Brief Report

The longitudinal course of bipolar disorder as revealed through weekly text messaging: a feasibility study

Objectives: To examine the feasibility of collecting course of illness data from patients with bipolar I and II disorder, using weekly text-messaged mood ratings, and to examine the time trajectory of symptom ratings based on this method of self-report.

Methods: A total of 62 patients with bipolar I (n = 47) or II (n = 15) disorder provided mood data in response to weekly cell phone text messages (n = 54) or e-mail prompts (n = 8). Participants provided weekly ratings using the Altman Self-Rating Mania Scale and the Quick Inventory of Depressive Symptoms–Self Report. Patients with bipolar I and II disorder, and men and women, were compared on percentages of time in depressive or manic mood states over up to two years.

Results: Participants provided weekly ratings over an average of 36 (range 1–92) weeks. Compliance with the procedure was 75%. Overall, participants reported depressive symptoms 47.7% of the time compared to 7% of entries reflecting manic symptoms, 8.8% reflecting both depressive and manic symptoms, and 36.5% reflecting euthymic mood. Participants with bipolar I disorder reported more days of depression and were less likely to improve with time than participants with bipolar II disorder. Gender differences observed at the beginning of the study were not observed at follow-up.

Conclusions: The results are similar to those of other longitudinal studies of bipolar disorder that use traditional retrospective, clinician-gathered mood data. Text-message-based symptom monitoring during routine follow-up may be a reliable alternative to in-person interviews.

Jedediah M Bopp[a], David J Miklowitz[b,c], Guy M Goodwin[c], Will Stevens[c], Jennifer M Rendell[c] and John R Geddes[c]

[a]Department of Psychology, University of Colorado, Boulder, CO, [b]Division of Child and Adolescent Psychiatry, University of California at Los Angeles School of Medicine, Los Angeles, CA, USA, [c]Department of Psychiatry, University of Oxford, Oxford, UK

Corresponding author:
David J. Miklowitz, Ph.D.
Division of Child and Adolescent Psychiatry
UCLA Semel Institute
760 Westwood Plasza
Room 58-217
Los Angeles, CA 90024-1759
Fax: +310-206-4446
E-mail: dmiklowitz@mednet.ucla.edu

Our limited understanding of the course of mood symptoms in bipolar disorder impedes the development of effective treatments. A greater understanding of weekly fluctuations in the course of mood symptoms in bipolar disorder could improve relapse prevention by allowing providers to intervene shortly after prodromal symptoms first appear. In the majority of longitudinal studies, illness course and treatment efficacy have been measured using observer-rated scales obtained under artificial conditions of incentive. Furthermore, most investigators collect mood data from patients retrospectively. Asking participants to recall mood episodes many weeks, months, or years in the past risks significant recall biases and can be quite time consuming.

There is a well-recognized need for a reliable method of prospectively recording mood variations and response to treatment. Given that mood monitoring is a key component of the most effective psychosocial treatment strategies for bipolar disorder (1), the development of novel and inexpensive methods to assess mood stability, without requiring elaborate training and reliability protocols, could be of considerable practical benefit to patients and clinicians. More generally, a convergence of outcome measures across studies has the potential to inform the often-lamented efficacy/effectiveness gap in clinical research.

It has been shown that new technologies such as home computers and personal digital assistants (PDAs) can be used effectively to monitor psychiatric symptoms. For example, computer-based, daily emotional and behavior monitoring can be employed when following adults with schizophrenia (2), children aged 7–12 years with attention-deficit hyperactivity disorder and their mothers (3, 4), smokers recording the relationship between smoking and mood (5), individuals with eating disorders tracking mood and binge/purge behavior (6), and patients with borderline personality disorder (7). All of these studies used PDAs, into which participants entered information multiple times over 24-hour periods. Data collection in these studies ranged from a single 24-hour period to 28 days.

There are only a handful of studies that take advantage of computer technology for collection of mood data in bipolar individuals. Schärer et al. (8) adapted the National Institute of Mental Health prospective Life-Chart Form for use on a handheld computer and found that patients preferred the device to paper and pencil charting, reported less social stigma when using the device to record their moods in public, gained knowledge about their disorder, and enjoyed playing a more active role in their treatment. Bauer et al. (9) recruited 80 individuals to use a software system installed on their home computers to record mood, medication, and sleep data. Participants only missed 6.1% of the 114 days for which data were requested, indicating that this method can have high compliance rates. Chinman et al. (10) compared in-clinic, computer-assisted, self-report mood ratings provided by 45 individuals with bipolar disorder to mood ratings made by trained interviewers, and found high correlations between the two sources. Taken together, these findings indicate that patients with bipolar disorder readily adopt technology-assisted symptom-reporting methods, and the resulting mood data may be as reliable as data gathered in person. However, the clinical utility of these methods is limited by their reliance on desktop or handheld computers, and some method of remote data capture.

In this study, we report on our experience using the Oxford University Symptom Monitoring System (SMS), in which patients with bipolar disorder used cell phones to respond via text message to weekly prompts for ratings of manic and depressive symptoms. Our primary objectives were: (i) to evaluate the feasibility of using this system for routine assessment of mood fluctuations during outpatient treatment and (ii) to explore the longitudinal trajectory of mood symptoms as collected using the method. We predicted that, as reported in previous large-scale longitudinal studies (11, 12), patients with bipolar disorder would report more days with depressive symptoms than with manic or hypomanic symptoms or euthymia.

Secondarily, we explored two predictors of the course of bipolar illness as revealed in weekly symptom monitoring—bipolar subtype and gender—to generate hypotheses for future research on factors that moderate illness course under routine treatment conditions. We predicted that weekly self-reporting of mood states would reveal a more treatment-refractory course of illness in bipolar I disorder (BDI) than in bipolar II disorder (BDII), and a more depressive course of illness in women than men.

Materials and methods

Participants

The sample consisted of 62 adult patients (35 women, 27 men, mean age = 34.30, SD = 10.66) with DSM-IV-TR (13) bipolar disorder (BDI, n = 47; BDII, n = 15) treated and followed between December 21, 2006 and September 30, 2008 in the outpatient mood disorders clinic of the Department of Psychiatry, Warneford Hospital, Oxford, UK. The mood disorders clinic is a secondary- and tertiary-care clinic. Patients are referred from primary care by their general practitioners or from secondary-care services for specialist assessment and management.

Patients could be in any clinical state or at any point in their pharmacological treatment, and were in an ongoing treatment relationship with a psychiatrist at the mood disorders clinic. The use of the anonymized, routinely collected data for this report was agreed upon by the Caldicott Guardian of Oxfordshire and Buckinghamshire Mental Health NHS Foundation Trust. The data analytic plan was separately approved by the Human Research Committee at the University of Colorado, Boulder, CO, USA.

Symptom monitoring system

At the Oxford clinical center, a central cell phone connected to a secure desktop computer sent weekly text messages prompting participants to complete ratings on the 5-item Altman Self-Rating Mania (ASRM) scale (14) as well as the 16-item Quick Inventory of Depressive Symptoms–Self Report (QIDS-SR) (15). Rather than receiving the scale questions each time, the participants were supplied with wallet-sized versions of each rating scale. Less than 5% of patients misplaced these cards. In the event that a card was misplaced, patients were able to obtain replacements at the clinic.

Participants responded to the text prompts first with the letter 'A', indicating that they were replying to the ASRM, or with the letter 'Q', indicating that they were replying to the QIDS-SR, and then with a sequence of digits (5 for the ASRM, 16 for the QIDS-SR), with each digit corresponding to an item response. For example, a subject completing the QIDS-SR would receive a text prompt, look at his or her rating card, and reply to the prompt 'Q' with the numerical rating for each of the 16 items (e.g., Q0330200001101111). If the text message contained errors (too few responses, scores out of range, etc.), the system sent a reply requesting that the subject resubmit his or her responses.

If the subject did not reply when first prompted, a reminder message was sent the following day and then again on the third day. If the third message did not prompt a response, the system did not send any more reminders. Patients were able to reply to messages after the second day, but all entries were date-stamped the day they were sent. Patients were not asked to spontaneously submit mood ratings; therefore, each rating was a direct response to prompts from the SMS system. Patients may have also responded to text prompts any number of days after they were sent, but responses were date-stamped when they arrived.

The default system captured one entry every seven days, but clinicians could request more frequent measures for clinical purposes. These entries were stored on a central computer in the form of raw ASRM and QIDS-SR scores for each week, as well as graphically, to display mood fluctuations over time. Clinicians were able to view individual patients' mood fluctuations on a week-to-week basis. Figure 1 illustrates a sample mood chart. Each time point on the mood chart represents the date on which a mood rating was made.

If they preferred, patients could use e-mail and Web form instead of SMS. Patients could not use the service if they did not own or have the ability to borrow a cell phone, if they did not know or were unwilling to learn how to use the text-messaging technology, or could not use e-mail.

The QIDS-SR is a 16-item measure of depression severity which covers the nine DSM-IV-TR symptoms related to the diagnosis of major depressive disorder (13). Participants were asked to rate each symptom on a 0–3 scale over the course of the past seven days. Specifically, the instructions read, 'Choose the one statement ... that best describes the way you have been feeling for the past week'. Depression scores on the QIDS-SR correspond to five levels of severity: none: 0–5; mild: 6–10; moderate: 11–15; severe: 16–20; very severe: 21–27. Scores below 6 typically indicate euthymia. The QIDS-SR has established psychometric properties for rating depressive symptom severity in individuals with chronic major depressive disorder (15) and with bipolar depression (16). It does not purport to distinguish between subsyndromal (e.g., dysthymic) and syndromal (e.g., major depressive) mood states.
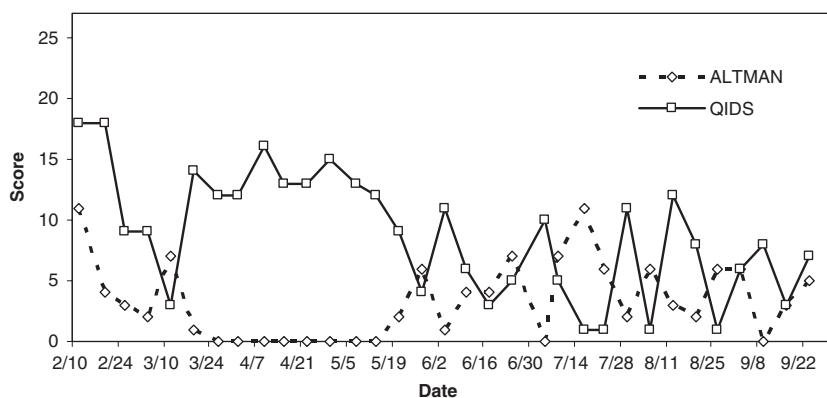

Fig. 1. Sample data for a single subject over seven months. ALTMAN = Altman Self-Rating Mania Scale; QIDS = Quick Inventory of Depressive Symptoms.

Participants rated their mania-related symptoms over the course of the previous week on five individual ASRM scales that range from 0–4. Specifically, patients were instructed to select, for each question, 'the one response to each item that best describes you for the past seven days'. Patients were additionally advised that the term *occasionally* meant once or twice, *often* meant several times or more, and *frequently* meant most of the time. Total ASRM scores range from 0–20, where scores of 6 or greater indicate significant manic or hypomanic symptoms. The ASRM has established psychometric properties for detecting mania in bipolar individuals (14) and for individuals with severe and even psychotic mania (17). Because the items do not measure the duration or functional impairment caused by symptoms, the ASRM does not distinguish between mania and hypomania. ASRM scores were used as continuous measures in this study in order to examine the weekly severity of manic/hypomanic symptoms. The terms *depressive symptoms* and *manic symptoms* are used hereafter and subsume both syndromal and sub-syndromal moods.

Data analyses

For each subject, we calculated the mean proportion of time that QIDS-SR and ASRM scores indicated significant depression, mania or mixed (simultaneous mania and depression) symptoms (as revealed by cutoffs of 6 or greater on each scale), or euthymic mood (both scale scores < 6), using total days in which ratings were supplied as the denominator. Using analyses of variance, we compared the percentage of time patients spent in one of four mood states across the two genders and bipolar subtypes.

To examine the relations between bipolar subtype and gender to changes in ASRM and QIDS scores, we created multilevel models using the PROC MIXED function in SAS® (18). For both the ASRM and QIDS variables, separate multilevel models were estimated. We treated ASRM and QIDS scores as continuous variables to examine degree of change over time. Because participants often did not make entries exactly seven days apart, the time variable ('day') was calculated as the number of days after the participant responded to the first text message. For the first level of the multilevel model, we estimated the mean ASRM (or QIDS) score: $ASRM_{ti} = \beta_{0i} + \epsilon_{ii}$, where $ASRM_{ti}$ is a subject's mania score at time t. For the second level, we created dichotomous variables for bipolar subtype (type I versus type II), $\alpha Diag$; and gender, $\alpha Gender$. Next, we included variables for the interactions between subtype and day, gender and day, and subtype and gender:
$$\beta_{0i} = \alpha_0 + \alpha_i Diag + \alpha_i Gender + \alpha_i Day + \alpha_i Diag \times \alpha_i Day + \alpha_i Day \times \alpha_i Gender + \alpha_i Diag \times \alpha_i Gender \; \epsilon_i.$$

Results

Compliance with the SMS procedure

Of the 62 patients, 54 (87%) opted to use the cell phone based SMS ratings and 8 (13%) opted to use e-mail. Fewer than 10% of the patients attending the clinic were unable to use the follow-up service due to being unable to use either cell phone texting or e-mail/Web. Patients were followed for an average of 36 weeks or 252 days (SD = 23.5, range: 1–93 weeks). This range reflects the fact that the study used rolling enrollment, which resulted in subjects with a wide range of time in the study. They occasionally completed multiple entries per week because the system accepted responses to texts at any time. Because participants varied in the number of weeks they were enrolled in the study, compliance was calculated by dividing the number of days the subject responded to mood rating requests by the total number of weeks the subject was enrolled in the study. The mean (SD) compliance proportion was 0.75 (SD = 0.24, range: 0.10–1.0), indicating a high level of consistency. In order to determine if compliance rates changed over the course of the study we calculated compliance at three points during the study. For subjects (n = 35) who were enrolled in the study for at least the mean number of weeks (36 weeks), we calculated compliance for weeks 1 through 12, weeks 13 through 24, and weeks 25 through 36. For these subjects, compliance for the first third of the study was 84%; for the second interval, 75%; and for the third interval, 73%.

Out of the 5,194 text-message responses received during the course of the study, there were 182 incorrectly formatted responses (or 3.5%); 4,311 (83%) of the text-message responses were received within 12 hours of the SMS prompt.

Depressive and manic symptoms over time

Across time and across bipolar subtype, participants reported depressive symptoms during 47.7% of the weeks compared to 7.0% with manic symptoms alone. Of the weekly entries, 8.8% reflected both depressive and manic symptoms. Participants were euthymic 36.5% of the time (Table 1).

Across time, participants with BDI reported depressive symptoms 53.6% of the time compared

Table 1. Mean percent time spent in mood states as reported by subtype and gender

Symptoms	Time in mood states across all subjects (n = 62)			
Mixed	8.77			
Manic	7.00			
Depressive	47.69			
Euthymic	36.50			
	Time in mood states by subtype			
	BDI (n = 47)	BDII (n = 15)	F	df	p-value
Mixed	8.15	8.62	0.019	1,61	0.89
Manic	7.06	6.00	0.122	1,61	0.73
Depressive	53.60	35.23	4.98	1,61	0.03
Euthymic	31.18	50.15	3.98	1,61	0.05
	Time in mood states by gender			
	Males (n = 27)	Females (n = 35)	F	df	p-value
Mixed	9.66	7.47	0.66	1,61	0.52
Manic	7.07	6.90	0.297	1,61	0.74
Depressive	45.99	53.79	1.61	1,61	0.21
Euthymic	37.26	31.84	1.79	1,61	0.18

BDI = bipolar I disorder; BDII = bipolar II disorder.

to 7.1% of weeks with manic symptoms alone, and 8.2% with both depressive and manic symptoms. Participants with BDI reported euthymia 31.2% of the time. Participants with BDII reported depressive symptoms 35.2% of the time compared to 6.0% with manic symptoms and 8.6% with both depressive and manic symptoms. Participants with BDII reported euthymic mood during 50.2% of the weeks. BDI patients spent more time in depressive states than BDII patients $[F(1,61) = 4.98, p = 0.029]$, whereas BDII patients spent somewhat more time euthymic than BDI patients $[F(1,61) = 3.98, p = 0.051]$ (Table 1). There were no differences between male and female participants in the proportion of weeks spent depressed (46.0% versus 53.8%), manic (7.1% versus 6.9%), mixed (9.7% versus 6.9%), or euthymic (37.3% versus 31.8%) (for all, p > 0.10; Table 1).

Hierarchical Linear Models indicated a significant interaction between BDI/BDII subtype and time on QIDS scores $[F(1,2849) = 6.55, p < 0.0005]$. Patients with BDI showed little change in QIDS scores over the duration of the reporting period (Fig. 2). In contrast, patients with BDII (n = 15) reported higher QIDS scores than patients with BDI (n = 47) during the first half of the reporting period, but lower QIDS scores during the second half of the period. Least squares mean QIDS ratings at the end of follow-up were 6.6 for patients with BDII compared to 8.3 for patients with BDI. These scores are within the 'mild' (scores between 6 and 10) range on the QIDS scale. The bipolar subtype × time interaction on mania symptoms was not significant $[F(1,3137) = 1.57, p = 0.21]$.

There was a significant gender × time interaction for QIDS scores $[F(1,2849) = 7.9, p < 0.05]$. Men initially reported higher mean depression scores than women (10.1 versus 7.9). Over time, these ratings decreased for men. In contrast, women reported a stable level of depression throughout the study, such that by the end of data collection, scores for women and men were nearly identical (5.52 for men versus 4.71 for women). There was also a significant interaction between gender and time on ratings of mania $[F(1,2882) = 11.19, p < 0.001]$. Women initially reported higher ASRM scores compared to men (least squared means: 6.5 versus 3.9), a difference which became less pronounced with time.

Discussion

We examined the utility of a cell phone text-messaging system to monitor the outcomes of 62
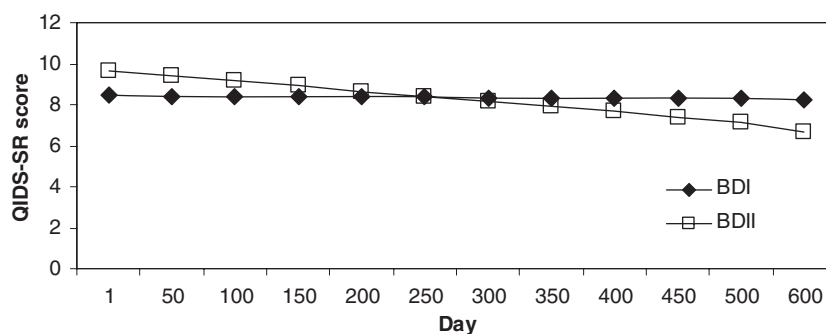

Fig. 2. Longitudinal course of Quick Inventory of Depressive Symptoms–Self Report (QIDS-SR) scores among bipolar I (BDI, n = 47) and bipolar II (BDII, n = 15) patients. Bipolar subtype × time interaction, $F(1,2849) = 6.55, p < 0.0005$.

outpatients with BDI and BDII. The prospective nature of data collection allowed us to examine the proportion of time patients spent in states of mania/hypomania, depression, mixed state, or euthymia, and time trends in the trajectory of symptoms. The compliance rate for adherence to the text-messaging protocol was 75% over an average of 36 weeks, with 83% of responses obtained within 12 hours of the prompting message. Fewer than 4% of the text messages sent by patients contained incorrectly formatted responses, suggesting that patients had little difficulty understanding the SMS requirements. These results suggest that this novel mood monitoring method is readily adopted by patients with BDI and BDII.

The proportions of time that patients reported spending in depressive, manic, and euthymic states were similar to the proportions reported in other longitudinal studies of bipolar patients (11, 12, 19, 20). Patients reported that a larger percent of their time (47.7%) was spent in depressed states, whereas only 36.5% was spent in euthymic mood. To test the validity of these weekly mood ratings, we would need to compare the current results to those obtained from clinician-gathered mood ratings, which were not available in the present study.

Depression was a more prominent feature of BDI than BDII. Participants with BDI reported depressive symptoms 53.6% of the time, whereas participants with BDII reported depressive symptoms 35.2% of the time. Patients with BDI reported relatively stable levels of minor depression, whereas patients with BDII initially reported higher levels of depression that decreased to below the levels of depression reported by patients with BDI. Possibly, depression among patients with BDI in this sample was less treatment responsive than depression among patients with BDII.

Differences in the proportion of time spent in depressive and manic states between BDI and BDII patients have varied from study to study. Kupka et al. (12) and Joffe et al. (19) found no significant differences between patients with BDI and patients with BDII on proportion of time spent with depressive symptoms. In contrast, Judd et al. (11) found that patients with BDI spent 30.6% of weeks with depressive symptoms, whereas patients with BDII spent 51.9% of weeks with depressive symptoms. In an 18-month study in Finland (21), patients with BDII spent 58.0% of their time in states of depression compared to 41.7% among BDI patients. The variable results across studies may reflect different assessment methods and differences in the length of the retrospective or prospective reporting intervals. Weekly prospective monitoring may expose more

subtle but pervasive subsyndromal depressive symptoms among BDI patients that are less apparent among BDII patients.

Although men began with higher initial depression scores than women, the two genders were nearly identical in levels of depression by the end of the study. In contrast, mania scores among women were initially higher than mania scores among men, but decreased over time to levels below the manic or hypomanic threshold (6 or above). In contrast, in a three-year follow-up of 19 men and 37 women with bipolar disorder, Christensen et al. (22) found that women had more depressive episodes than men, and men had more manic episodes than women. Future studies should examine whether gender differences remain robust once differences in age at onset, rapid cycling status, and pharmacological treatments are statistically controlled.

The group differences in depression scores at the end of the study (mean 6.6 ± 7.8 on the QIDS-SR for patients with BDII, compared to 8.3 ± 7.3 for patients with BDI), although statistically significant, may not be clinically significant given that both scores fell within the QIDS range reflecting 'moderate depression'. Minor week-to-week fluctuations in QIDS or ASRM scores will not always signal the need for modifications of the treatment regimen, unless such minor changes have historically heralded the development of fully syndromal recurrences in that patient.

Perhaps the most significant barrier to conclusive interpretation of these results is lack of standardization of the pharmacological and psychological treatments. Patients were followed naturalistically, and could be enrolled in any treatment modality. As a consequence, the course patterns we observed cannot be exclusively attributed to the effects of gender or BDI/BDII status. It is also possible that the present results simply reflect differences in reporting behavior between patients with BDI and BDII, or between men and women. Future research should explore moderators of illness course as measured using the SMS, such as comorbid diagnoses, treatment regimen or modality, illness severity, or age at onset.

The findings are also limited by the choice of the ASRM, which is designed as a categorical measure of mania and has not been validated as a continuous measure. There are no brief, self-report mania scales that differentiate between mania and hypomania. Neither the ASRM nor the QIDS is intended as a diagnostic measure; hence, scores above 6 should not necessarily be construed as meeting DSM-IV criteria for a manic, hypomanic, or major depressive episode. Because both meas-

ures had identical cutoff scores, it was possible to consider each measure categorically as a guide to the presence or absence of mania or depression. However, each measure is scaled in a different metric, so it was not possible to determine the temporal relationships between fluctuations between the two scales or to measure patients' cycling between the two mood states.

Whereas data were collected on a weekly basis, the SMS monitoring technique does not entirely solve the problem of retrospective recall. Patients were prompted to describe their mood over the past seven days. A truly contemporaneous measure would ask patients to enter their mood at the very moment they are prompted. Whereas ecological momentary assessment techniques (6) may reduce recall bias, they may introduce other problems. Notably, prompting frequently enough to establish daily or even within-day fluctuations could feel laborious or intrusive to some patients.

The SMS methodology is best suited to studies of illness course and, potentially, treatment outcome, particularly if its validity relative to face-to-face measures of symptom course can be established. In the future, a parallel effort should be undertaken to improve the SMS technology for clinical purposes. In addition to alerting treatment providers to the patient's mood shifts in 'real time', it may be possible to customize each system for each patient, allowing for the monitoring of variables such as hours slept, quality of sleep, levels of stress, medication adherence, exercise, and food intake. Correlations within each individual between these variables and mood change could then be determined.

The finding that the SMS was readily adopted by a clinical population and resulted in findings similar to those from prior longitudinal studies suggests that this technology could be developed both as a research and everyday clinical tool. The convergence of outcome measures will make much more feasible the comparison of data from clinical trials with data from ordinary practice. In bipolar disorder, short-term mood instability is pervasive and could be a much more sensitive measure of outcome than, for example, the frequency of syndromal illness episodes. The technology also appears to be empowering for its users and has the potential to reduce the need for more frequent face-to-face interviewing. The resource implications in terms of clinical time and travel could be considerable. We predict that text-messaging technology will have an important future in the clinical care of patients with bipolar disorder.

Acknowledgements

References

1. Miklowitz DJ, Goodwin GM, Bauer MS, Geddes JR. Common and specific elements of psychosocial treatments for bipolar disorder: a survey of clinicians participating in randomized trials. J Psychiatr Pract 2008; 14: 77–85.
2. Kimhy D, Delespaul P, Corcoran C, Ahn H, Yale S, Malaspina D. Computerized experience sampling method (ESMc): assessing feasibility and validity among individuals with schizophrenia. J Psychiatr Res 2006; 40: 221–230.
3. Whalen CK, Henker B, Ishikawa SS et al. An electronic diary study of contextual triggers and ADHD: get ready, get set, get mad. J Am Acad Child Adol Psychiatry 2006; 45: 166–174.
4. Whalen CK, Henker B, Jamner LD et al. Toward mapping daily challenges of living with ADHD: maternal and child perspectives using electronic diaries. J Abnorm Child Psychol 2006; 34: 115–130.
5. Delfino RJ, Jamner LD, Whalen CK. Temporal analysis of the relationship of smoking behavior and urges to mood states in men versus women. Nicotine Tobacco Res 2001; 3: 235–248.
6. Smyth JM, Wonderlich SA, Heron KE et al. Daily and momentary mood and stress are associated with binge eating and vomiting in bulimia nervosa patients in the natural environment. J Consult Clin Psychol 2007; 75: 629–638.
7. Trull TJ, Solhan MB, Tragesser SL et al. Affective instability: measuring a core feature of borderline personality disorder with ecological momentary assessment. J Abnorm Psychol 2008; 117: 647–661.
8. Schärer LO, Hartweg V, Valerius G et al. Life charts on a palmtop computer: first results of a feasibility study with an electronic diary for bipolar patients. Bipolar Disord 2002; 4: 107–108.
9. Bauer M, Grof P, Rasgon N, Gyulai L, Glenn T, Whybrow P. New computer based tool for the longitudinal study of bipolar disorder. Aspects Affect 2005; 1: 101–108.
10. Chinman M, Young AS, Schell T, Hassell J, Mintz J. Computer-assisted self-assessment in persons with severe mental illness. J Clin Psychiatry 2004; 65: 1343–1351.
11. Judd LL, Schettler PJ, Akiskal HS et al. Long-term symptomatic status of bipolar I vs. bipolar II disorders. Intl J Neuropsychopharmacol 2003; 6: 127–137.
12. Kupka RW, Altshuler LL, Nolen WA et al. Three times more days depressed than manic or hypomanic in both bipolar I and bipolar II disorder. Bipolar Disord 2007; 9: 531–535.
13. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision. Arlington: American Psychiatric Association, 2000.
14. Altman EG, Hedeker D, Peterson JL, Davis JM. The Altman Self-Rating Mania Scale. Biol Psychiatry 1997; 42: 948–955.
15. Rush AJ, Trivedi MH, Ibrahim HM et al. The 16-item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry 2003; 54: 573–583.

16. Trivedi MH, Rush AJ, Ibrahim HM et al. The inventory of depressive symptomatology, clinician rating (IDS-C) and self-report (IDS-SR), and the quick inventory of depressive symptomatology, clinician rating (QIDS-C) and self-report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. Psychol Med 2004; 34: 73–82.
17. Altman E. Rating scales for mania: is self-rating reliable? J Affective Disord 1998; 50: 283–286.
18. Ger D, Everitt BS. Handbook of Statistical Analyses Using SAS, 2nd Edition. London: CRC Press, 2001.
19. Joffe RT, MacQueen GM, Marriott M, Young LT. A prospective, longitudinal study of percentage of time spent ill in patients with bipolar I or bipolar II disorders. Bipolar Disord 2004; 6: 62–66.
20. Judd LL, Akiskal HS, Schettler PJ et al. A prospective investigation of the natural history of the long-term weekly symptomatic status of bipolar II disorder. Arch Gen Psychiatry 2003; 60: 261–269.
21. Mantere O, Suominen K, Valtonen HM et al. Differences in outcome of DSM-IV bipolar I and II disorders. Bipolar Disord 2008; 10: 413–425.
22. Christensen EM, Gjerris A, Larsen JK et al. Life events and onset of a new phase in bipolar affective disorder. Bipolar Disord 2003; 5: 356–361.