Designing Mobile Health Technology for Bipolar Disorder: A Field Trial of the MONARCA System

Jakob E. Bardram, Mads Frost, Károly Szántó
The Pervasive Interaction Technology Laboratory
IT University of Copenhagen, Denmark
{bardram,madsf,ksza}@itu.dk

Maria Faurholt-Jepsen, Maj Vinberg and Lars Vedel Kessing
Psychiatric Center Copenhagen,
University Hospital of Copenhagen, Denmark
<firstname.lastname>@regionh.dk

ABSTRACT
An increasing number of pervasive healthcare systems are being designed, that allow people to monitor and get feedback on their health and wellness. To address the challenges of self-management of mental illnesses, we have developed the MONARCA system – a personal monitoring system for bipolar patients. We conducted a 14 week field trial in which 12 patients used the system, and we report findings focusing on their experiences. The results were positive; compared to using paper-based forms, the adherence to self-assessment improved; the system was considered very easy to use; and the perceived usefulness of the system was high. Based on this study, the paper discusses three HCI questions related to the design of personal health technologies; how to design for disease awareness and self-treatment, how to ensure adherence to personal health technologies, and the roles of different types of technology platforms.

Author Keywords
Bipolar disorder; mental health; personal health systems; mobile application

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
According to WHO, mental illness is one of the most pressing healthcare concerns worldwide [34]. Bipolar disorder in particular, has a community lifetime prevalence of 4% [16] and is associated with high morbidity and disability [25]. Personal health technologies hold promise for helping bipolar patients to monitor their mood patterns and symptoms, recognize so-called 'early warning signs', and to handle medication. Health technologies can – based on subjective and objective sensor input – provide timely feedback to the patient and thereby increase their awareness of the disease. Smartphones are a promising platform for such personal feedback systems due to their ubiquitous availability and connectivity. Consequently, a number of personal monitoring and feedback systems have been suggested for the management of a wide range of health-related conditions. In general, these types of systems help users by enabling them to monitor and visualize their behavior, keeping them informed about their physical state, reminding them to perform specific tasks, providing feedback on the effectiveness of their behavior, and recommending healthier behavior or actions.

However, introducing new technology for patients with psychiatric disorders, who often have a low coping ability, may be stressful for them and introduce a high cognitive load. Unfortunately, mentally ill patients tend to be socially challenged as well, having a larger tendency for unemployment and alcohol or substance abuse [14]. As such, designing for this group of users is challenging, and the introduction of new technology may not be well adopted and used. Hence, a core research question is to what degree systems for mentally ill patients can be designed, to what degree such technologies will be adopted and used, and how it will lead to new ways for the patients and clinicians to treat this group of patients, compared to the existing approaches.

In this study, we examined the use of a personal health monitoring and feedback system for patients suffering from bipolar disorder, called the MONARCA system [1]. The system lets patients enter self-assessment data, it collects sensor data, it provides feedback on the data collected, and helps them manage their medicine. The study is based on a 14 week field trial of the system. The main objective of this study was to establish the feasibility and usefulness of the system by looking into whether; (i) it is sufficiently stable for general use; (ii) the usability of the system, focusing especially on investigating if it is as easy to use as the currently used paper-based self-assessment forms; (iii) the general usefulness of the system in terms of helping bipolar disorder patients cope with their disease; and (iv) the system – if used on a daily basis by bipolar patients – will be useful to them in the future.

The results of this study are used in a discussion of the broader implications for design of personal health technologies by addressing three questions; (i) how to design for disease awareness and self-treatment; (ii) how to ensure adherence to using personal health technologies, and (iii) what are the roles of different types of technology platforms.

BACKGROUND AND RELATED WORK
Bipolar disorder is a mental illness characterized by recurring episodes of both depression and mania. Treatment of bipo-

CHI 2013, April 27–May 2, 2013, Paris, France.

lar disorder aims to reduce symptoms and prevent episodes through a combination of (i) pharmacotherapy where mood is stabilized, and symptoms are controlled, using a customized and difficult to determine combination of several drugs like antidepressants, antipsychotics, and mood stabilizers; (ii) psychoeducation where patients are taught about the complexities of bipolar disorder, causes of recurrence of episodes, and how to manage their illness, and (iii) psychotherapy where patients are coached to be aware of their symptoms and find practical ways to prevent episodes through actionable behavioral and lifestyle choices, such as routine, sleep, and social activity.

However, patients' decreased recognition and insight into the illness and poor adherence to medication [15] are major challenges which increase the risk of recurrence in bipolar disorder. Therefore, continuous mood tracking and graphing [30], recognizing and controlling so-called Early Warning Signs (EWS), activity logging, and medication compliance training are core ingredients in cognitive behavioral training (CBT) for the experienced, but not yet stable bipolar disorder patient [2].

Paper-based mood charting forms are frequently applied, but possess significant limitations; they are inconvenient to fill out, are highly subjective, and they are filled out inconsistently due to forgetfulness or symptoms. Research has shown that paper-based charting suffers from a range of problems [4, 23, 32]: low adherence rates, unreliable retrospective completion of diaries, and time intensive data entry.

Various electronic monitoring systems have been presented for patients with bipolar disorder including self-monitoring of medication, mood, sleep, life events, weight, and menstrual data. PC-based [33, 4] Web-based, SMS-based [5], and mobile phone [23, 27] solutions exist, which have been shown to make data reporting easier for patients, thereby reducing data inconsistency and increasing adherence rates. So far none of these monitoring systems have included combined self-monitoring and objective system recording of the disorder, and none of them have built-in mechanisms for providing historical data visualization or personal feedback directly to the patient on the phone.

Personal Health Technologies

Recently, several research projects have investigated the use of personal technology to encourage healthy behavior. Personal health technologies can be grouped into three broad categories.

The first set of systems can be labeled 'wellness' applications, which seek to 'persuade' users to make healthy behavior change such as increased physical activity [20, 7], healthy eating habits [26], or better sleep [3]. For example, Fish'n Steps [20] and UbiFit Garden [7] seek to encourage physical activity; the Time to Eat! iPhone application is a persuasive game encouraging healthy eating habits [26]. Lately, systems like the BeWell application have proposed a more comprehensive Smartphone-based approach that can track activities that impact physical, social, and mental wellbeing – namely, sleep, physical activity, and social interactions – and provides intelligent feedback to promote better health [17].

The second category comprises systems targeted management of chronic somatic diseases like diabetes [21], chronic kidney disease [31], and asthma [18].

The third category contains systems which address issues of mental health and illness such as stress [12] and depression [29, 6], and more general-purpose mobile phone systems for mood charting [23, 24]. Two systems of particular relevance to our study are the Mobile Mood Diary [23] and Mobilyze! [6]. The Mobile Mood Diary uses a mobile phone to allow patients to report mood, energy, and sleep levels, which can then be accessed on a website. The study showed increased patient adherence to mood charting using the phone as compared to using paper-based forms. Mobilyze! is a mobile phone application using machine learning models to predict patients moods, emotions, cognitive/motivational states, activities, environmental contexts, and social contexts, based on phone sensor data like GPS, ambient light, recent calls, etc. The website contains graphs illustrating patients self-reported states, as well as didactic tools, that teach patients behavioral activation concepts. The feedback mechanism uses telephone calls and emails from a clinician in order to promote adherence. A small feasibility study showed, among other things, that accuracy rates of up to 91% were achieved when predicting categorical contextual states (e.g., location), but that for states rated on scales, especially mood, predictive capability was poor. The study showed, however, that patients were satisfied with the phone application and it improved on their self-reported depressive symptoms.

The MONARCA system belongs to this third category. In some ways, it is similar to the Mobile Mood Diary and Mobilyze! systems, but the MONARCA system is targeted to bipolar disorder, rather than uni-polar depression. The MONARCA system also seeks to provide direct and timely feedback to the patient by visualizing self-assessment data and objectively sensed data directly on the phone. Moreover, rather than having clinicians phone or email patients, the MONARCA system has a built-in trigger and notification feature. The system is thus designed to scale better, organizationally, since the feedback mechanisms are not tied solely to a human actor (i.e. clinician).

THE MONARCA SYSTEM

The MONARCA system was designed in a user-centered participatory design process [13] involving a group of three psychiatrists and seven patients suffering form bipolar disorder. Specifically, we used the Patient-Clinician-Designer (PCD) Framework [22], which outlines how key principles of user-centered design – including user focus, active user involvement, evolutionary systems development, prototyping, and usability champions – can be applied in the context of designing for mental illness. Through the PCD process, patients and clinicians were instrumental in making decisions about system features through collaborative design workshops and iterative prototyping. Three-hour workshops were held every other week for six months. The designers led each workshop by facilitating discussion about particular design goals and issues, system features and functionality, and feedback on mockups and prototypes of the system.

The final design of the MONARCA system[1] contains 5 features that are targeted specifically to help bipolar patient manage their illness: (i) self-assessment of self-reported data like mood, sleep, and alcohol; (ii) activity monitoring in terms of sampling sensor data from the phone; (iii) historical overview of self-assessment and sensed data; (iv) coaching & self-treatment based on customizable triggers and detection of early warning signs (EWS); and (v) data sharing between the patient and the clinician.

The overall approach is that self-assessment and review of various parameters can support bipolar illness management. For example, patients and their clinicians can use the data to determine adherence to medications, investigate illness patterns and identify early warning signs for upcoming affective episodes, or test potentially beneficial behavior changes. Through monitoring and feedback, the MONARCA system may be able to help patients implement effective short-term responses to warning signs and preventative long-term habits.

Similar to other personal health technologies, the design of the MONARCA system employs a mobile phone application as the main component. Using a mobile phone was an obvious design choice since they were already used extensively by all patients.

Android Phone Application

Figure 1 shows the 5 main screens of the MONARCA Android application; (i) inputting self-assessment data, (ii) historic data visualizations, (iii) prescribed medicine; (iv) activated 'triggers' and suggestions for 'actions to take'; and (v) a screen for various settings, such as an alarm reminding the patients to enter their self-assessment data. In addition to collecting 'subjective' self-reported data, the application also collects 'objective' sensor data.

Subjective and Objective Data Sampling

A significant part of the design process was spent on designing the self-assessment form. First of all, it was important that relevant data for bipolar disorder patients were collected. As discussed in [22], there is a tradeoff between the clinicians' need to collect clinically relevant and 'objective' data, and the patients' need for collecting more personalized data. Second, significant effort has been put into designing the self-assessment form on the phone application, so that it is as simple and short as possible. A core requirement from the patients involved in the design process, was that the list of self-reported items should be kept to a minimum and that self-assessment should be done quickly.

Based on thorough design discussions, the final version of the MONARCA self-assessment form contains a minimum set of things to monitor, which can be divided into a set of *mandatory* self-assessment items, which is absolutely crucial for clinicians to collect over time in the treatment of a bipolar patient, and a set of *optional* self-assessment items, which supplement the mandatory ones.

[1]The design of the MONARCA system and its technical implementation has been presented in [1], where many more technical details can be found. This section provides an overview of the system design and its core features.

Figure 1. The MONARCA Android application user interface.

The mandatory self-assessment items are:

- *Mood* measured on a 7-point HAMD scale spanning from highly depressed (-3) to highly manic (+3). As a mood-disorder illness, self-reported mood is the main data parameter to follow for bipolar disorder patients.
- *Sleep* indicated in half-hour intervals. Significant clinical evidence shows a direct link between the mood of a bipolar patient and the amount of sleep (sleep increases during depressed periods, and decreases during manic periods).
- *Subjective Activity* on a 7-point scale spanning from totally inactive (-3) to highly active (+3). Bipolar disorder patients report themselves to be more active during manic periods and less active during depressed periods, and self-reported activity level is, like sleep, a very good indicator of the state of the illness.
- *Medicine Adherence* by specifying whether prescribed medicine has been taken as prescribed, have been taken with modification, or not taken at all. Since medical treatment of bipolar disorder is very effective and can significantly help stabilize the patient's mood, keeping track of medicine adherence is core to medical treatment.

The optional self-assessment items include:

- *Universal Warning Signs*, which are signs that a psychiatric clinic can set up for all its patients. Such signs can e.g. include experience of so-called 'mixed mood', 'cognitive problems', or 'irritability'.
- *Early Warning Signs (EWS)*, are *personal* signs that are tailored specifically to each patient, and inform them of things to look out for. For example, if a patient begins to sleep in the living room, rather than the bed room, this is a sign for him that a manic phase is under way.
- *Alcohol*, as measured in number of drinks. For some bipolar patients, alcohol and drug abuse can be associated with their illness.

- *Stress* measured on a 6-point scale from 0 to 5. Self-perceived stress can be a significant trigger of a depressive period, and for some patients monitoring their stress level is important.

- *Note*, a free text entry done with an on-screen keyboard. This can be used to associate a note to the data entered or for entering more generic comments for this day.

On a daily basis, an alarm on the phone reminds the patient to report self-assessment data for that day. Only data for the current day can be entered and the system does not allow patients to revise earlier entries.

In addition to self-assessment data, the phone samples behavioral data via sensors in the phone. This includes *physical activity* data as measured by the accelerometer, and *social activity* as measured by the number of in- and outgoing phone calls and text messages. This sensor sampling is aggregated into two simple figures reflecting the level of physical and social activity on a given day, and is visualized as simple graphs on the data visualization screen.

Feedback Mechanisms – Visualization and Triggers

The visualization screen (Figure 1(ii)) is designed to be the main feedback mechanism to the patient. This screen is shown when the patient has entered his or her self-assessment data. The graph visualization display is designed to be very simple, while giving an overview of the self-assessment and sensed data. The phone only shows data for the past 14 days, whereas longer periods of data can be seen on the website.

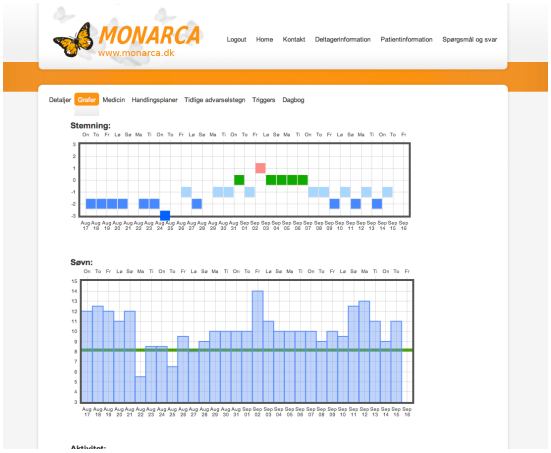

Figure 2. Website - Patient Data Visualization, showing the mood and sleep graphs.

The second feedback mechanism is the automatic trigger feature. A 'trigger' consists of a set of rules that apply to any self-assessment data being entered. For example, a trigger can be set up to trigger if the patient reports that he has been sleeping less that 6 hours, 3 days in a row. So-called 'actions-to-take' can be associated with a trigger. Actions-to-take are simple behavioral suggestions to a patient in different situations. For example, in case the patient sleeps to little, he or she can try using sleeping pills, or make sure to sleep in a cold and dark room. Triggers and actions-to-take are personalized to each patient.

When a trigger is activated, a notification is posted using Android's notification mechanism. The trigger is then displayed as an item in the notification view on the Android phone (typically in the top pull-down curtain). When clicking the notification, the patient is taken to the *Actions-to-Take screen* (Figure 1(iv)), which lists all active triggers and their associated actions-to-take.

Automatic triggers are designed to play a core role in continuous and automatic feedback to the the patient, since they consistently track patterns over time and can warn both the patient and the psychiatrist about things to be aware of.

Website

The MONARCA system can also be accessed via a website, which is designed to be used by patients and clinicians. Patients can see and update their personal data as shown in Figure 2, manage personal triggers and early warning signs, and configure the system. Clinicians are shown a dashboard that provides an overview of their patients and how they are doing on the core parameters of mood, activity, sleep, and medicine adherence for the last 4 days. From this dashboard, they can access detailed data on each patient and customize the system according to the needs of each individual patient, including medication.

FIELD TRIAL OF MONARCA

Clinical trials of inventions for mental health are very resource consuming, and often a staged evaluation strategy investigating both clinical and HCI issues can be more beneficial and informative [8, 23]. Hence, the MONARCA system was deployed in a single-arm feasibility trial that tested feasibility rather than efficacy. The study ran from May to August 2011, a total of 14 weeks. The study design was approved by the Danish National Committee on Health Research Ethics and the security and data handling was approved by the Danish Data Protection Agency, ensuring that everything was done according to standards. Informed consent was obtained from all patients.

The main objective of this study was to gauge the feasibility of the system as used by patients suffering from bipolar disorder. This was done by focusing on the following four questions:

$Q1$ – Is the system sufficiently stable for general use?

$Q2$ – How usable is the system and is it better than existing approaches for self-assessment and data collection?

$Q3$ – What is the usefulness of the system in terms of helping bipolar disorder patients in coping with their disease?

$Q4$ – Will this system – if used on a daily basis by bipolar patients – be useful to them in the future?

An important part of the study was to investigate the usefulness of the system during the trial ($Q3$) and the perceived usefulness in the future ($Q4$), and as such the study aimed at establishing the benefit for bipolar patients in managing their disease. As such, the goal of this study was to establish the feasibility of the system, and if this study was positive, to move into a clinical trial afterwards. Thus, the focus was on the patients using the system, and not the clinicians or the system used in treatment.

Trial Setup and Participant Recruitment

The study had three inclusion criteria: the patients should be; (i) between 18 and 65 years, (ii) able to use a mobile phone and a website; (iii) stable patients. No exclusion criteria were set up. Potential patients were referred to the study by doctors in the Affective Disorder Clinic at a university hospital. The doctor associated with the MONARCA study initially phoned each patient, introduced the project, and asked if they would be interested in participating. If so, they would meet with the doctor and a technician at the clinic, during which they were further informed of the project. If the patient was still interested in participating, an informed consent form was signed by the patient and the doctor. The patient then got a thorough introduction to the MONARCA system, received a printed user guide, and was issued a standard HTC Desire Smartphone. They were additionally helped to insert their SIM card, set up 3G internet, and transfer relevant content such as contacts. No compensation was paid to the patients, but they were reimbursed for their 3G internet subscription.

The patients received contact information (phone + email) for the MONARCA doctor and technician, whom the patients could contact if they had any questions or problems with the system. The MONARCA doctor would oversee the data on a daily basis and could contact the patients by phone if needed. The patients were still treated by their own doctor – who also had access to the patient's data via the website – but given that this was a feasibility study, the system was not fully integrated into the usual treatment at the hospital, as it would not be ethically sound to do this with an untested system. The MONARCA doctor would not engage in any treatment of the patients, but would notify the patient's doctor if necessary.

Methods

In order to answer the four questions above, we applied several methods. First, adherence to self-assessment was measured in two ways. Adherence to the paper-based forms was gauged by collecting and analyzing the paper-based mood assessment forms, used by the participants in 62 days from March to May 2011 (i.e., just prior to the launch of the MONARCA system). Adherence to the MONARCA system was measured by the number of daily self-assessments extracted from the system database. Second, we measured the usability of the MONARCA system by applying the IBM Computer System Usability Questionnaire (CSUQ) [19] online. Third, we issued an online questionnaire asking questions about the usefulness of the system during the trial period. Fourth, an online questionnaire containing the same questions, but now in future tense, was issued to investigate the *perceived* usefulness of the system in the future. There is a significant correlation between users' perceived usefulness of a system, and its actual future usefulness [9], and this analysis can hence be used to gauge the potential of the MONARCA system in a future clinical deployment. Fifth, we did semi-structured follow-up interviews with all participants at the end of the trial.

RESULTS

Participant Characteristics

28 patients were contacted initially, whereof 17 were interested and came to the clinic for further interviews. 14 patients were enrolled in the trial, of which 2 dropped out; one in week 2 because she wanted to go back to her iPhone, and one in week 7 due to a lack of time. Thus, a total of 12 bipolar patients participated in the field trial. The left column of Table 1 shows the list of participants and their demographic backgrounds. We were able to recruit a very diverse set of patients of different gender (5 male, 7 female); age (20–51 years); IT skills; and mobile phone experience. All participants were selected among stable patients having initial HAMD mood scores in the range of -1 to $+1$.

Participation

Figure 3 shows the use of the system during the trial period, including the number of phones reporting self-assessment ('Subjective') data as well as sensor ('Objective') data on a daily basis. The graph illustrates that phones were deployed during May, and usage peaked in mid June to mid August, and that almost all 12 phones reported data on a daily basis. During August, we experienced an error in the Android Market that locked the application. This was not discovered until the trial was over, and based on post-trial interviews, this error seems to explain the decline in use during August.

Adherence Results

Table 1 shows the rate of self-assessment when using both the paper-based forms as well as the MONARCA system. From this table we can observe several things.

First, on average, the length of the paper-based and phone-based trials are comparable (62 and 69 days), with some variation in the system trial. If the phone is working (i.e., charged), the application will sample objective data. On average, sampling was done 63 out of 69 days and the application was hence running 92% of the trial period.

Second, we can compare the adherence to self-assessment using the paper-based forms and the MONARCA system. In the paper-based forms, we count the number of days that any information is noted on the paper – irrespective of the level of detail. When using paper-based forms, the raw adherence percentage is 58%, which is similar to what was found in the Mobile Mood Diary study [23]. But, if not counting the four participants who did not fill in their self-assessment at all (P48;P63;P67;P70), the average adherence is 87%. The general adherence percentage when using the MONARCA system is 80% for all of the involved 12 participants. If we only take into consideration the days where the system was actually working (63 instead of 69), the adherence rate is 87%. Hence, the adherence rate for the paper-based and phone-based systems are comparable if only counting the days where data can be recorded and only involving participants who also reported data on their paper-based forms. However, the interviews revealed that paper-based forms were subject to significant retrofitting and we will discuss this further in the Discussion section below.

Third, looking specifically at the four mandatory self-assessment parameters of mood, sleep, activity, and medicine we see that all patients have high compliance scores (3.76 out of 4.00 is 94%), which is very positive since acquiring this data is a core goal of the system. On the paper-based self-assessment form, only 3 out of these four parameters can be

Participant	*Gender*	*Age*	*IT skills (1–5)*	*Mobile phone*	*Years of phone use*	*Mood score*	*Occupation*	*# self-assessment days*	*% adherence*	*# primary scores*	*# days participated*	*# self-assessment days*	*# days reporting data*	*% general adherence*	*% adherence (run. system)*	*# of 4 primary scores*	*# visits to website*
P48	m	29	4	iPhone	15	-1	Student	0	n/a	n/a	65	50	57	76	87	4.00	0
P49	f	50	4	Nokia	20	0	Unemployed	59	95	2.88	40	26	38	65	68	3.80	1
P55	m	29	2	Nokia	13	0	Shipping	59	95	2.97	99	60	60	60	100	3.87	0
P57	f	35	4	SonyE.	15	0	Accountant	62	100	3.00	90	76	86	84	88	4.00	0
P58	f	34	2	Samsung	10	0	Teacher	43	69	2.04	98	96	97	97	98	3.31	6
P59	f	38	4	Nokia	13	-1	Unemployed	43	69	2.77	98	74	92	75	80	3.77	0
P61	f	34	4	iPhone	14	0	Self-employed	59	95	2.34	68	64	68	94	94	3.56	1
P63	f	20	5	HTC	9	0	Student	0	n/a	n/a	22	8	14	36	57	3.38	0
P64	f	51	2	Nokia	14	1	Pensioner	59	95	2.22	77	77	77	100	100	3.87	15
P66	m	45	4	SonyE.	12	0	Student	50	80	3.00	70	61	69	87	88	3.80	51
P67	m	37	5	iPhone	15	1	Ph.D. student	0	n/a	n/a	53	46	52	86	88	3.78	11
P70	m	37	4	iPhone	15	-1	Musician	0	n/a	n/a	49	47	49	95	95	4.00	2
Avr.		36			14			36	87	2.65	69	57	63	80	87	3.76	7

Table 1. Participation in the MONARCA trial study. From left: participation ID; demographic data; and data from the normal paper-based self-assessment forms, and usage data for the 14 weeks trial study of the MONARCA system.

CSUQ item	*Description*	*avg.*	*sd.*
OVERALL	Overall satisfaction	2.60	1.01
SYSUSE	System usefulness	1.93	0.42
INFOQUAL	Information quality	3.32	1.10
INTERQUAL	Interface quality	2.71	0.93

Table 2. The CSUQ usability results on a Likert scale from 1–7: 1=Highly agree; 7=Highly disagree.

reported (activity is missing), and we see a slightly lower degree of compliance (2.65 out of 3.00 is 88%).

Finally, looking at the use of the website, it is quite evident that this is not used by most of the participants. P66, P64 and P67 show moderate use. The extensive use by P66 is due to this patient being interested in the data visualization, where he accessed the website constantly, just to look at the graphs.

System Usability

Table 2 shows the usability scores as measured by the IBM Computer System Usability Questionnaire (CSUQ) on a 7-point Likert scale from 'Strongly Agree' (1) to 'Strongly Disagree' (7). From these scores we can conclude that the overall usability of the system is good (OVERALL = 2.60) and the users found the system very useful (SYSUSE = 1.93). This reflects a low score in simplicity, comfortability and learnability, and efficiency. The information quality score is lower (INFOQUAL = 3.25) which can be ascribed to problems with the error messages of the system, which scored 5.33, and did not help users fix the problems they may have experienced. Finally, the system scores well in interface quality in general (INTERQUAL = 2.86), but the study showed that it did not have all the functions and capabilities that patients expected.

	System Usefulness		*Perceived Usefulness*	
	avg.	*sd.*	*avg.*	*sd.*
Disease Mgmt.	3.16	1.55	2.16	1.02
Self-assessment	2.21	1.06	1.73	0.72
Visualization	2.22	1.39	1.66	0.78
Alarms	2.34	1.44	2.13	1.88
Triggers	3.59	1.31	2.71	1.02
Early Warning Signs	3.44	1.18	2.36	0.78
Actions to take	3.25	1.52	2.34	0.88
Medication	4.30	1.50	3.17	1.51
Website	3.00	1.70	2.63	1.76

Table 3. Questionnaire results on 'System Usefulness' as used in the trial period and 'Perceived Usefulness' in the future. Users reported on a 1–7 point Liket scale on the question of "The MONARCA system is useful for ...": 1=Highly agree; 7=Highly disagree.

This was also mentioned in the interviews, and we will return to this in the discussion below.

Usefulness and Perceived Usefulness

In the usefulness questionnaire we asked 38 questions divided into 10 categories, using the 7-point Likert scale. The categories and average scores are shown in Table 3. The usefulness of the system for disease management during the trial scored 3.16. This means that patients agree (though not strongly) that MONARCA helped them in managing their bipolar disorder. This category addressed wether the patients became better at managing their disease (2.92), wether they were made more aware of their disease (2.50), the specific usefulness of the application for disease management (3.33), the usefulness of the website (4.33), and if the application

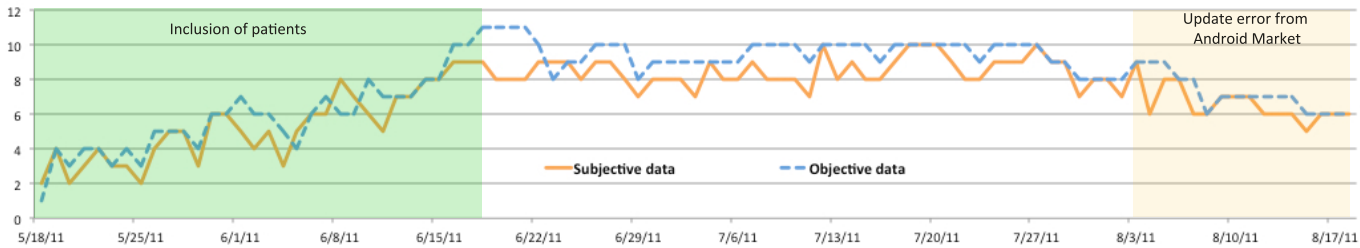

Figure 3. The usage of the MONARCA system during the trial period, showing the number of phones reporting subjective self-assessment data and objective sensor data on a daily basis.

made them change their lifestyle (4.00). Hence, a clear indication was that the system as such, did not have an effect on changing lifestyle, but more on disease management and awareness. When asking the same questions, but on perceived future usefulness, the patients generally agreed that the system would help them cultivate better disease management and awareness. Even the question on whether the system would make patients change their lifestyle scored relatively well (2.67). Hence, patients think that this system – if used in the future – would also assist in changing their lifestyle. The questionnaire also inquired about more specific aspects of the MONARCA system. Table 3 shows that the features of 'self-assessment', 'visualization', and 'alarms' were found to be the most useful features now and in the future. Hence, the patients found it very useful to be reminded to enter self-assessment data and see a temporal visualization of it. Moreover, the use of visualizations is perceived as the most useful feature of the system in future use. Features like 'triggers', 'early warning signs' and 'actions to take' are found to be useful, though less useful compared to the self-assessment functionality. Managing medication in the MONARCA system was found not to be useful. The usefulness of the website seems to be rather low and as shown in Table 1, it was not accessed by many patients.

DISCUSSION

Personal health technologies hold promises for both clinical and qualitative improvements of the healthcare model of the Western world. Even though this study did not provide unequivocal clinical evidence that the MONARCA system had any effect (positive or negative) on the patients' disease, the field trial has shown that the system is definitely feasible to use for the management of bipolar disorder. To be successful, a core requirement for such systems is that they are easy to use and are perceived as useful by patients. As such, the design of personal health technologies is a highly important topic for human-computer interaction.

Based on the insights from this trial test of the MONARCA system, combined with existing human-computer interaction studies of the design and use of personal health technologies, we have identified three core questions, which have to be addressed in the design of personal health technologies:

1. How to design for disease awareness and self-treatment?
2. How to ensure adherence to using personal health technologies?
3. What are the roles of different types of technology platforms?

This section discusses these core human-computer interaction design questions for personal health technologies in greater detail.

Designing for Disease Awareness and Self-treatment

As described in the Background section, continuous mood tracking, recognizing and controlling early warning signs, activity logging, and medication compliance training are core ingredients in cognitive behavioral training (CBT) for bipolar disorder patients [30, 2].

The field trial indicates that this kind of disease awareness and self-treatment was supported by the MONARCA system. The usability and usefulness scores show that the patients found the system very usable and useful in disease management, which was also reflected in the interviews:

> *"I am surprised how much I like it! [...] [MONARCA] is filled with substance and I have really benefited from it in relation to my illness. I have never used the [paper-based] mood charts that much, and I have never had much awareness about my [data] history [...] so I have been extremely happy with it, and I really think it is great."* [P70]

This quote also hints at the main reason behind the usefulness of the system, as P70 argues that the awareness of the historic data is highly useful. P70 continues;

> *"What I saw [in the trial] is that it helped me keep on track. I try to keep track of the triggers [early warning signs], and my history – and in that way it has helped me enormously. Previously, I went into periods where I encountered random mood swings, up and down, and I did not have any history [data] to relate to, so it kind of surprised me. But now I can actually follow how I'm doing – also back in time – and what caused it. It has really been great, and I think I have been able to keep track of myself."* [P70]

This insight is also reflected in the study data. Table 3 shows that patients found the ability to enter self-assessment data and later to review them in the visualization to be the most useful features of the system.

In the design of the MONARCA system, several visualization techniques and metaphors were discussed. Similar to other personal health systems that use fishes and flowers as metaphors, we were looking for an appropriate metaphor for bipolar disorder. Many attempts were tried, including using metaphors like a scale, an equalizer, a river, a volcano, a dart board, and a radar, but we always had the case that some

patients prefered one visualization, and others hated it – as put by P67 one day; "*I do not want my disease reduced to a game!*" In order to avoid problems, we designed the MONARCA system using a neutral graph visualization metaphor, but it seems that there is a need for allowing patients to use different visualizations. In general, it seems that support for tailorability and personalization of what self-assessment data to collect and how to visualize it, is important in the design of these types of systems.

Features like 'triggers', 'early warning signs' and 'actions to take' are also found to be useful for maintaining an awareness of the development of the disease, and to react if something change. The perceived usefulness of these features in the future is even higher.

The medicine overview was, however, not found to be particularly useful. In the interviews, patients explained that the overview was fine, but it did not increase their awareness of their medication. The patients wanted to be able to adjust their medication themselves; something they are allowed to do by the psychiatrist in order to fine-tune their medicine intake according to changes in mood and other symptoms. Also, the level of detail in the self-reporting on medicine intake was too coarse-grained; they could basically just specify if they took the prescribed amount of a drug, not if they took more or less.

In summary, the MONARCA system seems to be successful in providing the self-assessment and awareness of core disease parameters – except medication – which clinical studies have shown to have a positive effect on CBT treatment of bipolar patients.

Ensuring Adherence to Technology Use

As mentioned before, we saw that entering self-assessment data and using the system on a daily basis is core in building clinically beneficial disease awareness and self-treatment skills. This obviously then requires that the patient enters this data on a regular basis, and thus the whole complexity of 'adherence' becomes important in the design of personal health technologies.

The adherence rate of 87% in the MONARCA system is higher than the 65% adherence found in the Mobile Mood Diary system [23], which, however, was tested in a much longer period and may suffer from long-term effects that we did not encounter. We found that the adherence rate for the system was comparable to the paper-based forms. But several patients reported in the interviews that they actually retrofitted their paper-based self-assessment. As P49 said: *"The paper is more inaccurate – I sometimes put in data for several days at once, because I forget it"*. And P58 stated that: *"I used to fill out the paper for the whole week just before meeting with my doctor"*. This verifies the findings in the evaluation of the Mobile Mood Diary system, as well as other research, which has shown that paper-based charting suffers from a range of problems [4]: low adherence rates, unreliable retrospective completion of diaries, and time intensive data entry [32]. Note also, that in MONARCA, the patients can only fill in self-assessment scores on the current day; there is no way to report data back in time. As such, the MONARCA system provides much more valid day-by-day self-assessment data.

In the field trial, all patients reported that it is much easier to use the phone based self-assessment approach rather than using paper charts. As P49 explained:

> *"It is much easier to use the phone than the paper. I have the phone with me at all times, and I don't have to worry about the paper getting lost. It is very convenient that you can enter data when you experience things instead of having to recall it all when you fill in the paper"*. [P49]

An important factor in ensuring adherence to using the technology seems to be related to the 'alarm' feature. Initially, we feared that the alarm would be too obtrusive for the patients, but the evaluation showed that the patients found it very useful in reminding them to fill in the the self-assessment once a day. As P49 continues:

> *"I have to have this alarm on, otherwise I forget to fill in my self-assessment — I often forgot it using the paper. I could use a normal alarm, but I think it is nicer to have it as a part of the system"*. [P49]

In sum, ensuring adherence to the use of personal health technologies is core to their success. Based on our findings, adherence can be ensured by designing for simple and limited data input and using a reminder mechanism, like the alarm feature in the MONARCA system.

Technology Platforms for Personal Health Systems

This study showed that using a Smartphone that is always with the patient, ensures better adherence compared to paper-based charts. In this section we will discuss the applicability of different technology platforms for personal health systems, focusing specifically on the web and Smartphone platforms.

Several personal health systems rely on using web sites for treatment of mental disorders [28, 11] and have reported good results. However, the MONARCA study showed that only a few patients accessed the website. Information on medicine, warning signs, triggers, and general actions-to-take are configured for each patient together with the doctor at the clinic, but once this has been set up, the patients reported that there was little need to go trough the trouble of logging into the website; the settings seldomly needed to be changed and most of the data collected is available on the phone. As explained by P61:

> *"I logged on to the website the same evening I got the phone, but I didn't do anything there since my warning signs and triggers were already there. I have actually not visited it since, as I haven't had the need to alter anything, and I don't feel like I have enough data yet to actually go back and explore it."* [P61]

The usability scores (Table 2) show that the website scored lower than the rest of the system. The patients stated that it is partly because they did not have any real use for it, but also because they feel the design could be better. P57 states that *"compared to the phone, the website feels more disease-like, and not that personal."*

Turning to the Smartphone as a technology platform, the trial clearly indicates that this platform is well-suited to personal health systems. As argued by P49, "it is much easier to use the phone than the paper – I have the phone with me at all times". Studies show that Smartphones are within close distance to a user 90% of the time [10], and the fact that a Smartphone is a personal device which is always with the patient, makes it a good platform for these kind of systems.

This general finding is in line with other recent studies of personal health systems. For example, the Mobile Mood Diary study showed that a mobile phone was very useful for self-reporting of data and systems like the UbiFit Garden, Fishn Steps, BeWell, and Mobilize! have shown that Smartphones are useful for sensor data collection and visualization.

In contrast to the Mobile Mood Diary, however, the trial of the MONARCA system showed that text entry on the phone was not used. This again highlights that the design of personal health systems should take great care to limit the amount of self-reported data needed. However, during the interviews it became apparent that patients were not always satisfied with the current set of items; a fact that is also reflected in the mediocre score in the CSUQ information quality (INFOQUAL) score in Table 2. Given that bipolar disorder is an individual and diverse disease, it is important to be able to set up and track individual items other than the early warning signs. As argued by P67: *"I need to keep track of energy level and coffee"*, and P70: *"I feel there should be an anxiety item in the self-assessment, as it is important to me. I think all patients have different items that are important to them personally."*

CONCLUSION

This paper has reported one a 14 week field deployment and study of the MONARCA system, used by 12 patients. The MONARCA system helps patients suffering from bipolar disorder to do daily self-assessment and to get timely feedback on how they are doing using through visualization and feedback mechanisms. In the trial, we studied different aspects, including how the system was used and adopted as compared to the existing paper-based self-assessment forms, the usability of the system, and the usefulness of the system for the patients in managing their illness. The results were positive; compared to using paper-based forms, the adherence to self-assessment improved; the system was considered very easy to use; and the perceived usefulness of the system was high. As such, we can conclude that it is possible to design and deploy pervasive monitoring systems for mental illness such as bipolar disorder.

Based on the trial of the MONARCA system, we have discussed three core questions to address in the design of personal health technologies: (i) how to design for disease awareness and self-treatment; (ii) how to ensure adherence to using personal health technologies; and (iii) what is the role of different types of technology platforms. The presented trial of the MONARCA system showed that the system promoted disease awareness through self-assessment and temporal data visualization, adherence via the reminder system and the use of a Smartphone technology platform which is always with the patient. As a personal device, however, personal health systems should allow for personalization both in terms of what data to collect (self-reported and sensor-based) as well as how to visualize this data.

ACKNOWLEDGEMENT

We would like to thank the CHI reviewers, whose diligent reviews significantly improved the paper. The MONARCA project is funded as a STREP project under the FP7 European Framework program. More information can be found at `http://monarca-project.eu/` and `http://pit.itu.dk/monarca`.

REFERENCES

1. J. E. Bardram, M. Frost, K. Szántó, and G. Marcu. The monarca self-assessment system: a persuasive personal monitoring system for bipolar patients. In *Proc. ACM SIGHIT International Health Informatics Symposium*, IHI '12, pages 21–30, New York, NY, USA, 2012. ACM.
2. M. R. Basco and A. J. Rush. *Cognitive-Behavioral Therapy for Bipolar Disorder*. The Guilford Press, 2nd edition edition, 2005.
3. J. Bauer, S. Consolvo, B. Greenstein, J. Schooler, E. Wu, N. F. Watson, and J. Kientz. Shuteye: encouraging awareness of healthy sleep recommendations with a mobile, peripheral display. In *Proc. ACM CHI 2012*, CHI '12, pages 1401–1410, New York, NY, USA, 2012. ACM.
4. M. Bauer, P. Grof, N. Rasgon, T. Glenn, M. Alda, S. Priebe, R. Ricken, and P. C. Whybrow. Mood charting and technology: New approach to monitoring patients with mood disorders. *Current Psychiatry Reviews*, 2(4):423–429, 2006.
5. J. Bopp, D. Miklowitz, G. Goodwin, W. Stevens, J. Rendell, and J. Geddes. The longitudinal course of bipolar disorder as revealed through weekly text messaging: a feasibility study. *Bipolar Disord*, 12(3):327–334, 2010.
6. M. Burns, M. Begale, J. Duffecy, D. Gergle, C. Karr, E. Giangrande, and D. Mohr. Harnessing context sensing to develop a mobile intervention for depression. *J Med Internet Res*, 13(3), 2011.
7. S. Consolvo, D. W. McDonald, T. Toscos, M. Y. Chen, J. Froehlich, B. Harrison, P. Klasnja, A. LaMarca, L. LeGrand, R. Libby, I. Smith, and J. A. Landay. Activity sensing in the wild: a field trial of ubifit garden. In *Proc. ACM CHI 2008*, CHI '08, pages 1797–1806, New York, NY, USA, 2008. ACM.
8. D. Coyle, N. McGlade, G. Doherty, and G. O'Reilly. Exploratory evaluations of a computer game supporting cognitive behavioural therapy for adolescents. In *Proc. ACM CHI 2011*, CHI '11, pages 2937–2946, New York, NY, USA, 2011. ACM.
9. F. D. Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3):319–339, September 1989.

10. A. K. Dey, K. Wac, D. Ferreira, K. Tassini, J.-H. Hong, and J. Ramos. Getting closer: an empirical investigation of the proximity of user to their smart phones. In *Proc. UbiComp 2011*, UbiComp '11, pages 163–172, New York, NY, USA, 2011. ACM.

11. G. Doherty, D. Coyle, and J. Sharry. Engagement with online mental health interventions: an exploratory clinical study of a treatment for depression. In *Proc. ACM CHI 2012*, CHI '12, pages 1421–1430, New York, NY, USA, 2012. ACM.

12. P. Ferreira, P. Sanches, K. Höök, and T. Jaensson. License to chill!: how to empower users to cope with stress. In *Proc. NordiCHI 2008*, NordiCHI '08, pages 123–132, New York, NY, USA, 2008. ACM.

13. J. Greenbaum and M. Kyng, editors. *Design at Work - Cooperative Design of Computer Systems*. Lawrence Erlbaum Associates Publishers, Hillsdale, NJ, 1991.

14. L. Kessing. The effect of comorbid alcoholism on recurrence in affective disorder: a case register study. *J Affect Disord*, 53(1):49–55, 1999.

15. L. Kessing, L. Søndergard, K. Kvist, and P. Andersen. Adherence to lithium in naturalistic settings: results from a nationwide pharmacoepidemiological study. *Bipolar Disord*, 9(7):730–736, 2007.

16. T. A. Ketter. Diagnostic features, prevalence, and impact of bipolar disorder. *J Clin Psychiatry*, 71(6), June 2010.

17. N. D. Lane, T. Choudhury, A. Campbell, M. Mohammod, M. Lin, X. Yang, A. Doryab, H. Lu, S. Ali, and E. Berke. BeWell: A Smartphone Application to Monitor, Model and Promote Wellbeing. In *Proc. Pervasive Health 2011*. IEEE Press, 2011.

18. H. R. Lee, W. R. Panont, B. Plattenburg, J.-P. de la Croix, D. Patharachalam, and G. Abowd. Asthmon: empowering asthmatic children's self-management with a virtual pet. In *Extended Abstracts of CHI 2010*, CHI EA '10, pages 3583–3588, New York, NY, USA, 2010. ACM.

19. J. R. Lewis. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7(1):57–78, 1995.

20. J. Lin, L. Mamykina, S. Lindtner, G. Delajoux, and H. Strub. Fish'n'steps: Encouraging physical activity with an interactive computer game. In P. Dourish and A. Friday, editors, *UbiComp 2006: Ubiquitous Computing*, volume 4206 of *Lecture Notes in Computer Science*, pages 261–278. Springer Berlin / Heidelberg, 2006.

21. L. Mamykina, E. D. Mynatt, and D. R. Kaufman. Investigating health management practices of individuals with diabetes. In *Proc. ACM CHI 2006*, CHI '06, pages 927–936, New York, NY, USA, 2006. ACM.

22. G. Marcu, J. E. Bardram, and S. Gabrielli. A Framework for Overcoming Challenges in Designing Persuasive Monitoring Systems for Mental Illness. In *Proc. Pervasive Health 2011*, pages 1–10. IEEE Press, 2011.

23. M. Matthews and G. Doherty. In the mood: engaging teenagers in psychotherapy using mobile phones. In *Proc. ACM CHI 2011*, CHI '11, pages 2947–2956, New York, NY, USA, 2011. ACM.

24. M. Morris, Q. Kathawala, T. Leen, E. Gorenstein, F.Guilak, M. Labhard, and W. Deleeuw. Mobile therapy: Case study evaluations of a cell phone application for emotional self-awareness. *Journal of Internet Medical Research*, J Med Internet Res 2010;12(2):e10:12, 2010.

25. C. J. L. Murray and A. D. Lopez. Global mortality, disability, and the contribution of risk factors: Global burden of disease study. *Lancet*, 349:1436–42, 1997.

26. J. Pollak, G. Gay, S. Byrne, E. Wagner, D. Retelny, and L. Humphreys. It's time to eat! using mobile games to promote healthy eating. *Pervasive Computing, IEEE*, 9(3):21 –27, july-sept. 2010.

27. P. Prociow and J. Crowe. Towards personalised ambient monitoring of mental health via mobile technologies. *Technol Health Care*, 18(4-5):275–284, 2010.

28. J. Proudfoot, C. Ryden, B. Everitt, D. A. Shapiro, D. Goldberg, A. Mann, A. Tylee, I. Marks, and J. A. Gray. Clinical efficacy of computerised cognitivebehavioural therapy for anxiety and depression in primary care: randomised controlled trial. *The British Journal of Psychiatry*, 185(1):46–54, 2004.

29. L. Robertson, M. Smith, and D. Tannenbaum. Case management and adherence to an online disease management system. *J Telemed Telecare*, 11 Suppl 2:73–75, 2005.

30. G. S. Sachs. Bipolar mood disorder: Practical strategies for acute and maintenance phase treatment. *J. Clin. Psychopharm.*, 16(2):32–47, Mar 1996.

31. K. A. Siek, K. H. Connelly, Y. Rogers, P. Rohwer, D. Lambert, and J. L. Welch. When do we eat? an evaluation of food items input into an electronic food monitoring application. In *Proc. Pervasive Health 2006*, pages 1 –10, 29 2006-dec. 1 2006.

32. A. A. Stone, S. Shiffman, J. E. Schwartz, J. E. Broderick, and M. R. Hufford. Patient non-compliance with paper diaries. *BMJ*, 324(7347):1193–1194, 5 2002.

33. P. Whybrow, P. Grof, L. Gyulai, N. Rasgon, T. Glenn, and M. Bauer. The electronic assessment of the longitudinal course of bipolar disorder: the chronorecord software. *Pharmacopsychiatry*, 36 Suppl 3:244–249, 2003.

34. World Health Organization. *The global burden of disease: 2004 update*. WHO; 1 edition, 2009.