Published in final edited form as:

Bipolar Disord. 2014 November ; 16(7): 669–677. doi:10.1111/bdi.12218.

The prediction of study-emergent suicidal ideation in bipolar disorder: a pilot study using ecological momentary assessment data

Wesley K Thompson^{a,b}, Anda Gershon^{c,d}, Ruth O'Hara^{c,d}, Rebecca A Bernert^{c,d}, and Colin A Depp^{a,b,e}

^aStein Center for Research on Aging, University of California, San Diego School of Medicine, La Jolla, CA, USA

^bDepartment of Psychiatry, University of California, San Diego, La Jolla, CA, USA

^cDepartment of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford University, Stanford, CA, USA

^dVA Sierra-Pacific Mental Illness Education and Clinical Center, Veterans Affairs Health Care System, Palo Alto, CA, USA

^eSan Diego Veterans Affairs Healthcare System, La Jolla, CA, USA

Abstract

Objectives—Bipolar disorder is associated with idiosyncratic precursors of clinically important states such as suicidal ideation. Ecological momentary assessment (EMA)—high frequency data collection in a subject's usual environment—provides the potential for development of temporal, individualized prediction of risk states. The present study tested the ability of EMA data to predict individual symptom change in clinician-rated suicidal ideation.

Methods—Thirty-five adults diagnosed with inter-episode bipolar disorder completed daily measures of affect in their home environments using diaries administered over an eight-week assessment timeline. Suicidal ideation was assessed monthly at in-person visits using the Inventory of Depressive Symptomatology–Clinician Rated. We used a novel application of *functional linear models* (FLMs) to generate prospective predictions of suicidal ideation at in-person clinician assessments based on intensively sampled trajectories of daily affect.

Results—Eight instances of suicidal ideation scores > 0 were recorded during the study period on six participants. Utilizing trajectories of negative and positive affect, cross-validated predictions attained 88% sensitivity with 95% specificity for elevated suicidal ideation one week prior to in-person clinician assessment. This model strongly outperformed prediction models using cross-sectional data obtained at study visits alone.

Corresponding author: Colin A. Depp, Ph.D. Stein Institute for Research on Aging Department of Psychiatry (0664) University of California, San Diego 9500 Gilman Drive La Jolla, CA 92093-0664, USA Fax: 858-534-5475 cdepp@ucsd.edu.

Disclosures The authors of this paper do not have any commercial associations that might pose a conflict of interest in connection with this manuscript.

Conclusions—Utilizing EMA data with FLM prediction models substantially increases the accuracy of prediction of study-emergent suicidal ideation. Prediction algorithms employing intensively sampled longitudinal EMA data could sensitively detect warning signs of suicidal ideation to facilitate improved suicide risk assessment and the timely delivery of preventative interventions.

Keywords

bipolar disorder; ecological momentary assessment; functional linear models; intensively sampled longitudinal data; mobile technology; prediction; suicidal ideation

High frequency data collection in a participant's usual environment [ecological momentary assessment (EMA)] offers a promising approach for understanding, and potentially intervening on, precursors of clinically important states and behaviors in psychiatric disorders such as suicide (1–3). An estimated 25–50% of those with bipolar disorder will attempt suicide in their lifetime (4), and evidence suggests that suicide attempts of individuals with bipolar disorder are also associated with greater lethality or suicide potential compared to unipolar depression (5).

Suicidal ideation is a reliable risk factor for death by suicide and suicide attempt incidence across psychiatric disorders (6). However, the temporal course of precursors to suicidal ideation in individuals with bipolar disorder is poorly understood (7). In particular, it is unclear what symptoms or states may place persons at elevated risk for suicidal ideation and the earliest that meaningful changes in these precursors may be detected. Although some surges in suicidal ideation may be impulsive and without precedent, it is likely that, for many patients with bipolar disorder, changes in mood states or comorbidities predict later emergence of suicidal ideation (8). A goal of some recent studies employing EMA data collection methodology is to intensively assess changes in potential risk factors in psychiatric disorders to determine when and how accurately they predict suicidal ideation (9, 10).

Statistical prediction models are well suited for application to intensively sampled risk factors obtained from EMA data (11–13) because of their utility in discovering patterns of temporal variation antecedent to and predictive of clinical outcomes. An important criterion for evaluating EMA-based prediction models is robust sensitivity and specificity of predictions across heterogeneous samples. Moreover, prediction models should ideally identify risk states (e.g., mood changes) in advance of the observed outcome (e.g., suicidal ideation) for the application of timely preventative interventions. The greater temporal resolution of EMA data and associated prediction models should also provide for greater accuracy and/or earlier detection than would data obtainable from routine intervals, such as in clinic visits or typical longitudinal clinical research studies.

There have been a number of studies utilizing EMA data for the prediction of suicidal behaviors over the short term (9, 14–16); however, none have examined prediction of risk in individuals with bipolar disorder. Prediction of suicide risk in patients with bipolar disorder may be particularly challenging; long-term prospective studies indicate substantial heterogeneity in both symptoms (17, 18) and precursors to exacerbations (19, 20). However,

negative affect appears to be a common precursor to symptom worsening in mood disorders (21), providing a rationale for using EMA negative affect ratings in statistical prediction models to detect early signs of suicide risk within bipolar disorder specifically.

In the present report, we evaluated EMA-derived negative affect from a pilot study of individuals with inter-episode bipolar disorder across an eight-week assessment period. Applying statistical prediction models, we sought to use EMA data to prospectively predict study-emergent suicidal ideation prior to in-person assessments. Specifically, we employed *functional linear models* (FLMs) (22, 23) which utilize intensively sampled longitudinal data to generate empirically derived, updatable predictions. We hypothesized that use of individual daily affect ratings as predictors in FLMs would provide more accurate individualized predictions of suicidal ideation compared to data from more infrequently scheduled in-person assessments alone.

Materials and methods

Sample and procedures

Data were obtained from a parent study assessing daily affect and sleep in bipolar disorder relative to healthy controls (24). The sample presented here consisted of 35 adults, aged 18–64 years, diagnosed with bipolar disorder. Participants were recruited from the community via online advertisements and flyers posted in the San Francisco Bay Area. Interested individuals responding to recruitment postings completed a preliminary screening interview over the telephone. Callers appearing to be potentially eligible for the study were invited to the laboratory for an in-person clinical assessment. At this initial visit, participants were interviewed using the Structured Clinical Interview for DSM-IV-TR (SCID) (25) and assessed for severity of mood symptoms using the Young Mania Rating Scale (YMRS) (26) and the Inventory of Depressive Symptomatology–Clinician Rated (IDS-C) (27). To be included in the study, participants were required to meet diagnostic criteria for bipolar disorder type I or II according to the DMS-IV-TR (SCID), and to remain inter-episode throughout the eight-week study period. Inter-episode status was defined as: (i) the absence of depressive or (hypo)manic episode (SCID based), (ii) a score of < 12 on the YMRS, and (iii) a score of < 24 on the IDS-C. Participants were also required to have no suspected diagnosis of substance or alcohol abuse disorder in the six months preceding the baseline visit.

Eligible subjects were asked to complete eight consecutive weeks of daily diaries. Each evening, prior to bedtime, participants completed the Positive and Negative Affect Scale (PANAS) (28) questionnaire. To ensure timely completion of the diaries, participants were asked to call a voicemail box when they completed the evening diary. Participants who missed three consecutive calls were contacted by study staff and encouraged to resume calling. Eighty-eight percent of participant calls were completed as requested.

Participants returned for two additional in-person laboratory visits at four and eight weeks following their initial visit. During each visit, the SCID mood module, the IDS-C, and the YMRS were re-administered to confirm continued inter-episode status. Two participants relapsed during the assessment period. Our focus on the inter-episode period restricted

participant eligibility to individuals maintaining low levels of symptoms throughout the course of the prospective assessment period. The two participants who relapsed were paused from participation, assessed for safety, provided monetary compensation for their time, and with permission, re-contacted for reassessment. Both participants agreed to re-enroll in the study upon achieving inter-episode status, and both subsequently completed eight weeks of daily diary. Given the additional laboratory visit and four weeks of diary data collected, a total of 87 days of affect diaries as well as three two post-baseline in-person suicidal ideation assessments were available for both cases. All study procedures were approved by the University of California's Committee for the Protection of Human Subjects and all participants provided written informed consent to participate in this study.

Measures

Psychiatric diagnosis—The SCID and the Duke Structured Interview for Sleep Disorders (DSISD) (29) were used to assess psychiatric and sleep disorder history. The SCID is a semi-structured interview designed to assess DSM-IV-TR diagnostic criteria for Axis I disorders, and has been shown to have good reliability (30, 31). A post-doctoral fellow or a trained psychology doctoral candidate administered the SCID to all participants to assess current and lifetime Axis I disorders. The DSISD is a semi-structured interview designed to ascertain Research Diagnostic Criteria for sleep disorder diagnoses, consistent with DSM-IV-TR and the International Classification of Sleep Disorders (ICSD) nosologies. The DSISD has good reliability and validity (29).

Suicidal ideation—Suicidal ideation was assessed using Item 18 of the Inventory for Depressive Symptoms—Clinician Version (IDS-C), a widely-used 30-item instrument, designed to assess depressive symptoms. Item 18 consists of the following questions: (i) 'In the past week, have you felt that life was not worth living?'; (ii) 'Do you have thoughts of death or suicide?'; (iii) 'How often do these thoughts come?'; (iv) 'How long do they stay?'; (v) 'What have you thought about?'; (vi) 'Have you thought of a plan in the last week?'; and (vii) 'Have you done anything to hurt yourself?'. Ratings are made on a 0–3 scale (e.g., 0 = *does not think of suicide or death* and 3 = *thinks of suicide/death several times a day in depth, or has made specific plans, or attempted suicide*).

Daily affect ratings—The PANAS (28) is a 20-item self-report scale widely used to assess positive and negative dimensions of emotional experience. It consists of ten positive adjectives (attentive, interested, alert, excited, enthusiastic, inspired, proud, determined, strong, active) and ten negative adjectives (distressed, upset, hostile, irritable, scared, afraid, ashamed, guilty, nervous, jittery). Each adjective is rated using a 5-point scale (1 = *very slightly or not at all* and 5 = *extremely*). Scores range from 5 to 50 for each of the two scales. This measure has good psychometric properties, and has been used extensively in mood research including studies with bipolar samples [e.g., (32, 33)].

Statistical analyses

Two important and related aims of statistical prediction using EMA data are (i) determining temporal relationships among multiple, concurrently, and repeatedly assessed variables; and (ii) maximizing prediction accuracy of individual subject clinical outcomes. Aim (i) is useful

for understanding the mechanisms of fluctuation in symptoms over time for a given subject. If changes in one variable predict subsequent changes in another variable, this provides some indication of a mechanistic relationship between the two variables (13, 34). Aim (ii), the primary focus of the current study, is useful for prognosis and hence for delivering relevant interventions in a timely fashion. Predictions must be of acceptable accuracy and with sufficient lead-in time to enable preventative interventions. For accuracy, a balanced trade-off in sensitivity (true positives) and specificity (true negatives) is generally desirable, although a higher level of sensitivity may be required for more severe outcomes such as a high level of suicidal ideation.

One potential approach to statistical prediction utilizing EMA data is to apply standard least squares multiple linear regression, using intensively sampled risk factors as predictors in regression models predicting clinical outcomes (35). However, a major problem with standard least squares regression is that the number of potential predictors (prior daily assessments) can be quite large compared to the number of outcomes. In this setting, standard multiple linear regression models can perform poorly when predicting outcomes in *de novo* samples, due to over-fitting [e.g., (36)]. This difficulty can be addressed by employing penalized least square regression methods such as FLMs (37–39). FLMs utilize intensively sampled longitudinal trajectories in regression models predicting outcomes. The central assumption of FLMs is that, although EMA data are sampled frequently, the regression weights relating prior daily assessments to outcomes vary smoothly as a function of time. The degree of smoothness is empirically determined to maximize cross-validated prediction accuracy, as detailed below. This results in an effective dimensionality that may be much less than the number of days of data used to predict outcomes. To determine if *volatility* (non-smooth variation) in EMA-derived trajectories also contributes predictive power, we included trajectory variance as an additional predictive measure in the FLMs.

To be specific, let $X_i(d)$ be the EMA assessment for the i th subject on day d , where d is the number of days until the next in-person clinician assessment of suicidal ideation. In our example, we employ $X_i(d)$ evaluated over two weeks: from three weeks ($d = -21$) to one week ($d = -7$) prior to the in-person assessment ($d = 0$). The FLM for suicidal ideation, here denoted as y_i , is given by:

$$y_i = \beta_0 + \int_{-21}^{-7} X_i(d) \beta_1(d) \delta d + \varepsilon_i$$

where $\beta_1(d)$ is a regression weight function and the ε_i are sampling errors. The regression weights $\beta_1(d)$ are assumed to vary smoothly as a function of time (d), so that its effective degrees of freedom is much less than 14 (the number of days) (40). The smoothness requirement limits the effective degrees of freedom and thereby avoids over-fitting the data, improving prediction accuracy in *de novo* samples.

The number and timing of measurements may differ across subjects either by design or because of missing data, both of which can be addressed directly by FLMs (41). However, for simplicity we impute missing data in a preliminary step by computing the mean value for a given predictor for each day in the two-week period ($d = -21$ to $d = -7$), conditional on the

number of missing data points the subject had during this period. In general, high levels of missing data are likely to lead to less accurate predictions; however, if missingness is related to the outcome, the number or pattern of missing observations for a given trajectory can itself be entered into the FLM as an additional predictor to improve accuracy.

We calculated accuracy (sensitivity and specificity) of FLMs to predict a rating > 0 on the IDS-C suicidal ideation item as follows. To compute the predicted level of suicidal ideation for a given subject, that subject's data were omitted from the dataset. An FLM was fit the data from the remaining subjects, and suicidal ideation scores were predicted from this model based on his/her relevant EMA trajectories. This leave-one-subject-out, cross-validation technique ensures unbiased estimates of predicted scores in *de novo* samples. These predicted scores were then thresholded on a range of cut-offs to produce receiver operating characteristic (ROC) curves.

We compared the prediction accuracy of FLMs using EMA-derived data to the accuracy of predictions based on in-person data alone. Specifically, we used variables collected on the same date as the previous in-person visit (in usual linear regression models) to prospectively predict suicidal ideation at the following in-person visit four weeks later. We also included the IDS-C suicidal ideation score from the previous visit as a predictor in these comparative models. We then compared the accuracy of the prediction model applied to data from in-person visits alone with that achieved utilizing FLMs in conjunction with EMA data.

Results

Sample characteristics

Table 1 presents demographic characteristics, illness history, and medication use for all participants. Thirty-two participants were diagnosed as having bipolar I disorder and three as having bipolar II disorder.

IDS-C suicidal symptom endorsement—At baseline: seven of the 35 subjects were rated as 1 ('Feels life is empty or is not worth living') and one participant was rated as 2 ('Thinks of suicide/death several times a week for several minutes'). The remaining participants ($n = 27$) were rated as 0 ('Does not think of suicide or death'). Post-baseline: there were eight instances rated with a score > 0 , across six participants, only two of whom had experienced suicidal ideation at baseline. The remaining ($n = 29$) were rated as 0.

Application of FLMs to prediction of suicidal ideation

We chose two potential predictors from the EMA data: daily self-ratings of PANAS positive affect and negative affect total scores. Mean trajectories for Weeks 3 through Week 1 preceding in-person clinician assessments for all three variables (standardized to have zero mean and unit standard deviation) are displayed in Figure 1. The mean of trajectories in Figure 1 ending in IPS-C suicidal ideation scores > 0 are elevated for negative affect compared to trajectories ending in suicidal ideation scores $= 0$, especially near the end of this two-week period. Positive affect is generally less well differentiated between the two types of trajectories.

We fit three FLMs. *Model 1* utilized PANAS positive affect scores as a predictor. *Model 2* used PANAS negative affect scores. *Model 3* used both positive and negative affect scores. For each FLM, we computed one-week-ahead predictions of IDS-C Item 18 suicidal ideation scores, using leave-one-subject-out cross-validation.

Cross-validated predictions and corresponding ROC curves (obtained by thresholding predicted scores over a range of cutoffs) for each model are displayed in Figure 2. PANAS negative affect strongly outperforms positive affect as a predictor of suicidal ideation. For example, to achieve an 88% true positive rate (7/8 correct predictions), Model 1 (negative affect alone) achieves a maximum of 87% specificity (52/60 true negatives); whereas Model 2 (positive affect alone) achieves a maximum of only 35% specificity (21/60 true negatives). Model 3 (both negative and positive affect) performed somewhat better than Model 1, with 95% specificity for 88% sensitivity. ROC curves validate the superior performance of negative over positive affect as a predictor of suicidal ideation scores over the range of possible cutoffs. As an additional model comparison metric, area under the curve (AUC) was computed; AUC = 0.85 for Model 1, AUC = 0.73 for Model 2, and AUC = 0.88 for Model 3. Finally, squared Spearman correlations of predicted versus actual IDC-S suicidal ideations scores for the three models were 0.15, 0.06, and 0.22 for Models 1, 2, and 3, respectively, also indicating that the most accurate predicted scores were from Model 3.

The beta weight functions for Model 3, displayed in Figure 3, possess roughly 5.8 effective degrees of freedom.

To further evaluate the utility of using EMA data to make predictions of suicidal ideation, we compared the results of Model 3 with a linear model employing data points collected only on days of in-person assessments. Specifically, we used the same variables as in Model 3 (PANAS negative and positive affect and their interaction) from the day of the previous in-person visit to predict IDS-C-assessed suicidal ideation at the subsequent visit. We also included the IDS-C suicidal ideation score from the previous visit as a predictor in the linear regression (total model degrees of freedom = 4). Both negative affect ($p < 0.001$) and suicidal ideation (IDS-C Item 18) at the previous visit ($p < 0.0001$) were highly significant predictors of suicidal ideation at the subsequent visit, as was the interaction of negative and positive affect ($p = 0.01$). However, cross-validated prediction accuracy was notably poorer compared to Model 2 (e.g., 88% sensitivity, with a true negative rate of only 35%). Additionally, AUC = 0.79, and squared Spearman correlation was 0.11 for this model, both indicating worse predictive performance than Models 2 and 3.

Using Model 3, we also computed alternative two-week ahead cross-validated predictions of IDS-C suicidal ideation based on EMA data from Week 4 to Week 2 prior to visits, achieving 87% specificity (52 out of 60) for 88% sensitivity. Thus, the two-week ahead predictive performance of Model 1 is nearly as good as its one-week-ahead performance. Model 3 achieved 57%, 80%, and 93% specificity for 88% sensitivity using one week's worth of EMA data from Weeks 4 to 3, Weeks 3 to 2, and Weeks 2 to 1 prior to the subsequent in-person visit, respectively. Thus, predictions become more accurate as they utilized more days of EMA data and/or time points in closer proximity to the day of the in-person assessment of suicidal ideation.

Finally, we examined whether adding a measures of volatility (trajectory variance) added predictive power to FLMS. We re-ran Model 3 adding in negative and positive affect trajectory variance. The predictive performance of this model was very similar to that of Model 3 (AUC = 0.93, squared Spearman correlation = 0.23).

Discussion

We found that EMA data in conjunction with functional linear models were able to predict future occurrence of clinician-assessed suicidal ideation with high sensitivity and specificity. For one-week-ahead predictions (i.e., prior to clinician assessment), the combination of daily negative affect predicted seven of eight instances where IDS-C suicidal ideation scores were rated as greater than zero (revealing only 3/60 false positives). For two-week-ahead predictions, the same combination revealed 87% specificity at 88% sensitivity. This level of accuracy was obtained robustly, using cross-validated predictions. The accuracy of predictions using EMA assessments was greater compared to those using in-person data alone, with an increase in model complexity of only approximately two degrees of freedom.

These results are consistent with the literature on other psychiatric disorders assessed with EMA implicating a relationship between negative affect and increased risk for suicidal behaviors (9, 14–16). Our findings extend prior work on prospective changes in suicidal ideation in bipolar disorder by showing that temporal associations between negative affect several weeks before increases in suicidal ideation were detected during in-person clinician assessments, suggesting that level negative affect may be considered a proximal warning sign of future suicidal ideation in bipolar disorder. This finding is underscored by the fact that most subjects who endorsed elevated symptoms were not experiencing suicidal ideation at baseline. Although we are unable to examine whether suicidal ideation was present prior to when it was assessed by clinicians at study visits (i.e., EMA protocol did not include questions about suicidal symptoms specifically); it does appear that EMA, in combination with FLMS, can be used to more accurately predict suicidal ideation and risk prior to clinician assessment compared to more infrequent in-person data alone. EMA approaches could be useful identifying both changes in intensity or severity of precursors like affect and volatility in those same predictors as potential indicators of impulsivity (although we did not see variance as a predictor in this study). Although evidence-based interventions for suicide (e.g., safety planning) can be implemented to mitigate risk once risk is recognized, the ability to detect and, especially, predict risk remains by comparison alarmingly poor. Though small in sample size, the present findings are distinct in offering a data collection methodology and statistical modeling framework that are able to delineate risk factors and possible warning signs antecedent to in-person clinical assessments, though these results will need to be replicated in a larger sample.

A number of limitations should be noted to guide future studies in this area. First, participants were selected for inter-episode status and no substance abuse comorbidity; thus, our results may not apply to more severely ill subjects or patients with substance-use comorbidities. Second, the study is based on a small sample. Replication with a larger and more representative sample is warranted. Third, suicidal ideation was not assessed as part of our EMA protocol. Future studies would benefit from improved symptom monitoring, on a

daily basis, to inform temporal variation in suicide ideation. Fourth, although we achieved a high completion rate of daily diaries, compliance may have been affected by mood and symptoms; including completion rate as a term in prediction models did not increase accuracy in the present example, but this may be a crucial issue in long-term monitoring of patients. Fifth, assessment of the utility of prediction models in the context of other variables (e.g., life events) across outcomes (e.g., manic/depressive episodes) are needed. Sixth, a primary limitation of this study is the reliance on a single item assessment of suicidal ideation. Replication using validated suicidal symptom instruments that would more comprehensively and reliably assess suicide risk according to established risk assessment frameworks is desirable. However, despite these limitations, we hope that based on our findings future could specifically address the use of such prediction models to prognostic and treatment outcomes.

In conclusion, few interventions currently capitalize upon continuous data collected after an initial assessment (42). For heterogeneous, chronic and relapsing illnesses such as bipolar disorder, EMA data in conjunction with predictive models such as FLMs may be useful to identify and intervene on identified high-risk states in efforts toward secondary prevention of suicide.

Acknowledgments

This work was supported by National Institute of Mental Health (NIMH) grants MH091260 and MH100417 (CAD), P30MH090333 (WKT), and NIMH Ruth L. Kirschstein National Research Service Award Postdoctoral Fellowship F32MH76339; and the Department of Veterans Affairs Advanced Fellowship Program in Mental Illness Research and Treatment (AG).

References

1. Shiffman S, Stone AA, Huffors MR. Ecological momentary assessment. *Annu Rev Clin Psychol.* 2008; 4:1–32. [PubMed: 18509902]
2. Cain AE, Depp CA, Jeste DV. Ecological momentary assessment in aging research: a critical review. *J Psychiatr Res.* 2009; 43:987–996. [PubMed: 19272611]
3. Depp CA, Mausbach B, Granholm E, et al. Mobile interventions for severe mental illnesses: Design and preliminary data from three interventions. *J Nerv Ment Dis.* 2010; 198:715–721. [PubMed: 20921861]
4. Goodwin, G.; Jamison, K. *Manic Depressive Illness.* Oxford University Press; New York: 2007.
5. Zalsman G, Braun M, Arendt M, et al. A comparison of the medical lethality of suicide attempts in bipolar and major depressive disorders. *Bipolar Disord.* 2006; 8:558–565. [PubMed: 17042829]
6. Kuo WH, Gallo JJ, Tien AY. Incidence of suicide ideation and attempts in adults: the 13-year follow-up of a community sample in Baltimore, Maryland. *Psychol Med.* 2001; 31:1181–1191. [PubMed: 11681544]
7. Oquendo MA, Currier D, Mann JJ. Prospective studies of suicidal behavior in major depressive and bipolar disorders: what is the evidence for predictive risk factors? *Acta Psychiatr Scand.* 2006; 114:151–158. [PubMed: 16889585]
8. Goldstein TR, Ha W, Axelson DA, et al. Predictors of prospectively examined suicide attempts among youth with bipolar disorder. *Archiv Gen Psychiatry.* 2012; 69:1113–1122.
9. Ben-Zeev D, Young MA, Depp CA. Real-time predictors of suicidal ideation: Mobile assessment of hospitalized depressed patients. *Psychiatry Res.* 2012; 197:55–59. [PubMed: 22397912]
10. Humber N, Emsley R, Pratt D, Tarrier N. Anger as a predictor of psychological distress and self-harm ideation in inmates: A structured self-assessment diary study. *Psychiatry Res.* 2013; 210:166–173. [PubMed: 23643187]

11. Hedeker D, Mermelstein RJ, Demirtas H. Modeling between-subject and within-subject variances in ecological momentary assessment data using mixed-effects location scale models. *Stat Med*. 2012; 31:3328–3336. [PubMed: 22419604]
12. Tschacher W, Ramseyer F. Modeling psychotherapy process by time-series panel analysis (TSPA). *Psychother Res*. 2009; 19:469–481. [PubMed: 19585371]
13. Wild B, Eichler M, Friederich H-C, Hartmann M, Zipfel S, Herzog W. A graphical vector autoregressive modelling approach to the analysis of electronic diary data. *BMC Med Res Methodol*. 2010; 10:28. [PubMed: 20359333]
14. Anestis MD, Silva C, Lavender JM, et al. Predicting nonsuicidal self-injury episodes over a discrete period of time in a sample of women diagnosed with bulimia nervosa: an analysis of self-reported trait and ecological momentary assessment based affective lability and previous suicide attempts. *Int J Eat Disord*. 2012; 45:808–811. [PubMed: 21744378]
15. Nock MK, Prinstein MJ, Sterba SK. Revealing the form and function of self-injurious thoughts and behaviors: a real-time ecological assessment study among adolescents and young adults. *J Abnorm Psychol*. 2009; 118:816–827. [PubMed: 19899851]
16. Palmier-Claus JE, Taylor PJ, Gooding P, Dunn G, Lewis SW. Affective variability predicts suicidal ideation in individuals at ultra-high risk of developing psychosis: an experience sampling study. *Br J Clin Psychol*. 2012; 51:72–83. [PubMed: 22268542]
17. 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. [PubMed: 12622659]
18. Judd LL, Akiskal HS, Schettler PJ, et al. The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Arch Gen Psychiatry*. 2002; 59:530–537. [PubMed: 12044195]
19. Jones S. Circadian rhythms, multilevel models of emotion and bipolar disorder-- an initial step towards integration? *Clin Psychol Rev*. 2001; 21:1193–1209. [PubMed: 11702512]
20. Johnson SL, Cuellar AK, Ruggero C, et al. Life events as predictors of mania and depression in bipolar I disorder. *J Abnorm Psychol*. 2008; 117:268–277. [PubMed: 18489203]
21. Thompson RJ, Berenbaum H, Bredemeier K. Cross-sectional and longitudinal relations between affective instability and depression. *J Affect Disord*. 2011; 130:53–59. [PubMed: 20951438]
22. James GM. Generalized linear models with functional predictors. *J R Statist Soc B*. 2002; 64:411–432.
23. Müller H-G, Stadtmüller U. Generalized functional linear models. *Ann Stat*. 2005; 33:774–805.
24. Gershon A, Thompson WK, Eidelman P, McGlinchey E, Kaplan KA, Harvey AG. Restless pillow, ruffled mind: sleep and affect coupling in inter-episode bipolar disorder. *J Abnorm Psychol*. 2012; 121:863–873. [PubMed: 22845651]
25. First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JBW. Structured Clinical Interview for DSM-IV Axis I disorders (SCID I). Biometrics Research, New York State Psychiatric Institute; New York: 1997.
26. Young RC, Biggs J, Ziegler V, Meyer D. A rating scale for mania: reliability, validity, and sensitivity. *Br J Psychiatry*. 1978; 133:429–435. [PubMed: 728692]
27. Rush AJ, Bernstein IH, Trivedi MH, et al. An evaluation of the quick Inventory of Depressive Symptomatology and the Hamilton Rating Scale for Depression: a sequenced treatment alternatives to relieve depression trial report. *Biol Psychiatry*. 2006; 59:493–501. 15. [PubMed: 16199008]
28. Watson D, Clark LA, Carey G. Positive and negative affectivity and their relation to anxiety and depressive disorders. *J Abnorm Psychol*. 1988; 97:346–353. [PubMed: 3192830]
29. Edinger J, Wyatt J, Olsen M, Stechuchak K, Carney C, Chiang A. Reliability and validity of the Duke Structured Interview for Sleep Disorders for insomnia screening. *Sleep*. 2009; 32:810.
30. Williams JB, Gibbon M, First MB, et al. The Structured Clinical Interview for DSM-III-R (SCID): multisite test-retest reliability. *Arch Gen Psychiatry*. 1992; 49:630–636. [PubMed: 1637253]
31. Skre I, Onstad S, Torgersen S, Kringlen E. High interrater reliability for the structured clinical interview for DSM-III-R axis I (SCID-I). *Acta Psychiatr Scand*. 1991; 84:167–173. [PubMed: 1950612]

32. Hofmann BU, Meyer TD. Mood fluctuations in people putatively at risk for bipolar disorders. *Br J Clin Psychol.* 2006; 45:105–110. [PubMed: 16480569]
33. Lovejoy MC, Steuerwald BL. Subsyndromal unipolar and bipolar disorders: Comparisons on positive and negative affect. *J Abnorm Psychol.* 1995; 104:381–384. [PubMed: 7790640]
34. Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica.* 1969; 37:424–438.
35. Thompson, WK.; Choi, J-I.; Anderson, S. Clinical prediction models. In: Lavretsky, H.; Sajatovic, M.; Reynolds, CF., III, editors. *Late-life Mood Disorders.* Oxford University Press; New York: 2013. p. 704-713.
36. Hawkins DM. The problem of overfitting. *J Chem Information Computer Sci.* 2004; 44:1–12.
37. Green, PJ.; Silverman, BW. *Nonparametric regression and generalized linear models: a roughness penalty approach.* Chapman & Hall London; London: 1994.
38. Cardot H, Sarda P. Estimation in generalized linear models for functional data via penalized likelihood. *J Multivariate Analy.* 2005; 92:24–41.
39. Fan J, Zhang JT. Two-step estimation of functional linear models with applications to longitudinal data. *J R Statist Soc B.* 2000; 62:303–322.
40. Ramsay, JO.; Silverman, BW. *Functional Data Analysis.* Springer; New York: 2005.
41. Thompson WK, Rosen O. A Bayesian model for sparse functional data. *Biometrics.* 2008; 64:54–63. [PubMed: 17573864]
42. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *Br J Health Psychol.* 2010; 15:1–39. [PubMed: 19646331]

Fig. 1. Mean ecological momentary assessment negative and positive affect trajectories
 X-axis is number of days until the next in-person clinician assessment of suicidal ideation.
 Y-axis is standardized total score of negative and positive affect from the Positive and Negative Affect Scale. Black lines give mean levels for trajectories ending in a suicidal ideation score of 0, whereas red lines give mean levels for trajectories ending in suicidal ideation scores > 0.

Fig. 2. Prediction accuracy using ecological momentary assessment trajectories of negative and positive affect

Top panels show prediction results for Model 1 (negative affect alone). Middle panels show prediction results for Model 2 (positive affect alone). Bottom panels show prediction results for Model 3 (negative and positive affect together). Left panels show cross-validated one-week-ahead predicted standardized suicidal ideation scores. Right panels show corresponding receiver operating characteristic (ROC) curves formed by thresholding cross-validated predicted standardized scores over a range of cut-offs.

Fig. 3. Beta weight functions for functional linear model predicting suicidal ideation

The x-axis is number of days until the next in-person clinician assessment of suicidal ideation. Beta weight functions are from Model 3 including both Positive and Negative Affect Scale negative and positive mood total scores.

Table 1

Demographic and clinical characteristics

Characteristics	Mean (SD) or Frequency (%)
Age, years, mean (SD)	34.3 (10.7)
Gender, female, n (%)	22 (62.9)
Ethnicity, n (%)	
African American	3 (8.6)
Asian	3 (8.6)
Hispanic	3 (8.6)
Native American	1 (2.9)
White	23 (65.7)
Biracial/Other	2 (5.7)
Marital status, n (%)	
Single	24 (68.6)
Married/Live-in partner	2 (5.7)
Divorced/Separated/Widowed	9 (25.7)
Employment status, n (%)	
Full time	11 (31.4)
Part time	10 (28.6)
Student	6 (17.1)
Unemployed/Retired	8 (22.9)
Age at illness onset, years, mean (SD)	19.9 (7.6)
No. of lifetime manic episodes, mean (SD)	7.9 (15.3)
No. of lifetime depressive episodes, mean (SD)	8.4 (11.5)
History of psychiatric hospitalizations, mean (SD)	19.0 (63.3)
No. of lifetime psychiatric hospitalizations, mean (SD)	1.8 (3.0)
Medication status	
Overall somatotherapy score, mean (SD)	1.7 (1.4)
Any medication, n (%)	33 (94.3)
Lithium, n (%) ^a	5 (14.3)
Antidepressants, n (%) ^a	15 (42.9)
Valproic acid, n (%) ^a	3 (9.0)
Atypical antipsychotics, n (%) ^a	16 (45.7)
Alternative (e.g., lamotrigine), n (%) ^a	14 (40.0)

SD = standard deviation.

^aPresent treatment.