Evidence for mood instability in patients with bipolar disorder

Mildiner Moraga, Sebastian; Bos, Fionneke M; Doornbos, Bennard; Bruggeman, Richard; van der Krieke, Lian; Snippe, Evelien; Aarts, Emmeke

Published in:
Journal of psychopathology and clinical science

DOI:
10.1037/abn0000915

Publication date:
2024

Journal of Psychopathology and Clinical Science

ISSN: 2769-7541

2024, Vol. 133, No. 6, 456–468
https://doi.org/10.1037/abn0000915

Evidence for Mood Instability in Patients With Bipolar Disorder: Applying Multilevel Hidden Markov Modeling to Intensive Longitudinal Ecological Momentary Assessment Data

Sebastian Mildiner Moraga[1], Fionneke M. Bos[2, 3], Bennard Doornbos[4], Richard Bruggeman[2], Lian van der Krieke[2], Evelien Snippe[3], and Emmeke Aarts[1]

[1] Department of Methodology and Statistics, Faculty of Social and Behavioural Sciences, Utrecht University
[2] Department of Psychiatry, Rob Giel Research Centre, University Medical Centre Groningen, University of Groningen
[3] Department of Psychiatry, Interdisciplinary Centre Psychopathology and Emotion Regulation (ICPE), University Medical Centre Groningen, University of Groningen
[4] Lentis Research, Lentis Psychiatric Institute, Groningen, The Netherlands

Bipolar disorder (BD) is a chronic psychiatric condition characterized by large episodic changes in mood and energy. Recently, BD has been proposed to be conceptualized as chronic cyclical mood instability, as opposed to the traditional view of alternating discrete episodes with stable periods in-between. Recognizing this mood instability may improve care and call for high-frequency measures coupled with advanced statistical models. To uncover empirically derived mood states, a multilevel hidden Markov model (HMM) was applied to 4-month ecological momentary assessment data in 20 patients with BD, yielding ∼9,820 assessments in total. Ecological momentary assessment data comprised self-report questionnaires (5 × daily) measuring manic and depressive constructs. Manic and depressive symptoms were also assessed weekly using the Altman Self-Rating Mania Scale and the Quick Inventory for Depressive Symptomatology Self-Report. Alignment between HMM-uncovered momentary mood states and weekly questionnaires was assessed with a multilevel linear model. HMM uncovered four mood states: neutral, elevated, mixed, and lowered, which aligned with weekly symptom scores. On average, patients remained <25 hr in one state. In almost half of the patients, mood instability was observed. Switching between mood states, three patterns were identified: patients switching predominantly between (a) neutral and lowered states, (b) neutral and elevated states, and (c) mixed, elevated, and lowered states. In all, elevated and lowered mood states were interspersed by mixed states. The results indicate that chronic mood instability is a key feature of BD, even in "relatively" euthymic periods. This should be considered in theoretical and clinical conceptualizations of the disorder.

This article was published Online First June 3, 2024.
Daniel Fulford served as action editor.
Sebastian Mildiner Moraga https://orcid.org/0000-0001-6282-7394
Fionneke M. Bos https://orcid.org/0000-0002-9630-0440
Bennard Doornbos https://orcid.org/0000-0002-5013-3501
Richard Bruggeman https://orcid.org/0000-0002-3238-8471
Lian van der Krieke https://orcid.org/0000-0001-6745-5913
Evelien Snippe https://orcid.org/0000-0002-3003-7475
Emmeke Aarts https://orcid.org/0000-0002-2432-7564

This research was funded in whole, or in part, by the NWO (Grant EINF-2570). For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. This article is based on data published in Bos et al. (2020), a subset of which was also analyzed in Bos et al. (2022) using different methodologies. This study was financially supported by the Rob Giel Research Center. This work made use of the Dutch National e-Infrastructure Snellius with the support of the SURF Cooperative using Grant EINF-2570, which is (partly) financed by the Dutch Research Council (Nederlandse Organisatie voor Wetenschappelijk Onderzoek). All funding sources had no role in the design of this study, nor did they have any role during its execution, analyses, interpretation of the data, and the decision to submit results. We thank all patients and clinicians who have shared their data, thoughts, and ideas with us, the clinicians at Mental Healthcare (Geestelijke Gezondheidszorg) Drenthe and the University Center Psychiatry Groningen for their help with the recruitment of patients, and S. van der Tuin for helping with the data collection. The authors declare that they have no competing interests. The study was approved by the University Medical Center Groningen medical ethics committee (20150 1161). All participants signed informed consent. The data that support the findings of this study are available from the corresponding author upon reasonable request. The code behind this study has been made publicly available at Zenodo and can be accessed at https://doi.org/10.5281/zenodo.8175673.

Sebastian Mildiner Moraga served as lead for formal analysis, methodology, validation, and visualization. Fionneke M. Bos served as lead for data curation and investigation. Bennard Doornbos served in a supporting role for formal analysis. Evelien Snippe served as lead for funding acquisition and served in a supporting role for formal analysis and writing–review and editing. Emmeke Aarts served as lead for funding acquisition and supervision and contributed equally to methodology. Sebastian Mildiner Moraga and Emmeke Aarts contributed equally to software. Sebastian Mildiner Moraga, Fionneke M. Bos, Bennard Doornbos, Richard Bruggeman, Lian van der Krieke, and Emmeke Aarts contributed equally to writing–review and editing. Fionneke M. Bos, Bennard Doornbos, Richard Bruggeman, Lian van der Krieke, and Evelien Snippe contributed equally to conceptualization. Fionneke M. Bos and Emmeke Aarts contributed equally to formal analysis. Sebastian Mildiner Moraga and Fionneke M. Bos contributed equally to writing–original draft.

Correspondence concerning this article should be addressed to Sebastian Mildiner Moraga, Department of Methodology and Statistics, Faculty of Social and Behavioural Sciences, Utrecht University, Padualaan 14, 3584 CH Utrecht, The Netherlands. Email: s.mildinermoraga@uu.nl

General Scientific Summary

The article investigates bipolar disorder with advanced statistical methods and adds evidence challenging the traditional view of alternating discrete mood episodes with stable periods in-between. Instead, it suggests that chronic mood instability may be a significant aspect of bipolar disorder and calls for high-frequency assessments using advanced statistical models to better understand and improve care for individuals with the disorder.

Keywords: bipolar and related disorders, intensive ecological momentary assessment, hidden Markov model, multilevel analysis, mobile health

Supplemental materials: https://doi.org/10.1037/abn0000915.supp

Bipolar disorder (BD) is a chronic psychiatric condition characterized by large shifts in mood, energy, and cognitive functioning. These shifts range from manic episodes during which mood and energy are elevated to the opposite pattern during depressive episodes (Leboyer & Kupfer, 2010). Furthermore, patients can experience so-called mixed episodes during which manic and depressive symptoms co-occur (Abreu & Bragança, 2015). Although the *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition (*DSM-5*) conceptualization of BD still centers around discrete episodes, there have been recent calls to address the longitudinal and cyclical nature of the condition (Harrison et al., 2018). Indeed, mood instability in between episodes may lie at the heart of the disorder (Bonsall et al., 2012; Gottschalk et al., 1995; Hadaeghi et al., 2013). Patients with BD have been found to shift frequently between mood states, from week to week (Bonsall et al., 2012; Reich et al., 2012), day to day (i.e., ultra-rapid cycling), and even hour to hour (i.e., ultradian cycling; Kramlinger & Post, 1996). Such mood instability predicts worse outcomes on both a symptomatic and functional level and may require alternative treatment strategies (Carvalho et al., 2014; Sperry et al., 2020). Gaining a better understanding of the nature of mood instability in BD is therefore necessary.

Given its detrimental impact, it is unsurprising that there is a large body of literature on the nature and extent of mood instability. Roughly, studies can be divided into relatively infrequent (weekly or monthly) clinical assessments of symptoms and more frequent (daily or hourly) self-report assessments of mood and affective states. In the first type of studies (i.e., with assessments spaced further apart), the hidden Markov model (HMM) has been applied to uncover empirically derived states from longitudinal data and determine the probabilities of switching between states over time (Rabiner, 1989; Zucchini et al., 2017). The HMM is a latent longitudinal mixture modeling approach (Harrison et al., 2018; McInnis et al., 2022). For example, Cochran et al. (2016) showed that patients with Type I BD could be objectively assigned to stable, depressive, and rapid-cycling mood states and that individuals belonging to the same class shared common clinical features. Another recent study uncovered three mood states (euthymic, depressed, and mixed) from manic and depressive symptom questionnaires administered at five time points 3 months apart (Prisciandaro et al., 2019). Although the euthymic and depressed states were relatively prevalent and stable, the mixed state occurred infrequently and was unstable (i.e., was not consistent across time points) and was associated with rapid cycling, substance abuse, and psychosis. Finally, Yee et al. (2021) found five mood states (euthymic, two manic microstates, and two depressive microstates) in patients with Type I BD using weekly measurements.

The second approach uses more frequent assessments. In these studies, mood instability is often operationalized as the mean squared successive difference between the current assessment and the previous one, a higher mean squared successive difference signifying higher mood instability. Several studies focused on day-to-day self-reported mood using a single mood item, assessed for 9 months up to 2 years, which demonstrated that day-to-day mood instability is present in patients with BD, during manic, depressive, and euthymic episodes (Faurholt-Jepsen et al., 2019b, 2023; Stanislaus et al., 2020). Furthermore, mood instability has been associated with impaired quality of life and increased stress (Faurholt-Jepsen et al., 2019b, 2023; Stanislaus et al., 2020) and differed between bipolar Type I versus Type II (Faurholt-Jepsen et al., 2019b). Zooming in even further to within-day assessments, several ecological momentary assessment (EMA) studies have demonstrated evidence for increased mood instability in BD. EMA is an intensive self-monitoring method that assesses mood, symptoms, and behaviors 8–10 times daily on patients' smartphones, usually for a shorter time period ($\pm$14 days, Shiffman et al., 2008). A recent EMA study found higher (sad and anxious) mood instability in patients with bipolar patients versus healthy controls (Lamers et al., 2018). Furthermore, three studies showed that higher mood instability of positive and negative affect was associated with higher BD psychopathology in nonclinical individuals (Sperry & Kwapil, 2019, 2022; Sperry et al., 2020).

Taken together, these studies demonstrate mood instability is a key factor in BD on different timescales. However, the different methodologies each have their strengths and limitations. Longitudinal HMM studies focused on group-level patterns reducing our insight into individual patterns of mood instability. Furthermore, the infrequent nature of this data complicates the study of the extent and frequency of mood instability (e.g., hourly, daily, or weekly). Daily mood studies so far relied on a single item to assess mood and may miss relevant patterns when mood is assessed more broadly and preclude insight into within-day mood patterns. Finally, EMA studies are relatively short (i.e., 2 weeks) and focus on broader positive and negative affective states rather than mood and activity symptoms that are specifically relevant for BD. As such, longitudinal EMA data could bridge the gap between these studies in the current literature. In addition, by applying a multilevel approach to HMM, we can not only infer mood patterns shared among individuals but also uniquely provide insights into individual-specific mood patterns. This offers new insights into the heterogeneity in mood fluctuations. As such, this approach could offer insight into how patients move from one bipolar mood state to another in the space of a few hours, how long they remain within these states, and whether this differs across individuals.

The present exploratory study therefore applied multilevel HMM (MHMM) to 4-month EMA data in 20 bipolar patients to uncover empirically derived mood states, the probabilities of switching between states, and the duration patients remain within states. By comparing the results of this novel methodology to weekly administered traditional validated symptom questionnaires, we will be able to determine the validity of our approach. In this article, we will study (a) which momentary bipolar mood states can be uncovered, (b) the duration patients remain within each state, (c) the probabilities of switching between states, and (d) the alignment of these momentary states with weekly administered validated symptom questionnaires. Whereas previous studies could only examine average (group-level) dynamics, this is the first study to combine MHMM with fine-grained longitudinal EMA data to offer insight into individual mood dynamics in BD.

Materials and Method

Study Overview

The study methods have been described in detail elsewhere (Bos et al., 2022). Briefly, it included 20 patients with BD Type I/II of ≥18 years, currently in treatment, who reported at least two manic and/or depressive episodes in the previous year. Clinicians of two Dutch tertiary care institutions invited patients for the study until the intended cap of 20 participants was reached. In total, 28 patients were referred, two of whom were unreachable, and six declined to participate because they expected that study participation and the focus on mood would be too burdensome. This left 20 patients who started and finished the study. The study was approved by the University Medical Center Groningen Medical Ethics Committee (201501161).

The study was designed to assess the added value of long-term EMA in BD treatment (Bos et al., 2020) and to explore whether manic and depressive transitions can be anticipated with EMA data (Bos et al., 2022). To this end, participants completed five EMA questionnaires per day for 4 months, adding up to approximately 500 assessments of 29 EMA items per participant, yielding almost 300,000 observations in total. Starting in the morning at a self-chosen time, every 3 hr, participants received a text message with a link to the EMA on their smartphone, which took 1–2 min to complete. In addition, participants completed weekly validated questionnaires on manic and depressive symptoms. Questionnaires were securely administered and stored via RoQua in participants' electronic health records (https://www.roqua.nl). Participants completed on average 18 weeks of monitoring (range = 16–32), except for two participants who completed 12 weeks of monitoring because they were part of a pilot.

Measures

The EMA questionnaire (see Table S1 in the online supplemental materials) consisted of 29 items pertaining to momentary mood, symptoms, sleep, and activities. These were based on previous EMA research (Knapen, 2019; Krieke et al., 2016; Tyler et al., 2015) and interviews with three patients and a psychiatrist. In these interviews, patients were asked to reflect on what manic and depressive symptoms felt like in daily life and which words best described their experiences prior to and during episodes.

In the current study, we selected the 12 EMA items that were intended to measure manic symptoms ("agitated," "extremely well," "irritated," "full of ideas," "racing thoughts," "ability to focus/switch") and depressive symptoms ("down," "tired," "content," "dreading the rest of the day," "inadequate," "worry"). All items were measured on 0–100 visual analog scales (ranging from *not at all* to *very much*). The within- and between-patient composite reliability omega (ω^w and ω^b) obtained from two separate two-level confirmatory factor analyses as described in Geldhof et al. (2014) for the manic and depressive EMA items, were $\omega^w = .66$, 95% confidence intervals (CIs) = [0.64–0.67], and $\omega^b = .91$ (95% CIs = [0.85–0.98]), and $\omega^w = .78$ (95% CIs = [0.78–0.79]) and $\omega^b = .92$ (95% CIs = [0.86–0.98]), correspondingly. On average, the 20 participants completed 76% (SD = 14.1%) of the EMA questionnaires ($M = 491$ assessments, SD = 137.8).

Additionally, participants completed validated symptom questionnaires weekly. Although clinical interviews are generally the golden standard in terms of diagnostic accuracy, they can be burdensome, especially if episodes need to be assessed weekly over a 4-month period in addition to EMA. We therefore opted for self-report measures of depressive and manic episodes. For manic symptoms, this was the Altman Self-Rating Mania (ASRM) Scale, and for depressive symptoms, the Quick Inventory for Depressive Symptomatology Self-Report (QIDS-SR). The ASRM and QIDS-SR have shown adequate psychometric properties and are sufficiently able to classify individuals as experiencing manic, depressive, or euthymic episodes (see below). Furthermore, they have been combined previously as weekly mood-monitoring tools and have been deemed feasible for routine mood monitoring (Simon et al., 2017).

Manic Episodes

The Dutch version of the five-item ASRM (E. G. Altman et al., 1997) inquires on manic[1] symptoms on a 0–4 scale (sum score ranges 0–20). A score of ≥6 has been found to be indicative of a potential (hypo)manic episode in bipolar samples (Miller et al., 2009). A recent systematic review demonstrated that among mania self-rating scales, the ASRM has high clinical utility (9.6 out of 11), a composite score of reliability, validity, ease of use, and sensitivity to change (Cerimele et al., 2019). Indeed, the ASRM correlated .77 with the Clinician-Administered Rating Scale for Mania and showed high sensitivity (85%–93%) and specificity (82%–87%), indicating it can accurately identify the presence or absence of manic episodes in patients with BD (E. Altman et al., 2001; Meyer et al., 2020). Patients completed 95% of ASRM questionnaires in this study.

Depressive Episodes

The Dutch version of the 16-item QIDS-SR (Rush et al., 2003) inquires on depressive symptoms on a 0–3 scale (sum score ranges 0–48). Similar to the ASRM, a score of ≥6 has been found to be indicative of a potential depressive episode in bipolar samples (Bernstein et al., 2010). The QIDS-SR had a clinical utility score of 6.75 out of 11 and correlated .89–.92 with the clinician-rated version of the QIDS (QIDS-C; Bernstein et al., 2010; Cerimele et al., 2019). Although to the best of our knowledge no measures of sensitivity and specificity are available in bipolar samples, QIDS-SR and

[1] Although $N = 9$ patients had bipolar disorder Type II, suggesting they experience hypomanic episodes, we refer to both hypomanic and manic symptoms as "manic symptoms" for consistency.

QIDS-C scores reflected the diagnostic phases of the disorder (i.e., depressed, euthymic, or manic), as assessed by structured clinical interviews. Compliance in this study was 94%.

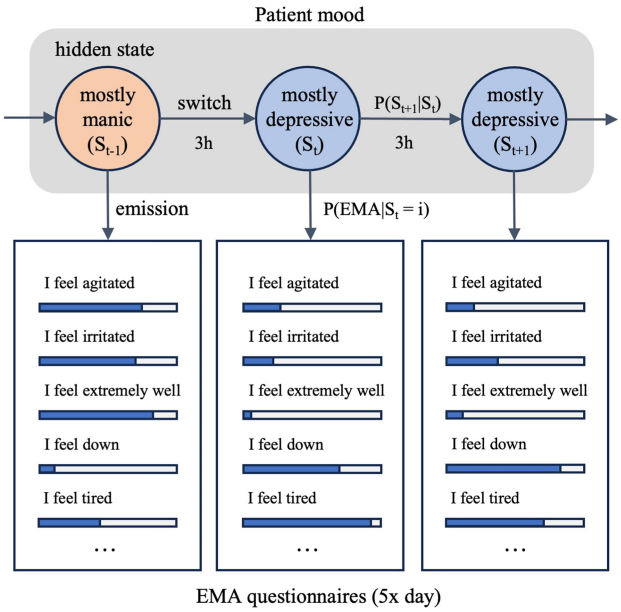
Statistical Analysis

The current study used a Bayesian MHMM (R. M. K. Altman, 2007; Shirley et al., 2010; Zhang & Berhane, 2014) to explore mood states' dynamics in patients with BD. HMMs are probabilistic models used to study processes characterized by switches between discrete latent (hidden) states, which are assumed to follow a first-order Markov chain (i.e., the probability of switching to a state only depends on the previous state). When applied to psychological processes, the HMM can be used to quantify information on the latent temporal dynamics into two sets of parameters: the probability of switching between latent states $P(S_t|S_{t-1})$ and the probability of emitting observations given the latent states $P(\text{Obs}_t|S_t)$. Since the EMA items are measured on a continuous scale, we adopted a normal distribution for the observations. Multiple items can be accommodated in the model at once, assuming conditional independence of the item scores given the states. Finally, the MHMM used here assumes time-homogeneous switching probabilities.

In the current study, latent states represent mood states inferred from the response patterns of patients with BD over the 12 manic and depressive EMA items simultaneously, as depicted in Figure 1.

Figure 1

Schematic Diagram of a Simple Hidden Markov Model



Note. The latent mood states (i.e., hidden states), represented by circles, are not directly observable and are assumed to follow a Markov process. As such, the probability switching to a next state S at Time Point $t+1$ only depends on the current state S at Time Point t: $P(S_{t+1}|S_t)$. The pattern of response on the 12-item EMAs filled out by patients every 3 hr, represented by rectangles, are assumed to be a realization of the mood state $S = i$ of the patient at that Time t: $P(\text{EMA}|S_t = i)$. EMA = ecological momentary assessment. See the online article for the color version of this figure.

Thus, the scores over the 12 EMA items at each measurement occasion are assumed to represent the mood state of the patient at that time. For instance, a patient in a mood state with predominantly depressive symptoms on the occasion t is expected to be more likely to experience low scores for "agitated" and "extremely well" and high scores on items such as "down" or "tired," and the opposite would be expected for a patient in a state with predominantly manic symptoms. By extending the model to the mixed-effect framework, the MHMM offers three advantages over conventional HMMs: (a) it provides each patient with individual parameters including an individual trajectory over time, while ensuring that the meaning of each mood state remains the same across patients, (b) it allows for the quantification of variability between patients with random effects, and (c) it is more robust to outliers (Gelman et al., 2013). See the online supplemental materials for a more detailed statistical model specification.

Model Fit

Fitting of the MHMM was performed with the statistical software R (R Core Team, 2021) and the R package mHMMbayes (Aarts, 2019) in a Bayesian framework (for details, see the online supplemental materials). Twelve-item intense longitudinal EMA data were used to train models with two to seven states based on the number of states found in previous studies (Cochran et al., 2016; Prisciandaro et al., 2019; Yee et al., 2021). Model selection was performed considering expert knowledge (i.e., a discussion with the team's psychiatrist using information on state composition), comparing relative model fit with Akaike information criterion (AIC), and the ability of the model to reproduce the original EMA data via Bayesian posterior predictive checks (for details, see Gelman et al., 2013) as advised in Pohle et al. (2017). The goodness of fit was further evaluated by examining the distribution, homoscedasticity, and autocorrelation of pseudoresiduals for each patient on the final model by visual inspection. The reliability of the model parameters and mood states uncovered with the MHMM (see below) was assessed with a small Monte Carlo simulation.

Missing Data

Missing data were assumed missing at random (MAR; Little & Rubin, 2019), so that the missingness mechanism is independent of the missing data and the hidden states given the observed data and model parameters. To check this, we calculated the proportion of missing observations across the states. Three overnight measurement occasions were added as missing values to fulfill the assumption of equally spaced observations and treated in the same way as missing observations (Asparouhov et al., 2018; Hamaker et al., 2018).

State Duration and Switching

The state durations in hours can be estimated from the self-switching probabilities using the geometric distribution implicitly assumed by the model. Alternatively, empirical state durations can be obtained from uncovered state sequences, less sensitive to implicit assumptions (Shappell et al., 2019). Expected geometric and empirical durations were computed for each patient and the aggregated sample. To facilitate between switching probabilities comparisons, we normalized the off-diagonal probabilities. Additionally, we calculated the relative switching frequency, representing the proportion of occasions with state switches, for each patient and aggregated data.

Uncovering Mood States and Alignment With Weekly Symptom Data

Uncovering the sequence of momentary mood states was done for each patient with the Viterbi algorithm (Forney, 1973) based on patient-specific parameters, which yields the most likely sequence of states based on the pattern of response of a patient over the 12 EMA items. We used the R package lme4 (Bates et al., 2015) to fit linear mixed-effects regressions with patient-specific random coefficients to assess the effect of inferred mood state on the weekly ASRM and QIDS symptom scores of each patient. The discrepancy in the measurement frequency between EMA and weekly data was dealt with by assignment of each week's ASRM and QIDS scores to the mood states uncovered on the 6 days prior to, 3 days prior to, or the same day of the administration of the weekly questionnaires. The models were estimated using data for the three windows, and the within-patient explained variance (R^2) was calculated.

Open Practices, Transparency, and Data Sharing

The data that support the findings of this study are available from the corresponding author upon reasonable request to the corresponding author. The code behind this study has been made publicly available at Zenodo and can be accessed at https://doi.org/10.5281/zenodo.8175673. No part of the research reported in this study has been preregistered.

Results

As can be observed in Table 1, most patients were female and used medication. Patients had been in treatment for on average 10.6 years ($SD = 8.8$). Almost half of the patients were diagnosed with bipolar Type I and reported a comorbid disorder.

Mood State Composition

The four-state MHMM was found to have the best fit, as indicated by the AIC and model stability, and a good match with previous literature and domain expertise. The mood states uncovered from the EMA data could be classified as neutral, elevated, mixed, and lowered states and are visualized in Figure 2.

Model Fit

Although AIC favored models with higher complexity (Table S2 in the online supplemental materials), models with >5 latent states ran into convergence issues. Posterior predictive checks for the four-state MHMM revealed a good fit for group-level and patient-level means of the 12 EMA items (Figures S1 and S2 in the online supplemental materials). Pseudoresiduals generally exhibited adequate distributional properties across patients. The proportion of missing observations was stable across states in the aggregate data (0.22–0.26) yet showed substantial variability over patients with no discernible trend (see Figure S3 in the online supplemental materials). As a result, we concluded no violation of the MAR assumption. Finally, the Monte Carlo simulation indicated a high accuracy on the uncovered mood states ($M = 87.4\%$, $SD = 2.9\%$) averaged over patients and simulation repetitions. The absolute percent bias in the self-transition and emission parameters were generally low (≤10% threshold), and parameter coverage was adequate ($M = 91.4\%$, $SD = 8.6\%$, and $M = 97.1\%$, $SD = 3.0\%$, respectively).

Table 1
Demographic and Clinical Characteristics of Patients

Characteristic	Patients (n = 20)
Gender (N)	
Male	4
Female	16
Age (N)	
20–35 years	9
36–50 years	8
51–65 years	3
Education level (N)	
Higher education	9
Secondary education	5
Secondary vocational education	3
Prevocational education	3
Years since BD diagnosis, M (SD)	6.4 (6.3)
BD diagnosis (N)	
BD Type I	9
BD Type II	11
Comorbid diagnoses (N)	
No comorbid axis I/II disorder	12
Attention deficit/hyperactivity disorder	1
Autism spectrum disorder	1
Sleep disorder	1
Alcohol/drug dependence	1
Personality disorder	6
Medication use at study start (N)	
None	2
Amphetamine	1
Antiepileptic	10
Atypical antipsychotic	10
Benzodiazepine	9
Thyreomimetica	2
Lithium	5
Monoamine oxidase inhibitor	3
Selective serotonin reuptake inhibitor	4
Tricyclic antidepressant	1
Manic and depressive symptom levels, M (SD)	
During manic periods (ASRM ≥ 6)	9.0 (3.3)
During depressed periods (QIDS ≥ 6)	11.5 (4.7)
During nonmanic periods (ASRM < 6)	1.3 (1.6)
During nondepressive periods (QIDS < 6)	3.3 (1.3)

Note. BD = bipolar disorder; ASRM = Altman Self-Rating Mania; QIDS = Quick Inventory for Depressive Symptomatology.

Composition

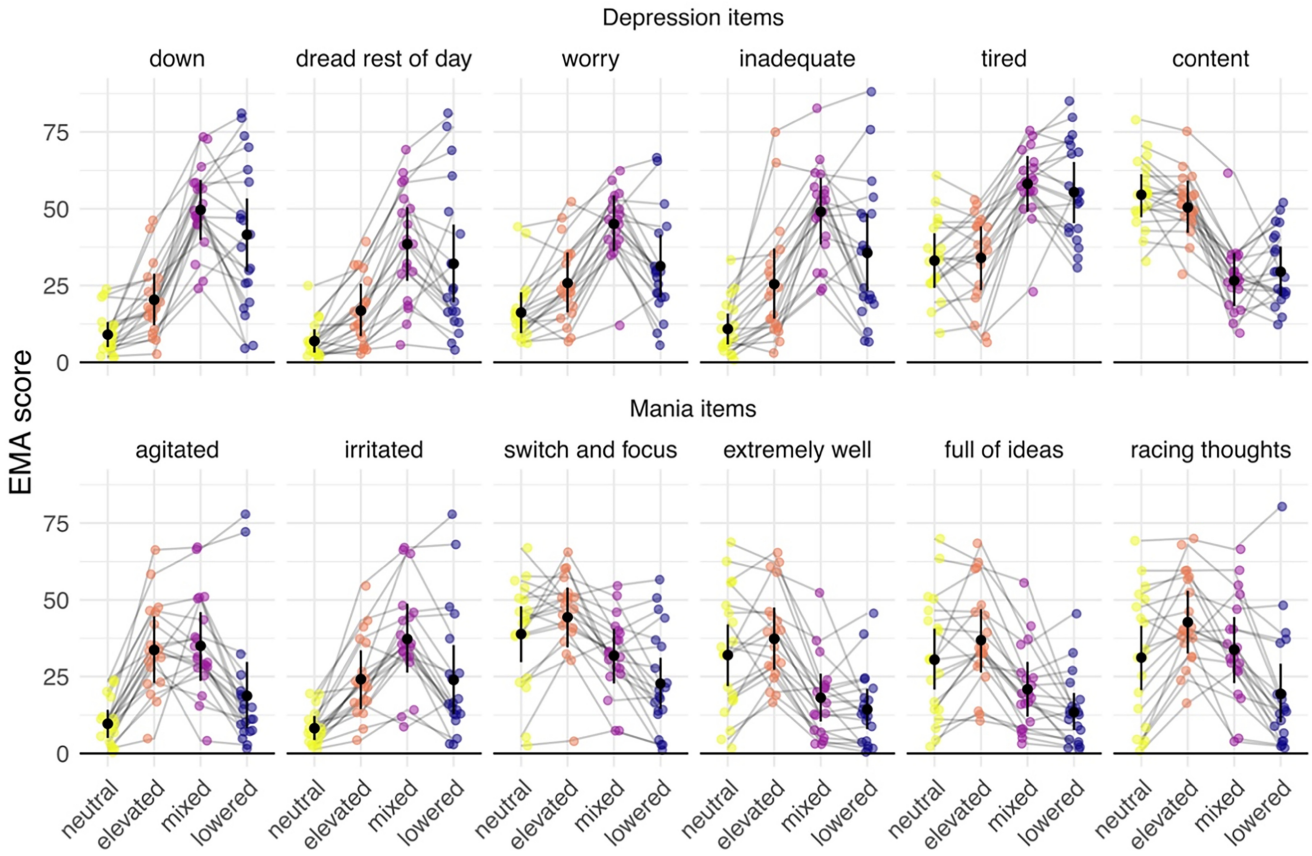
The neutral state was characterized by low scores on manic EMA items such as "agitated" and "irritated," as well as depressive items such as "down" and "worry." Relative to the other mood states, the elevated state exhibited high scores on manic items and low scores on most depressive items. The reverse pattern occurred for the lowered state, except for "irritated," which exhibited a relatively high score. Finally, the mixed state presented a combination of the patterns of elevated and lowered states.

Heterogeneity

Figure 2 also reveals some heterogeneity between patients on their mean scores over the items for the four mood states. The amount of between-patient variability appears to vary over items and states: for

Figure 2

Composition of the Four Mood States Uncovered With the MHMM on the EMA Data



Note. Black point and range indicate group-level, state-specific estimated emission scores for the 12 EMA items, along with 95% confidence intervals. Colored (gray) points indicate patient-specific, state-specific estimated emission scores for each of the 12 EMA items. Lines indicate that the parameters belong to the same patient. MHMM = multilevel hidden Markov model; EMA = ecological momentary assessment. See the online article for the color version of this figure.

example, "agitated" and "irritated" scores were uniformly low across patients for the neutral state but varied more between patients in the elevated, mixed, and lowered states. Nonetheless, the scoring pattern over states within EMA items was relatively constant across patients (i.e., patients generally scored lower or higher irrespective of the EMA item). This indicates that the mood states, found at the group level, represent the same construct and can be interpreted similarly across individual patients while accommodating patient heterogeneity.

Mood State Duration

On average, the empirical duration of each of the mood states obtained from the empirical state decoding was less than 25 hr (Figure 4): patients remained shortest in the mixed state (M = 12.8 hr, SD = 17.3), followed by elevated (M = 13.8 hr, SD = 19.5), lowered (M = 23.3 hr, SD = 121.5), and neutral (M = 24.9 hr, SD = 116.7). Over time and across patients, 12.5% of the measurements were categorized as mixed, 14.8% as elevated, 36.3% as lowered, and 36.4% as neutral.

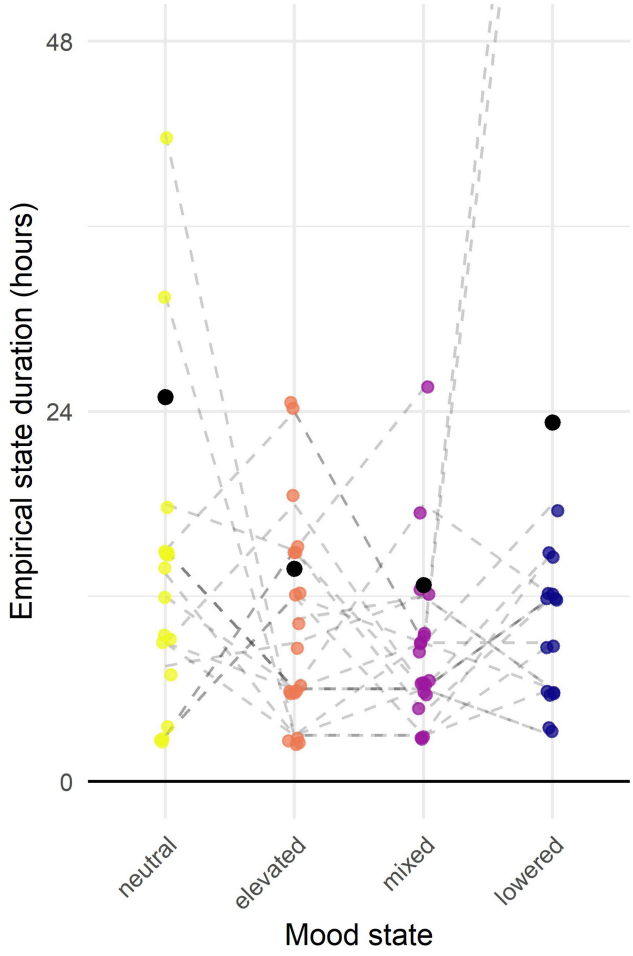
Considerable heterogeneity was observed between patients in terms of their mood state duration. Figure 3 illustrates that the time spent in each mood state varied between patients, ranging 3–372 hr for neutral, 3–38 hr for elevated, 3–43 hr for mixed, and 4–3,072 hr for lowered (see Table S4 in the online supplemental materials for all patient-specific state durations).

Switching Between Mood States

On average, patients switched to a different mood state on 15.9% (SD = 9.6) of the 4-month EMA occasions. Mood instability, measured as the relative number of switches between uncovered mood states, ranged 0.00–0.38 for patients in the sample (Table S3 in the online supplemental materials) with eight patients switching states on ≥20% of the occasions in the 4-month EMA period. Unlike the state composition and duration, switching between states varied to such an extent across patients that these dynamics could not be summarized using solely the group-level parameters. Ranking the patients by their normalized switching probabilities, three distinct switching patterns clearly emerged in the sample. To appreciate the heterogeneity between patients, we will describe the switching dynamics between states inferred from these patterns, summarized in Figure 4 (see Figures S4 and S5 in the online supplemental

Figure 3

Empirical Duration of the Mood States Uncovered With the MHMM on the EMA Data



Note. Colored (gray) points represent the median duration in hours for individual patients; black points represent mean duration for the combined sample. Note that six points corresponding to durations >48 hr were excluded from the plotting area to facilitate visualization. MHMM = multilevel hidden Markov model; EMA = ecological momentary assessment. See the online article for the color version of this figure.

materials for both the ranked normalized and nonranked nonnormalized switching probabilities of the complete sample).

First, the "neutral/lowered" pattern, observed in six patients, was characterized by a heightened probability of switching between neutral and lowered states ($M = 0.87$, range = 0.72–0.95; $M = 0.83$, range = 0.53–0.99). Furthermore, these patients showed a low probability to switch from neutral to elevated states ($M = 0.12$, range = 0.04–0.24). Overall, they showed a high frequency of neutral and lowered states (decoding incidence = 0.67, 0.27), with almost no occurrences of elevated and mixed states (0.03, 0.03).

Second, the "neutral/elevated" pattern, observed in six patients, was characterized by a high probability of switching between neutral and elevated states ($M = 0.60$, range = 0.25–0.87; $M = 0.65$, range = 0.36–0.92). They also exhibited a high probability of switching from lowered back to neutral ($M = 0.53$, range = 0.18–0.86) but a low probability of switching from neutral to lowered ($M = 0.28$, range = 0.01–0.46). In addition, these patients showed moderate probabilities of switching from mixed to elevated ($M = 0.36$, range = 0.21–0.61) and mixed to lowered ($M = 0.46$, range = 0.06–0.73). All states were present in the decoding of this pattern, with neutral being the most frequent mood state (0.44), followed by lowered (0.23), elevated (0.19), and mixed (0.14).

Third, the "mixed/elevated/lowered" pattern, observed in eight patients, was characterized by a substantial probability of switching between mixed and lowered states ($M = 0.60$, range = 0.47–0.78; $M = 0.78$, range = 0.60–0.94) and, to a lesser extent, between mixed and elevated states ($M = 0.36$, range = 0.19–0.51; $M = 0.39$, range = 0.12–0.90). Furthermore, these patients had a high probability of switching from neutral to elevated states ($M = 0.72$, range = 0.63–0.86) and a moderate probability of switching from elevated back to neutral states ($M = 0.41$, range = 0.60–0.94). However, switches between lowered and neutral states were rare for these patients ($M = 0.04$, range = 0.01–0.11; $M = 0.16$, range = 0.02–0.27). The frequency of elevated, mixed, and lowered states was higher (0.22, 0.20, and 0.54, respectively), with almost no occurrences of the neutral state (0.04).

Notably, switches between neutral and mixed ($M = 0.09$, range = 0.00–0.34; $M = 0.10$, range = 0.01–0.32) and elevated and lowered ($M = 0.19$, range = 0.02–0.50; $M = 0.15$, range = 0.01–0.57) were relatively infrequent for all patterns. These results suggest that when present, the mixed state mediated the transition between more stable elevated and lowered states (e.g., see Patients 4, 8, and 10 in Figure 5).

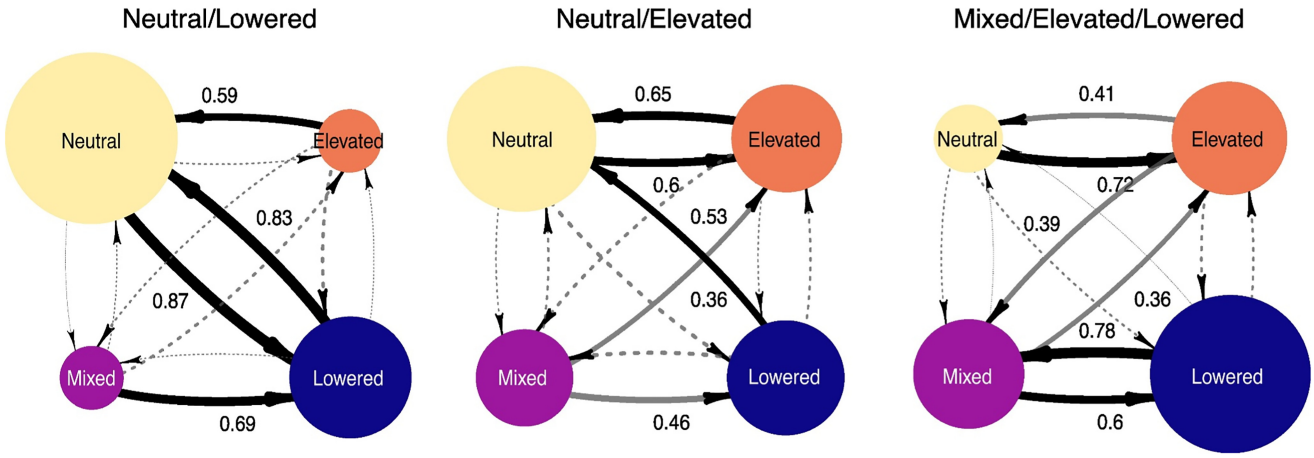
Alignment of Mood States to Validated Symptom Questionnaires

Average ASRM and QIDS scores aggregated by state on the full sample support the classification of the four mood states in neutral, elevated, mixed, and lowered. Overall, during momentary elevated states, ASRM manic scores were high, whereas during lowered states, QIDS depressive scores were high. The ASRM and QIDS scores were both low during neutral states (Table S2 in the online supplemental materials). The alignment between clinical questionnaires and momentary states is further supported by the linear mixed models, which revealed a significant effect of same-day mood state on the variability observed in both weekly scores (Table 2, Random coefficient model). For the ASRM weekly scores, only the elevated mood state showed a statistically significant higher ASRM score compared to the reference category neutral. Notably, for the QIDS weekly scores, both mixed and lowered mood states showed a significantly higher QIDS score compared to the reference category neutral. Mood states in the model explain 22% and 27% of the variability in ASRM and QIDS scores observed within patients, respectively.

Importantly, the linear mixed models revealed that alignment was best when the ASRM and QIDS were compared to momentary states assessed on the same day as the two clinical questionnaires. When the ASRM and QIDS were compared to aggregated momentary states of the 3 and 6 days prior to the completion of the weekly questionnaires, alignment was worse (see Tables S5 and S6 in the online supplemental materials). For these two windows, the elevated mood state did not exhibit a statistically significant effect on the ASRM score, although the elevated mood state showed a significantly higher QIDS score compared to the reference category neutral.

Figure 4

Diagrams of the Three Switching Patterns Between Mood States Described Based on Their Relative Switching Probabilities



Note. The thickness of the links is proportional to the size of the relative switching probability. Dashed lines indicate relative probabilities <.3, gray line relative probabilities on the range .3–.5, and black line relative probabilities >.5. Only relative switching probabilities >.3 have been labeled. The area of the circles is proportional to the incidence of each mood state in the corresponding pattern. See the online article for the color version of this figure.

Finally, Figure 5 visualizes that, overall and for the three mood patterns, there is good temporal alignment between momentary mood states and validated questionnaires (e.g., elevated manic symptoms on the ASRM co-occurred with elevated states). However, for eight patients, the uncovered states based on EMA data reveal more mood instability (i.e., switching between states) than the weekly symptom questionnaires suggest, as evidenced by a much more erratic "barcode" (e.g., Patients 1, 11, and 12; see Table S3 in the online supplemental materials for a quantification of switches).

Discussion

This exploratory study used MHMM and high-frequency EMA data to examine mood instability in 20 patients with BD. We found evidence of four momentary mood states: neutral, lowered, elevated, and mixed. Results demonstrate that patients in our sample switched more frequently between mood states than indicated by validated weekly self-report symptom questionnaires: patients displayed high mood instability, even in periods classified as euthymic by the symptom questionnaires. On average, patients only remained in the same mood state for 25 hr at maximum. We also showed that elevated and lowered mood states were interspersed by mixed states (with symptoms of both poles present) and that three patterns of switching between mood states could be identified. These results offer new insights into the dynamics of mood in BD and, if replicated in larger samples, suggest implications for the conceptualization of the disorder, as well as a higher frequency of mood assessment.

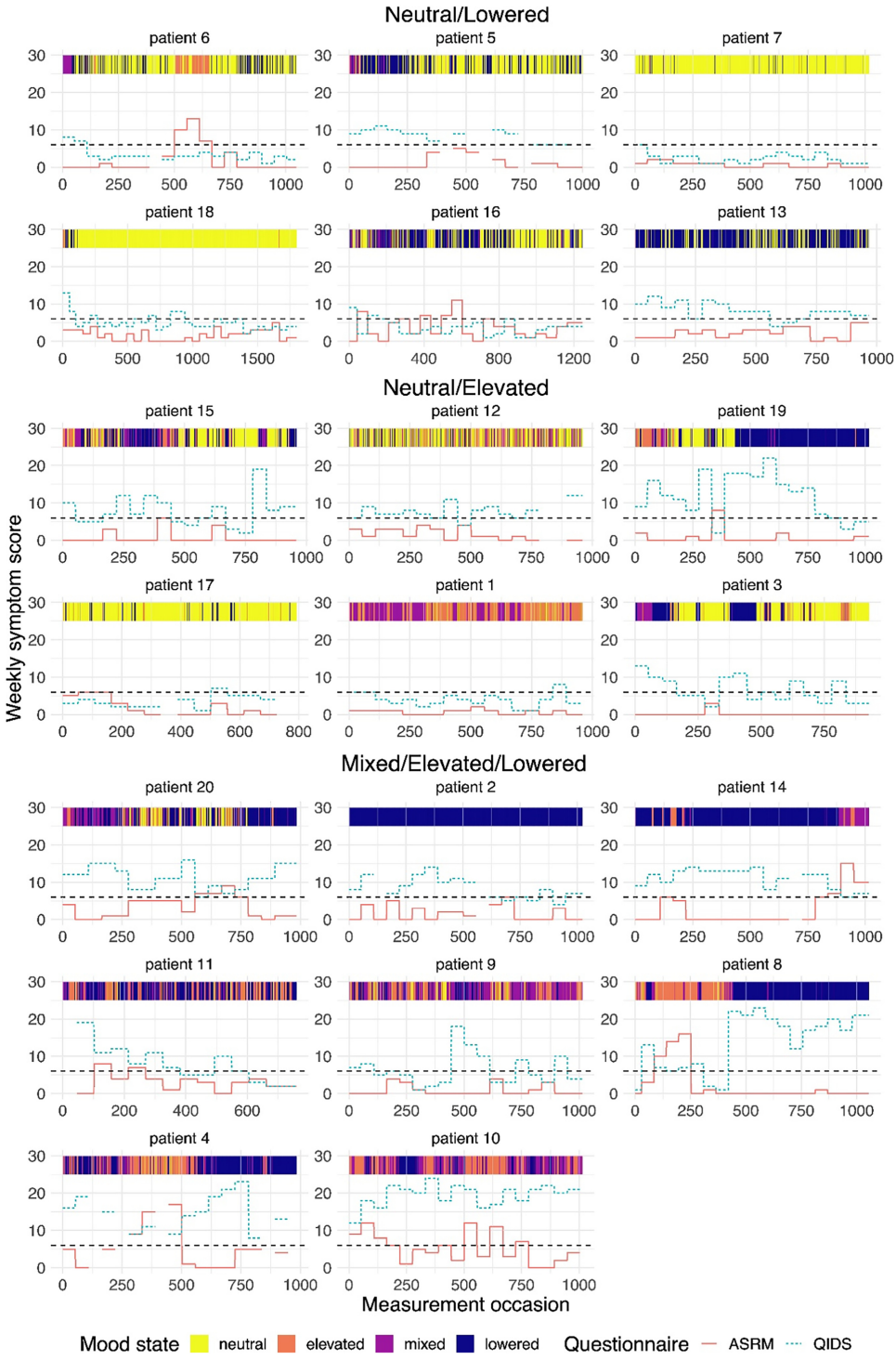
Our study extends findings from studies also using HMM on symptom questionnaire data in larger samples but with fewer data points per individual (Cochran et al., 2016; Prisciandaro et al., 2019; Yee et al., 2021). We uncovered similar mood states as in these studies, which suggests that the switches between manic, depressive, and mixed episodes so far only observed at the group level over weeks and months are also observed for individual patients and within days. Furthermore, in line with these studies, we found that the neutral and lowered states were relatively more prevalent and prolonged than the elevated and mixed states. Our results further correspond to studies using more frequent methods of assessment showing the key role of mood instability in BD, both day-to-day (Faurholt-Jepsen et al., 2019b, 2023; Stanislaus et al., 2020) and within days (Lamers et al., 2018; Sperry et al., 2020).

The design of our study allowed us to compare the fine-grained (3-hourly) EMA-derived mood states to weekly administered validated symptom questionnaires, taken here to reflect clinical episodes, which suggested that mood instability is more prevalent than is suggested by clinical questionnaires. Indeed, almost half of the patients in this sample (eight out of 20) would be characterized as stable by the ASRM and QIDS but demonstrated frequent switching between states within days. Such ultradian cycling has been observed before (Kramlinger & Post, 1996), but the extent to which patients experience this was thus far unknown (Carvalho et al., 2014). In our sample, mood states lasted shorter than 25 hr, indicating mood instability is frequent and present during and in between discrete manic and depressive episodes. Furthermore, the ASRM and QIDS were most closely related to momentary states assessed on the same day as the two clinical questionnaires versus 3 or 6 days prior. Taken together, this supports the hypothesis that, indeed, mood instability is a key feature of BD (Bonsall et al., 2012; Gottschalk et al., 1995; Hadaeghi et al., 2013).

The present study further provides insight into how patients might switch between states. Such hourly patterns have been theorized to, over time, evolve into more prolonged clinical episodes (Wichers et al., 2021). We identified three patterns in dominant feedback loops: patients switching predominantly between (a) neutral and lowered states, (b) neutral and elevated states, and (c) mixed, elevated, and lowered states, with almost no occurrence of neutral states. Previous work has also detected different patterns of mood instability for bipolar Type II during depressive versus manic episodes (Faurholt-Jepsen et al., 2019a). Notably, we found that depressive episodes (above cutoff scores on the QIDS) were associated not only with momentary lowered states but also with elevated and mixed states. Relatedly, patients exhibited higher scores on the

Figure 5

Uncovered Mood States Over Time and Temporal Correspondence Between Manic and Depressive Scores in Weekly Validated Questionnaires (ASRM and QIDS) Over the 4-Month Measurement Period for the 20 Patients in the Sample, Grouped by Switching Patterns



Note. Lines represent weekly scores of ASRM and QIDS, respectively. The uncovered momentary mood states are displayed on top of each panel. ASRM = Altman Self-Rating Mania; QIDS = Quick Inventory for Depressive Symptomatology. See the online article for the color version of this figure.

Table 2

Results of Linear Mixed Model Investigating the Relationship Between Mood States and Weekly Symptom Scores on the Same Day of Weekly Assessment

	ASRM			QIDS		
Model	Random intercept only	Fixed effect state	Random coefficients	Random intercept only	Fixed effect state	Random coefficients
Fixed effects						
Intercept	2.18*** (0.33)	2.14*** (0.34)	2.10*** (0.32)	8.48*** (0.93)	7.79*** (0.90)	6.80*** (0.65)
Elevated state		1.79*** (0.19)	1.60* (0.63)		−1.10*** (0.25)	0.60 (0.46)
Mixed state		0.32 (0.20)	0.29 (0.68)		1.42*** (0.26)	3.23** (0.95)
Lowered state		−0.79*** (0.15)	−0.44† (0.24)		1.91*** (0.20)	2.59* (0.91)
Random effects						
Residual	7.39 (2.72)	6.85 (2.62)	5.73 (2.39)	12.51 (3.54)	11.57 (3.40)	9.10 (3.02)
Patient	2.07 (1.44)	2.10 (1.45)	1.74 (1.32)	17.36 (4.17)	15.82 (3.98)	7.71 (2.78)
Elevated state			5.18 (2.28)			2.27 (1.51)
Mixed state			6.79 (2.61)			13.58 (3.69)
Lowered state			0.46 (2.39)			14.73 (3.84)
Explained variance within R^2		.07	.22		.08	.27

Note. Neutral state used as the reference category. Parentheses denote standard deviation of the estimator. ASRM = Altman Self-Rating Mania; QIDS = Quick Inventory for Depressive Symptomatology.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

EMA items "agitated" and "irritated" in the lowered state when compared to the neutral state. This corresponds to other studies demonstrating that bipolar depression is sometimes characterized by heightened arousal and irritation (Deckersbach et al., 2004; Faurholt-Jepsen et al., 2023; Judd et al., 2012). In our study, the mixed state (with higher levels of depressive EMA items and irritation/agitation) also appeared to mediate switches between more prolonged elevated and lowered states. Switching directly between elevated and lowered states was rare, suggesting that mixed features may be key to understanding switching between the two poles (Bell et al., 2020).

Taken together, our results tentatively suggest that the current conceptualization of BD in the *DSM-5* as a phenomenon characterized by mostly discrete episodes may not fully reflect the changing dynamic of mood in all patients and does not take into account the presence of frequent mood instability between major episodes. Classifying such mood instability is highly relevant as it is experienced as invalidating by patients (Crowe et al., 2012) and associated with worse outcomes, both on symptomatic and functional measures (Carvalho et al., 2014; Pallaskorpi et al., 2017; Sperry et al., 2020). It may also be informative for treatment. For example, antidepressant medication has been suggested to induce cycle acceleration and worsen manic symptoms (Carvalho et al., 2014) and is therefore not recommended for patients with mixed episodes (Fagiolini et al., 2015). Unfortunately, the limited number of patients in the current study precluded us from examining the effect of covariates on the uncovered mood dynamics and relating them to clinical outcomes. Future research should compare mood instability to a control group and examine whether different patterns of mood instability predict clinical outcomes. If so, future *DSM* classifications of BD might include mood instability as a specifier to facilitate timely diagnosis and adequate intervention (Bonsall et al., 2012).

The study is innovative in several ways. First, this study is the first to use intensive longitudinal EMA data to study mood instability in hourly assessed bipolar constructs. This enables the study of mood switches with more reliability and a higher temporal resolution than so far observed in the literature, addressing the limitations of retrospective reporting, such as recall bias (Ariens et al., 2020). Second, by comparing the alignment of EMA-based mood states and weekly symptom data, we are uniquely able to compare how momentary mood states relate to more prolonged clinical episodes. Third, the MHMM presented here offers an optimal balance, summarizing patterns shared among patients (i.e., group level) while elucidating individual-specific switching dynamics.

The innovative nature of the study warrants exploration of its limitations, relating to both clinical and methodological factors. First, the limited number of participants restricts the generalizability of our results. Importantly, the total sample size (Number of Participants × Number of Assessments) is comparable with previous studies that explore switches in latent mood states of patients with BD (e.g., Cochran et al., 2016; Prisciandaro et al., 2019; Yee et al., 2021). Indeed, previous research in MHMMs has demonstrated that parameters are accurately estimated with smaller samples when the number of assessments is sufficiently large (McClintock, 2021; Mildiner Moraga & Aarts, 2024). However, clinically, the sample was small and quite heterogeneous (e.g., both BD Type I and II, males and females, with different comorbidities and medication types), and participants were selected based on a clinical history of frequent manic or depressive episodes. It is likely that mood instability may differ across these factors (e.g., Faurholt-Jepsen et al., 2019a), and more studies are necessary to replicate our findings. Future research may study in further detail how other factors, such as medication use and comorbid (personality) disorders, may influence mood dynamics over the course of the day. Based on previous literature, we might expect that antidepressant medication use and the presence of comorbid borderline personality disorder may induce more switching, different switching patterns, and more mixed states (Carvalho et al., 2014; Reich et al., 2012; Santangelo et al., 2017). Likewise, our lack of a healthy control group precluded us from gaining insight into how the uncovered mood dynamics deviate from individuals without BD. Finally, we adopted the common practice in discrete-time models of treating overnight measures as missing data, since ignoring the longer overnight time gap between consecutive days would introduce bias in switching probabilities and state durations.

Although we found no deviations from MAR, the assumption is theoretically challenging to uphold for mood-related data. As a result, the findings of this study should be interpreted with caution until they are replicated and be viewed as suggestive of what an MHMM could reveal when paired with longitudinal EMA data.

Conclusion

To conclude, this exploratory study showed that mood states derived from longitudinal EMA data demonstrate frequent mood instability in BD in between and during episodes. On average, mood states lasted for less than 25 hr. The MHMM could reveal the dynamics of mood instability and visualize and quantify mood patterns over the span of a few hours, showing that the mixed state often mediated switches between elevated and lowered states. Our identification of distinct mood-switching patterns highlights the relevance of considering individual-specific dynamics. Taken together, this tentatively suggests that mood instability is a key feature of BD, which should be considered in theoretical as well as clinical conceptualizations of the disorder. Such a personalized approach has the potential to improve the care of patients with BD on a very individual scale.

References

Aarts, E. (2019). *mHMMbayes: Multilevel hidden Markov models using Bayesian estimation* (0.1.1.9002) [R]. https://cran.r-project.org/web/packages/mHMMbayes/mHMMbayes.pdf

Abreu, T., & Bragança, M. (2015). The bipolarity of light and dark: A review on bipolar disorder and circadian cycles. *Journal of Affective Disorders*, *185*, 219–229. https://doi.org/10.1016/j.jad.2015.07.017

Altman, E., Hedeker, D., Peterson, J. L., & Davis, J. M. (2001). A comparative evaluation of three self-rating scales for acute mania. *Biological Psychiatry*, *50*(6), 468–471. https://doi.org/10.1016/s0006-3223(01)01065-4

Altman, E. G., Hedeker, D., Peterson, J. L., & Davis, J. M. (1997). The Altman Self-Rating Mania Scale. *Biological Psychiatry*, *42*(10), 948–955. https://doi.org/10.1016/S0006-3223(96)00548-3

Altman, R. M. K. (2007). Mixed hidden Markov models: An extension of the hidden Markov model to the longitudinal data setting. *Journal of the American Statistical Association*, *102*(477), 201–210. https://doi.org/10.1198/016214506000001086

Ariens, S., Ceulemans, E., & Adolf, J. K. (2020). Time series analysis of intensive longitudinal data in psychosomatic research: A methodological overview. *Journal of Psychosomatic Research*, *137*, Article 110191. https://doi.org/10.1016/j.jpsychores.2020.110191

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 359–388. https://doi.org/10.1080/10705511.2017.1406803

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bell, E., Boyce, P., Porter, R. J., Bryant, R. A., & Malhi, G. S. (2020). Could irritability be the key to unlocking the enigma of mixed states? *Bipolar Disorders*, *22*(8), 781–784. https://doi.org/10.1111/bdi.13029

Bernstein, I. H., Rush, A. J., Suppes, T., Kyotoku, Y., & Warden, D. (2010). The Quick Inventory of Depressive Symptomatology (Clinician and self-report versions) in patients with bipolar disorder. *CNS Spectrums*, *15*(6), 367–373. https://doi.org/10.1017/S1092852900029230

Bonsall, M. B., Wallace-Hadrill, S. M. A., Geddes, J. R., Goodwin, G. M., & Holmes, E. A. (2012). Nonlinear time-series approaches in characterizing mood stability and mood instability in bipolar disorder. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1730), 916–924. https://doi.org/10.1098/rspb.2011.1246

Bos, F. M., Schreuder, M. J., George, S. V., Doornbos, B., Bruggeman, R., van der Krieke, L., Haarman, B. C. M., Wichers, M., & Snippe, E. (2022). Anticipating manic and depressive transitions in patients with bipolar disorder using early warning signals. *International Journal of Bipolar Disorders*, *10*(1), Article 12. https://doi.org/10.1186/s40345-022-00258-4

Bos, F. M., Snippe, E., Bruggeman, R., Doornbos, B., Wichers, M., & van der Krieke, L. (2020). Recommendations for the use of long-term experience sampling in bipolar disorder care: A qualitative study of patient and clinician experiences. *International Journal of Bipolar Disorders*, *8*(1), Article 38. https://doi.org/10.1186/s40345-020-00201-5

Carvalho, A. F., Dimellis, D., Gonda, X., Vieta, E., McIntyre, R. S., & Fountoulakis, K. N. (2014). Rapid cycling in bipolar disorder: A systematic review. *The Journal of Clinical Psychiatry*, *75*(06), e578–e586. https://doi.org/10.4088/JCP.13r08905

Cerimele, J. M., Goldberg, S. B., Miller, C. J., Gabrielson, S. W., & Fortney, J. C. (2019). Systematic review of symptom assessment measures for use in measurement-based care of bipolar disorders. *Psychiatric Services (Washington, D.C.)*, *70*(5), 396–408. https://doi.org/10.1176/appi.ps.201800383

Cochran, A. L., McInnis, M. G., & Forger, D. B. (2016). Data-driven classification of bipolar I disorder from longitudinal course of mood. *Translational Psychiatry*, *6*(10), Article e912. https://doi.org/10.1038/tp.2016.166

Crowe, M., Inder, M., Carlyle, D., Wilson, L., Whitehead, L., Panckhurst, A., O'Brien, A., & Joyce, P. (2012). Feeling out of control: A qualitative analysis of the impact of bipolar disorder. *Journal of Psychiatric and Mental Health Nursing*, *19*(4), 294–302. https://doi.org/10.1111/j.1365-2850.2011.01786.x

Deckersbach, T., Perlis, R. H., Frankle, W. G., Gray, S. M., Grandin, L., Dougherty, D. D., Nierenberg, A. A., & Sachs, G. S. (2004). Presence of irritability during depressive episodes in bipolar disorder. *CNS Spectrums*, *9*(3), 227–231. https://doi.org/10.1017/s1092852900009020

Fagiolini, A., Coluccia, A., Maina, G., Forgione, R. N., Goracci, A., Cuomo, A., & Young, A. H. (2015). Diagnosis, epidemiology and management of mixed states in bipolar disorder. *CNS Drugs*, *29*(9), 725–740. https://doi.org/10.1007/s40263-015-0275-6

Faurholt-Jepsen, M., Busk, J., Bardram, J. E., Stanislaus, S., Frost, M., Christensen, E. M., Vinberg, M., & Kessing, L. V. (2023). Mood instability and activity/energy instability in patients with bipolar disorder according to day-to-day smartphone-based data—An exploratory post hoc study. *Journal of Affective Disorders*, *334*, 83–91. https://doi.org/10.1016/j.jad.2023.04.139

Faurholt-Jepsen, M., Frost, M., Busk, J., Christensen, E. M., Bardram, J. E., Vinberg, M., & Kessing, L. V. (2019a). Differences in mood instability in patients with bipolar disorder type I and II: A smartphone-based study. *International Journal of Bipolar Disorders*, *7*(1), Article 5. https://doi.org/10.1186/s40345-019-0141-4

Faurholt-Jepsen, M., Frost, M., Busk, J., Christensen, E. M., Bardram, J. E., Vinberg, M., & Kessing, L. V. (2019b). Is smartphone-based mood instability associated with stress, quality of life, and functioning in bipolar disorder? *Bipolar Disorders*, *21*(7), 611–620. https://doi.org/10.1111/bdi.12796

Forney, G. D. (1973). The Viterbi algorithm. *Proceedings of the IEEE*, *61*(3), 268–278. https://doi.org/10.1109/PROC.1973.9030

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*(1), 72–91. https://doi.org/10.1037/a0032138

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/b16018

Gottschalk, A., Bauer, M. S., & Whybrow, P. C. (1995). Evidence of chaotic mood variation in bipolar disorder. *Archives of General Psychiatry*, *52*(11), 947–959. https://doi.org/10.1001/archpsyc.1995.03950230061009

Hadaeghi, F., Hashemi Golpayegani, M. R., & Moradi, K. (2013). Does "crisis-induced intermittency" explain bipolar disorder dynamics?

Frontiers in Computational Neuroscience, *7*, Article 116. https://doi.org/10.3389/fncom.2013.00116

Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research*, *53*(6), 820–841. https://doi.org/10.1080/00273171.2018.1446819

Harrison, P. J., Geddes, J. R., & Tunbridge, E. M. (2018). The emerging neurobiology of bipolar disorder. *Trends in Neurosciences*, *41*(1), 18–30. https://doi.org/10.1016/j.tins.2017.10.006

Judd, L. L., Schettler, P. J., Akiskal, H., Coryell, W., Fawcett, J., Fiedorowicz, J. G., Solomon, D. A., & Keller, M. B. (2012). Prevalence and clinical significance of subsyndromal manic symptoms, including irritability and psychomotor agitation, during bipolar major depressive episodes. *Journal of Affective Disorders*, *138*(3), 440–448. https://doi.org/10.1016/j.jad.2011.12.046

Knapen, S. E. (2019). *Rhythm & Blues: Chronobiology in the pathophysiology and treatment of mood disorders* [Thesis fully internal (DIV)]. Rijksuniversiteit Groningen.

Kramlinger, K. G., & Post, R. M. (1996). Ultra-rapid and ultradian cycling in bipolar affective illness. *The British Journal of Psychiatry: The Journal of Mental Science*, *168*(3), 314–323. https://doi.org/10.1192/bjp.168.3.314

Krieke, L. V. D., Jeronimus, B. F., Blaauw, F. J., Wanders, R. B. K., Emerencia, A. C., Schenk, H. M., Vos, S. D., Snippe, E., Wichers, M., Wigman, J. T. W., Bos, E. H., Wardenaar, K. J., & Jonge, P. D. (2016). Hownutsarethedutch (HoeGekIsNL): A crowdsourcing study of mental symptoms and strengths. *International Journal of Methods in Psychiatric Research*, *25*(2), 123–144. https://doi.org/10.1002/mpr.1495

Lamers, F., Swendsen, J., Cui, L., Husky, M., Johns, J., Zipunnikov, V., & Merikangas, K. R. (2018). Mood reactivity and affective dynamics in mood and anxiety disorders. *Journal of Abnormal Psychology*, *127*(7), 659–669. https://doi.org/10.1037/abn0000378

Leboyer, M., & Kupfer, D. J. (2010). Bipolar disorder: New perspectives in health care and prevention. *The Journal of Clinical Psychiatry*, *71*(12), 1689–1695. https://doi.org/10.4088/JCP.10m06347yel

Little, R. J. A., & Rubin, D. B. (2019). *Statistical analysis with missing data*. John Wiley & Sons.

McClintock, B. T. (2021). Worth the effort? A practical examination of random effects in hidden Markov models for animal telemetry data. *Methods in Ecology and Evolution*, *12*(8), 1475–1497. https://doi.org/10.1111/2041-210X.13619

McInnis, M. G., Andreassen, O. A., Andreazza, A. C., Alon, U., Berk, M., Brister, T., Burdick, K. E., Cui, D., Frye, M., Leboyer, M., Mitchell, P. B., Merikangas, K., Nierenberg, A. A., Nurnberger, J. I., Pham, D., Vieta, E., Yatham, L. N., & Young, A. H. (2022). Strategies and foundations for scientific discovery in longitudinal studies of bipolar disorder. *Bipolar Disorders*, *24*(5), 499–508. https://doi.org/10.1111/bdi.13198

Meyer, T. D., Crist, N., La Rosa, N., Ye, B., Soares, J. C., & Bauer, I. E. (2020). Are existing self-ratings of acute manic symptoms in adults reliable and valid?-A systematic review. *Bipolar Disorders*, *22*(6), 558–568. https://doi.org/10.1111/bdi.12906

Mildiner Moraga, S., & Aarts, E. (2024). Go multivariate: Recommendations on Bayesian multilevel hidden Markov models with categorical data. *Multivariate Behavioral Research*, *59*(1), 17–45. https://doi.org/10.1080/00273171.2023.2205392

Miller, C. J., Johnson, S. L., & Eisner, L. (2009). Assessment tools for adult bipolar disorder. *Clinical Psychology: A Publication of the Division of Clinical Psychology of the American Psychological Association*, *16*(2), 188–201. https://doi.org/10.1111/j.1468-2850.2009.01158.x

Pallaskorpi, S., Suominen, K., Ketokivi, M., Valtonen, H., Arvilommi, P., Mantere, O., Leppämäki, S., & Isometsä, E. (2017). Incidence and predictors of suicide attempts in bipolar I and II disorders: A 5-year follow-up study. *Bipolar Disorders*, *19*(1), 13–22. https://doi.org/10.1111/bdi.12464

Pohle, J., Langrock, R., van Beest, F. M., & Schmidt, N. M. (2017). Selecting the number of states in hidden Markov models: Pragmatic solutions illustrated using animal movement. *Journal of Agricultural, Biological, and Environmental Statistics*, *22*(3), 270–293. https://doi.org/10.1007/s13253-017-0283-8

Prisciandaro, J. J., Tolliver, B. K., & DeSantis, S. M. (2019). Identification and initial validation of empirically derived bipolar symptom states from a large longitudinal dataset: An application of hidden Markov modeling to the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) study. *Psychological Medicine*, *49*(07), 1102–1108. https://doi.org/10.1017/S0033291718002143

R Core Team. (2021). *R: A language and environment for statistical computing*. R Development Core Team. https://www.R-project.org

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–286. https://doi.org/10.1109/5.18626

Reich, D. B., Zanarini, M. C., & Fitzmaurice, G. (2012). Affective lability in bipolar disorder and borderline personality disorder. *Comprehensive Psychiatry*, *53*(3), 230–237. https://doi.org/10.1016/j.comppsych.2011.04.003

Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., Markowitz, J. C., Ninan, P. T., Kornstein, S., Manber, R., Thase, M. E., Kocsis, J. H., & Keller, M. B. (2003). 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. *Biological Psychiatry*, *54*(5), 573–583. https://doi.org/10.1016/S0006-3223(02)01866-8

Santangelo, P. S., Reinhard, I., Koudela-Hamila, S., Bohus, M., Holtmann, J., Eid, M., & Ebner-Priemer, U. W. (2017). The temporal interplay of self-esteem instability and affective instability in borderline personality disorder patients' everyday lives. *Journal of Abnormal Psychology*, *126*(8), 1057–1065. https://doi.org/10.1037/abn0000288

Shappell, H., Caffo, B. S., Pekar, J. J., & Lindquist, M. A. (2019). Improved state change estimation in dynamic functional connectivity using hidden semi-Markov models. *NeuroImage*, *191*, 243–257. https://doi.org/10.1016/j.neuroimage.2019.02.013

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*(1), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Shirley, K. E., Small, D. S., Lynch, K. G., Maisto, S. A., & Oslin, D. W. (2010). Hidden Markov models for alcoholism treatment trial data. *Annals of Applied Statistics*, *4*(1), 366–395. https://doi.org/10.1214/09-AOAS282

Simon, J., Budge, K., Price, J., Goodwin, G. M., & Geddes, J. R. (2017). Remote mood monitoring for adults with bipolar disorder: An explorative study of compliance and impact on mental health service use and costs. *European Psychiatry: the Journal of the Association of European Psychiatrists*, *45*, 14–19. https://doi.org/10.1016/j.eurpsy.2017.06.007

Sperry, S. H., & Kwapil, T. R. (2019). Affective dynamics in bipolar spectrum psychopathology: Modeling inertia, reactivity, variability, and instability in daily life. *Journal of Affective Disorders*, *251*, 195–204. https://doi.org/10.1016/j.jad.2019.01.053

Sperry, S. H., & Kwapil, T. R. (2022). Bipolar spectrum psychopathology is associated with altered emotion dynamics across multiple timescales. *Emotion (Washington, D.C.)*, *22*(4), 627–640. https://doi.org/10.1037/emo0000759

Sperry, S. H., Walsh, M. A., & Kwapil, T. R. (2020). Emotion dynamics concurrently and prospectively predict mood psychopathology. *Journal of Affective Disorders*, *261*, 67–75. https://doi.org/10.1016/j.jad.2019.09.076

Stanislaus, S., Faurholt-Jepsen, M., Vinberg, M., Coello, K., Kjærstad, H. L., Melbye, S., Sletved, K. S. O., Christensen, E. M., Frost, M., Bardram, J. E.,

& Kessing, L. V. (2020). Mood instability in patients with newly diagnosed bipolar disorder, unaffected relatives, and healthy control individuals measured daily using smartphones. *Journal of Affective Disorders*, *271*, 336–344. https://doi.org/10.1016/j.jad.2020.03.049

Tyler, E., Jones, S., Black, N., Carter, L.-A., & Barrowclough, C. (2015). The relationship between bipolar disorder and cannabis use in daily life: An experience sampling study. *PLOS ONE*, *10*(3), Article e0118916. https://doi.org/10.1371/journal.pone.0118916

Wichers, M., Riese, H., Hodges, T. M., Snippe, E., & Bos, F. M. (2021). A narrative review of network studies in depression: What different methodological approaches tell us about depression. *Frontiers in Psychiatry*, *12*, Article 719490. https://doi.org/10.3389/fpsyt.2021.719490

Yee, M. A., Yocum, A. K., McInnis, M. G., & Cochran, A. L. (2021). Dynamics of data-driven microstates in bipolar disorder. *Journal of Psychiatric Research*, *141*, 370–377. https://doi.org/10.1016/j.jpsychires.2021.07.021

Zhang, Y., & Berhane, K. (2014). Bayesian Mixed hidden Markov models: A multi-level approach to modeling categorical outcomes with differential misclassification. *Statistics in Medicine*, *33*(8), 1395–1408. https://doi.org/10.1002/sim.6039

Zucchini, W., Macdonald, I. L., & Langrock, R. (2017). *Hidden Markov models for time series: An introduction using R* (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/b20790

Received August 2, 2023
Revision received January 30, 2024
Accepted March 11, 2024 ■