):10 [FREE Full text] [doi: 10.1186/s40345-016-0051-7] [Medline: 27068058]

95. Bussmann JB, Ebner-Priemer UW, Fahrenberg J. Ambulatory activity monitoring. Eur Psychol 2009 Jan;14(2):142-152. [doi: 10.1027/1016-9040.14.2.142]

96. Dell'Osso L, Pini S, Cassano GB, Mastrocinque C, Seckinger RA, Saettoni M, et al. Insight into illness in patients with mania, mixed mania, bipolar depression and major depression with psychotic features. Bipolar Disord 2002 Oct;4(5):315-322. [Medline: 12479664]

Abbreviations

AMoSS: Automated Monitoring of Symptom Severity
ECG: electrocardiogram
EMA: Ecological Momentary Assessment
E-mental health: electronic mental health
GPS: global positioning system
HDRS-17: Hamilton Depression Rating Scale (17 item version)
HRV: heart rate variability
ID: identification
PHQ-9: Patient Health Questionnaire
PRISMA: Preferred Reporting Items for Systematic Review and Meta-Analysis
RCT: randomized controlled trial
SD: standard deviation
SMS: short messaging system
SRM-5: Social Rhythm Metric
YMRS: Young Mania Rating Scale