

Original Article

Cite this article: Prisciandaro JJ, Tolliver BK, DeSantis SM (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**, 1102–1108. <https://doi.org/10.1017/S0033291718002143>

Received: 9 February 2018
Revised: 16 May 2018
Accepted: 24 July 2018
First published online: 29 August 2018

Key words:

Bipolar disorder; hidden Markov model; latent variable modeling; longitudinal; mixed episode; mixed features; systematic treatment enhancement program for bipolar disorder

Author for correspondence:

James J. Prisciandaro,
E-mail: priscian@musc.edu

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

James J. Prisciandaro¹, Bryan K. Tolliver¹ and Stacia M. DeSantis²

¹Department of Psychiatry and Behavioral Sciences, Medical University of SC, Charleston, SC, USA and ²School of Public Health, University of Texas Health Science Center, Houston, TX, USA

Abstract

Background. Although bipolar disorder (BD) is a fundamentally cyclical illness, a divided model of BD that emphasizes polarity over cyclicity has dominated modern psychiatric diagnostic systems since their advent in the 1980s. However, there has been a gradual return to conceptualizations of BD which focus on longitudinal course in the research community due to emerging supportive data. Advances in longitudinal statistical methods promise to further progress the field.

Methods. The current study employed hidden Markov modeling to uncover empirically derived manic and depressive states from longitudinal data [i.e. Young Mania Rating Scale and Montgomery–Asberg Depression Rating Scale responses across five occasions from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) study], estimate participants' probabilities of transitioning between these states over time ($n = 3918$), and evaluate whether clinical variables (e.g. rapid cycling and substance dependence) predict participants' state transitions ($n = 3229$).

Results. Analyses identified three empirically derived mood states ('euthymic,' 'depressed,' and 'mixed'). Relative to the euthymic and depressed states, the mixed state was less commonly experienced, more temporally unstable, and uniquely associated with rapid cycling, substance use, and psychosis. Individuals assigned to the mixed state at baseline were relatively less likely to be diagnosed with BD-II (*v.* BD-I), more likely to present with a mixed or (hypo) manic episode, and reported experiencing irritable and elevated mood more frequently.

Conclusions. The results from the current study represent an important step in defining, and characterizing the longitudinal course of, empirically derived mood states that can be used to form the foundation of objective, empirical attempts to define meaningful subtypes of affective illness defined by clinical course.

Bipolar disorder (BD) is a fundamentally cyclical illness, not only in terms of its prototypical clinical course, characterized by alternating highs and lows of mood, energy, and behavior, but also in terms of its hypothesized neurobiological underpinnings (Abreu and Braganca, 2015). Credit for first describing BD (then named, 'folie circulaire') as a single, fundamentally cyclical disease entity is generally given to Falret in 1851 (Falret, 1851). Falret's conceptualization of folie circulaire was subsequently expanded on by later prominent psychiatrists, including Kraepelin with his renowned observation-based studies of 'manic-depressive illness' (Kraepelin, 1921), and generally dominated thinking on the matter until the 1960s, when Angst and others called for the division of manic-depressive illness into 'unipolar' (i.e. depression-only) and 'bipolar' (i.e. mania with or without depression) types based on emerging heritability data (Angst, 1966). This model of affective illness prioritized polarity over cyclicity, featuring unipolar *v.* bipolar types each defined by the lifetime occurrence of specific combinations of discrete mood episodes (DSM-III) (American Psychiatric Association, 1980), informed the development of the pre-eminent psychiatric diagnostic systems of the day that have managed to persist, largely unchanged, to the present (DSM-5) (American Psychiatric Association, 2013).

Nonetheless, since the 1980s, there has been a gradual return to Falret and Kraepelin's conceptualizations of affective illness in the BD research community (Marneros and Angst, 2001; Goodwin and Jamison, 2007), which focus on longitudinal course, cyclicity, and continuous 'spectrum models,' due to emerging supportive data (Akiskal, 1983; Prisciandaro and Roberts, 2005; 2011). For example, it is now recognized that even individuals with 'pure'

major depressive disorder (MDD) experience higher rates of sub-threshold hypomania than previously recognized (Zimmermann *et al.*, 2009; Angst *et al.*, 2010, 2011; Fiedorowicz *et al.*, 2011; Perugi *et al.*, 2015). As a result, the characterization of mixed states and/or features was significantly modified in DSM-V compared with previous editions. Increasingly, sophisticated new longitudinal statistical methods have been developed that allow investigation of syndromal and subsyndromal mood fluctuations over the course of illness of affective disorders (Cochran *et al.*, 2016). The current study employed one such statistical framework, hidden Markov modeling (HMM) (MacDonald and Zucchini, 1997), to uncover empirically derived manic and depressive symptom states from longitudinal data from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) study (Sachs *et al.*, 2003), estimate participants' probabilities of transitioning between these states over time, and evaluate whether clinical course variables and/or comorbid psychiatric disorders predict participants' state transitions.

Materials and methods

Study overview

The design and methods of the STEP-BD study have been described in detail previously (Sachs *et al.*, 2003). Briefly, STEP-BD was a multicenter naturalistic treatment study conducted at 21 US academic medical centers between 1998 and 2005 that prospectively evaluated clinical outcomes in patients with BD. Unlike most clinical trials that selectively enroll subjects in specific mood states and exclude individuals with substance abuse and other psychiatric comorbidities, STEP-BD was designed to study treatment response and clinical course in patients commonly encountered in clinical practice, regardless of current mood status, treatment regimen, or co-occurring conditions. The study allowed all participants to enter a standard care pathway in which clinic visits were scheduled as frequently as indicated and clinicians were encouraged to apply expert consensus guidelines for treatment but were free to prescribe any treatment they felt was indicated. Patients could also become eligible to voluntarily enter randomized controlled medication treatment for acute ($n = 366$) and refractory ($n = 66$) depression, or psychosocial treatment for acute depression ($n = 293$), but could return to the standard care pathway at any time. Over 7 years, 4360 participants aged 15 years and older who met DSM-IV criteria for bipolar I disorder, bipolar II disorder, BD not otherwise specified, cyclothymia, or schizoaffective disorder, bipolar type were enrolled in the STEP-BD study. Four thousand one hundred and eight of these participants (i.e. all participants enrolled through 30 November 2004 and scheduled for follow-up visits through 28 February 2005) were included in the publicly available STEP-BD database used for the current analyses. Of these available participants, 3918 were included in analyses because they had complete data on at least one of the dependent variables on at least one occasion. However, only 3229 participants were included in analyses involving baseline covariates, because the remaining 689 participants had missing data on at least one covariate.

Assessments

During their first year of participation, study patients completed quarterly independent assessment interviews at baseline and at months 3, 6, 9, and 12. Bipolar diagnoses and retrospective course

(e.g. history of suicide attempt, history of psychosis, history of antidepressant-associated switch, past year rapid cycling, and age of BD onset) were assessed by study clinicians at baseline using the Affective Disorders Evaluation (Sachs *et al.*, 2003), adapted from the mood and psychosis modules of the Structured Clinical Interview for DSM-IV; comorbid Axis I diagnoses were assessed using the Mini-International Neuropsychiatric Interview (Sheehan *et al.*, 1998). Quarterly interviews also included evaluation of manic and depressive symptoms using clinician-administered scales [i.e. the 11-item Young Mania Rating Scale (YMRS) (Young *et al.*, 1978) and the 10-item Montgomery-Asberg Depression Rating Scale (MADRS)].

Statistical analysis

The current study used HMM (MacDonald and Zucchini, 1997), a latent longitudinal mixture modeling approach, to uncover empirically derived manic and depressive symptom states from longitudinal data, estimate participants' probabilities of transitioning between these states over time, and evaluate whether clinical course variables and/or comorbid psychiatric disorders predict participants' state transitions. HMM has been prominent in recent substance use disorder research because it readily accommodates the chaotic nature of substance use in clinical studies (i.e. consistent stretches of use interspersed with unusual bursts) (Shirley *et al.*, 2010) and reduces the impact of measurement error on parameter estimates and standard errors that originates from patients' self-report of symptomatology. Because these methodological challenges are prominent in mood symptom reports from individuals with BD, HMM is arguably well-suited to modeling longitudinal symptom data from bipolar patients (Prisciandaro *et al.*, 2012). Additionally, unlike traditional approaches to defining the boundaries of bipolar symptom states (e.g. expert consensus and latent class analysis of cross-sectional data), typically referred to as mood 'episodes,' HMM identifies latent mood states using longitudinal mood symptom data.

In the current study, manic and depressive symptoms, in terms of YMRS and MADRS total scores, were assumed to be manifestations of an empirically determined number of latent states, and participants were allowed to freely transit among these states across time. An optimal number of latent states was determined by comparing the relative fit [i.e. using sample-size adjusted Bayesian Information Criteria (aBIC)] (Nylund *et al.*, 2007) and interpretability of several HMM models, each featuring a different number of states. To ensure that HMMs would be estimable, interpretable, and reliable, we imposed two additional constraints on selecting the optimal number of latent states: (i) the selected number of latent states, the mean of each latent state, and the probabilities of transitioning among states had to be the same for every time-point, and (ii) each state had to contain at least 5% of the sample. The probability of remaining in, or transitioning to, a given latent state from one quarter to the next was calculated from the best fitting HMM model. Finally, the best-fitting HMM model was re-estimated with covariates, selected to facilitate construct validation of the obtained symptom states, allowed to predict participants' transition probabilities. Figure 1 presents a graphical depiction of the final estimated model (Wall and Li, 2009); M_{it} and Y_{it} represent the mean MADRS and YMRS scores, respectively, for participant i during study week t , C_{it} represents the unobserved mood state of patient i at study week t , and X_i represents a vector of baseline covariates (e.g. age of onset and history of suicide attempt) for patient i .

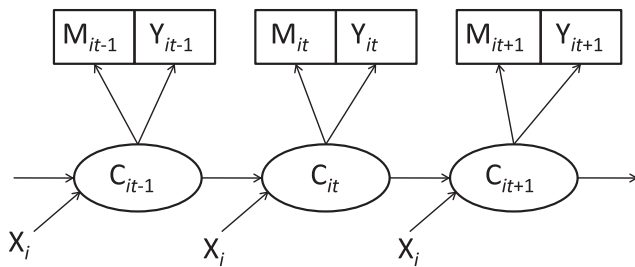


Fig 1. Hidden Markov model schematic. M_{it} and Y_{it} represent the mean MADRS and YMRS scores, respectively, for participant i during study week t , C_{it} represents the unobserved mood state of patient i at study week t , and X_i represents a vector of baseline covariates (e.g. age of onset) for patient i .

All models were estimated using an estimator (Maximum Likelihood with Robust standard errors; MLR), which uses all available data and yields unbiased parameters when missing data are missing at random, in MPlus 6.1 software (Muthen and Muthen, 2011); study site was included as a clustering variable for all models. Model parameterization of HMM as implemented in MPlus is described in detail in Langeheine and van de Pol (2002; pp. 323–329) as well as the MPlus user's guide (Example 8.12) (Muthen and Muthen, 2011).

Results

As described further below, HMM models optimally supported the existence of three latent mood states in the STEP-BD sample; baseline demographics and select clinical characteristics for the overall sample and for individuals in each of the latent states are presented in Table 1. Fit and estimated means for HMM models containing between 2 and 6 latent states are presented in Table 2. Although sample-size adjusted BIC suggested that the model with the largest estimable number of classes (i.e. 6) best fit the data, models with >3 latent states had at least one state containing <5% of the sample at some point in time. The three-state HMM consisted of 'euthymic' (i.e. MADRS $M=7.6$, YMRS $M=3.3$), 'depressed' (i.e. MADRS $M=23.6$, YMRS $M=6.7$), and 'mixed' (i.e. MADRS $M=20.1$, YMRS $M=20.0$) symptom states. Although all estimated HMM models contained 'euthymic' and 'depressed' states, additional states suggested by the four- through six-state models were further 'mixed' manifestations; for example, the four-state model, contained 'euthymic' and 'depressed' states, as well as two additional states with MADRS $M=22.0$, YMRS $M=23.7$ and MADRS $M=17.7$, YMRS $M=12.4$, respectively. In sum, a three-state HMM model was retained because it was maximally interpretable, fit better than an alternative two-state model, and consisted of classes that each contained $\geq 5\%$ of the sample.

The probability of transitioning between, or remaining in, latent states across time, along with the estimated proportion of participants classified in each state at each time point, are presented in Table 3. Membership in the 'euthymic' and 'depressed' states was prevalent and relatively stable over time; 94% of participants classified as euthymic at any given time point were classified as 'euthymic' at the next time point, and 72% of participants classified as 'depressed' were classified as 'depressed' at the next time point. Averaging across time, approximately 58% of the sample was classified as 'euthymic' and 32% was classified as 'depressed.' Conversely, membership in the 'mixed' state was limited (9% of the sample, on averaging across time) and highly

variable; approximately one-third of participants classified as 'mixed' at any given time point were classified as 'mixed' at the next time point, with the remaining two-thirds split roughly equally between 'depressed' and 'euthymic' states.

Given that the 'euthymic' and 'depressed' states were well-defined, stable, and conceptually clear, we primarily selected baseline covariates for the final HMM model that have been shown in past research to be associated with mixed episodes in order to clarify the conceptual nature of the obtained 'mixed' state. These variables included: history of suicide attempt, history of psychosis, history of antidepressant-associated switch, past year rapid cycling, age of BD onset, lifetime substance dependence, and panic disorder (Perlis *et al.*, 2004, 2010; Schneck *et al.*, 2008; Goldberg *et al.*, 2009; Ostacher *et al.*, 2010; Niitsu *et al.*, 2015); a wider range of covariates was not considered in order to maintain a reasonable experimentwise alpha level. Because mood symptoms were best represented by three latent states, associations between covariates and latent states were estimated in a multinomial logistic regression framework. As depicted in Fig. 1, reported covariate effects are independent of potentially correlated first order autoregressive effects. To obtain the full range of desired between-state comparisons, the conditional model was estimated twice, once with 'euthymic' as the reference state and again with 'depressed' as the reference state. Regression coefficients and odds ratios are presented in Table 4. As can be seen in Table 4, history of suicide attempt, past year rapid cycling, panic disorder, lifetime substance use disorder, and younger age of onset were all significantly associated with transitioning to or remaining in the 'mixed' v. 'euthymic' state. Most of these covariates (i.e. history of suicide attempt, past year rapid cycling, panic disorder, and lifetime substance use disorder) were also significantly associated with transitioning to or remaining in the 'depressed' v. 'euthymic' state. In contrast, history of psychosis was significantly associated with transitioning to or remaining in the 'euthymic' v. 'depressed' state. Importantly, however, three covariates (past year rapid cycling, lifetime substance use disorder, and history of psychosis) incrementally predicted transitioning to or remaining in the 'mixed' v. 'depressed' state. In sum, whereas history of suicide attempt, panic disorder, and age of BD onset similarly predicted transitioning to or remaining in the 'mixed' or 'depressed' state, past year rapid cycling, substance use disorder, and history of psychosis predicted transitioning to or remaining in the 'mixed' state significantly more strongly than they predicted transitioning to or remaining in the 'depressed' state.

Discussion

The current study used advanced longitudinal statistical modeling to simultaneously identify three latent mood states ('euthymic,' 'depressed,' and 'mixed') and characterize transitions among these mood states over time in 3229–3918 outpatients (depending on analysis) with BD. Unlike the euthymic and depressed latent mood states, the mixed state was relatively unstable across time (i.e. assignment to the mixed state at a given time point equally predicted assignment to any of the three identified mood states at the subsequent time point) and was uniquely associated with rapid cycling, substance use, and psychosis. Descriptively, individuals assigned to the latent mixed state at baseline were relatively more likely to be diagnosed with a current mixed, manic, or hypomanic episode, and reported spending more time experiencing irritable and elevated mood in the preceding year relative to individuals assigned to euthymic or depressed states. Although they were also relatively less likely to be diagnosed with bipolar II (v.

Table 1. Participant characteristics in the overall sample and by assigned latent state at baseline

Variable	Overall (<i>n</i> = 3918)	Sample		
		'Euthymic' (<i>n</i> = 1976)	'Depressed' (<i>n</i> = 1453)	'Mixed' (<i>n</i> = 489)
Demographics				
Female, %	57.7	56.5	59.2	58.1
Caucasian, %	90.2	90.2	90.2	90.2
Graduated college, %	43.6	42.6	44.5	44.9
Employed, %	31.6	31.7	30.9	33.3
Married, %	33.6	34.1	33.6	31.6
Age (M, s.d.)	39.7, 12.9	39.9, 13.8	40.0, 11.7	37.9, 12.4
Cigarettes/day (M, s.d.)	6.0, 12.0	5.5, 11.2	7.7, 13.0	7.3, 11.5
Bipolar diagnosis				
Bipolar I, %	64.6	64.5	63.2	69.3
Bipolar II, %	26.9	26.4	28.8	23.3
Bipolar NOS, %	6.9	7.2	6.6	6.1
Current episode status				
Subthreshold, %	55.9	78.6	32.6	33.5
MD episode, %	30.6	15.0	54.4	22.5
Manic episode, %	2.5	1.4	1.6	9.8
Mixed episode, %	7.3	2.3	9.5	20.9
Hypomanic episode, %	3.7	2.6	1.9	13.3
Current mood (M, s.d.)				
YMRS	7.1, 6.7	3.7, 3.6	7.0, 4.1	20.7, 4.7
MADRS	16.3, 10.9	8.5, 6.3	25.3, 7.9	20.8, 10.1
Past year mood (M, s.d.)				
% Days irritable	31.6, 30.3	23.2, 26.8	39.1, 40.0	43.4, 32.0
% Days elevated	19.4, 21.2	16.5, 19.8	19.3, 20.1	31.3, 25.4
% Days depressed	43.0, 29.9	32.1, 27.8	57.1, 27.1	45.5, 28.2
% Days anxious	34.9, 33.7	25.5, 30.2	44.5, 34.3	43.6, 34.6

NOS, not otherwise specified; subthreshold, not diagnosed with an MD, manic, mixed, or hypomanic episode; MD, major depressive; YMRS, Young Mania Rating Scale; and MADRS, Montgomery-Asberg Depression Rating Scale.

Note: Statistical tests are not included in order to maintain a reasonable experimentwise alpha level.

I) disorder, nearly 25% of individuals assigned to the mixed state at baseline were diagnosed with bipolar II disorder nonetheless, supporting the changes in how mixed symptoms are characterized in DSM-V relative to DSM-IV (which precluded the diagnosis of mixed episodes in bipolar II disorder by definition).

These findings are concordant with a recent study that identified 'stable' (i.e. primarily euthymic), 'depressive' [i.e. more prominent depression (i.e. 23% of the time) than the 'stable' class (i.e. 8% of the time)] and 'rapid-cycling' [i.e. somewhat less prominent depression (i.e. 17% *v.* 23% of the time) but more prominent mania (i.e. 13% *v.* 7% of the time) relative to the 'depressive' class] classes of mostly female outpatients with bipolar I disorder (Cochran *et al.*, 2016); Cochran's study used a similar statistical methodology as in the current study, but with a relatively smaller sample (i.e. *n* = 209) combined with a large number of retrospectively obtained data points [i.e. taken from a single administration of the Longitudinal Interval Follow-up Evaluation (LIFE) (Denicoff *et al.*, 2000)]. Their primary finding

was that 'rapid-cycling' participants experienced a greater degree of functional impairment and disability relative to 'stable' and 'depressive' participants. With continued empirical support, this emerging literature could provide clinicians and researchers with important predictive information about their patients and participants.

Based on both the current findings and the extent of literature, individuals with BD who present to the research and/or treatment clinic with elevated levels of both depression and mania are likely experiencing a state that is transient yet clinically dangerous (Pallaskorpi *et al.*, 2017). They may be at high risk for substance use and further mixed and depressive mood episodes, and they are more likely to have a history of psychosis. In short, they are different from individuals presenting with depression or euthymia in important ways that indicate elevated severity and risk for deleterious outcomes. Clearly, in the case of patients presenting to the clinic, these predictions could provide important guidance to clinicians regarding treatment and likely course. Results from the

Table 2. Fit and estimated means for hidden Markov models

No. of states	<i>n</i> Adj BIC	Estimated YMRS/MADRS means (% assigned to state at baseline)					
		State 1	State 2	State 3	State 4	State 5	State 6
2	156 473.3	3.9/8.0 (51.3)	9.5/24.3 (48.7)				
3	153 496.7	3.3/7.6 (45.8)	6.7/23.6 (40.6)	20.0/20.1 (13.6)			
4	152 236.9	2.5/7.4 (40.3)	5.0/23.7 (30.2)	23.7/22.0 (7.6)	12.4/17.7 (22.0)		
5	151 280.3	2.3/6.7 (37.2)	4.5/23.6 (30.0)	23.9/20.8 (7.3)	11.4/12.3 (15.6)	12.9/29.0 (10.0)	
6	150 767.2	2.1/6.6 (35.1)	4.1/23.4 (27.5)	20.2/19.5 (8.6)	10.2/11.9 (16.2)	11.7/28.4 (10.5)	29.1/23.2 (2.0)

n Adj BIC, sample-size-adjusted Bayesian Information Criterion.

Note: Whereas number of states and estimated means were fixed across time points within a given model, the % of the sample assigned to each state was not fixed across time points because individuals were free to transit between latent states across time points; only the percentage of the sample assigned to each state at baseline is provided above for brevity.

Table 3. Proportion of subjects classified in each state at each time point and latent state transition probabilities

(A) Proportion of subjects in each state/time point			
Time (<i>t</i>)	State	<i>n</i>	Proportion
1 (Baseline)	'Mixed'	532	0.14
	'Depressed'	1591	0.41
	'Euthymic'	1795	0.46
2 (Month 3)	'Mixed'	389	0.10
	'Depressed'	1409	0.36
3 (Month 6)	'Euthymic'	2120	0.54
	'Mixed'	329	0.08
4 (Month 9)	'Depressed'	1236	0.32
	'Euthymic'	2353	0.60
5 (Month 12)	'Mixed'	297	0.08
	'Depressed'	1097	0.28
	'Euthymic'	2524	0.64
	'Mixed'	277	0.07
	'Depressed'	992	0.25
	'Euthymic'	2649	0.68
(B) Latent state transition probabilities			
	$P(M'_{t+1})$	$P(D'_{t+1})$	$P(E'_{t+1})$
	0.37	0.37	0.26
Time (<i>t</i>)	0.09	0.72	0.19
	0.03	0.04	0.94

Proportion, proportion of the sample classified in a given latent state at a given point in time; P, probability of being classified in a particular latent state at the subsequent time point; M, Mixed; D, Depressed; E, Euthymic.

Note. Latent state transition probabilities were fixed across contiguous pairs of time points (e.g., transition probabilities for $t = 1$ to $t = 2$ were the same as probabilities for $t = 2$ to $t = 3$ and so on).

current study could also be used to define potential ranges and/or cut-offs on the most commonly utilized clinician-administered bipolar mood assessments (i.e. YMRS and MADRS) that could be used to diagnose euthymic, depressive, and mixed mood states in outpatients with BD. These cut-offs would represent a relatively objective and opinion-agnostic method of mood state ('episode') definition derived from longitudinal mood data on a very large sample of bipolar outpatients, in contrast to the current method that consists of using an assessment with more limited psychometric properties (i.e. the Structured Clinical Interview for DSM-IV Disorders) to diagnose mood states that have been defined by expert opinion (e.g. at least five of nine clinically significant depressive symptoms, necessarily including depressed mood and/or anhedonia, that have been present most of the time for at least 2 weeks – 'Major Depressive Episode' ala DSM-5). Using these empirically defined mood states as building blocks, future studies could examine more frequently collected (i.e. weekly) YMRS and MADRS data to identify characteristic longitudinal patterns of mood dysfunction which could be used to classify patients into meaningful subtypes that are defined by mood state course. We have avoided doing so in the current study because reliable and valid assessments of bipolar mood symptoms were obtained only quarterly in the STEP-BD study, precluding a more temporally granular examination of transitions between mood states. Cochran and colleagues' investigation may provide a first step toward such granular examination of mood state transitions suggesting individuals with bipolar I disorder can be divided into 'stable' (i.e. predominantly euthymic), 'depressive' (i.e. predominantly euthymic but depressed approximately 20% of the time), and 'rapid-cycling' (i.e. predominantly euthymic but with more time spent in mania relative to depressive patients and more transient mood episodes overall) (Cochran et al., 2016). However, given the relatively small sample, over-representation of females (i.e. potential lack of generalizability), and psychometric weaknesses of the retrospectively assessed symptomatology information, further research may be needed to provide more confidence to their findings. Many suitable data sets for the suggested investigation already exist in the form of completed clinical trials for BD, many (and perhaps most) of which used weekly administrations of the MADRS and YMRS to evaluate treatment outcome.

Table 4. Associations between baseline covariates and probabilities of remaining in or transitioning between latent states from one time point to the next

Predictor	Reference	Comparison	Logit	<i>p</i>	OR
Hx of suicide attempt	'Euthymic'	'Mixed'	0.42	0.001	1.53
		'Depressed'	0.52	<0.001	1.68
Hx of psychosis	'Euthymic'	'Mixed'	-0.10	0.37	0.91
	'Depressed'	'Mixed'	0.28	0.001	1.32
Antidepressant-associated switch Hx	'Euthymic'	'Mixed'	0.07	0.49	1.07
	'Depressed'	'Mixed'	-0.04	0.58	0.96
Past year rapid cycling	'Euthymic'	'Mixed'	0.89	<0.001	2.43
		'Depressed'	0.43	<0.001	1.53
	'Depressed'	'Mixed'	0.46	<0.001	1.59
Age of BD onset	'Euthymic'	'Mixed'	-0.19	0.01	0.83
		'Depressed'	-0.11	0.10	0.92
	'Depressed'	'Mixed'	-0.08	0.20	0.92
Substance dependence	'Euthymic'	'Mixed'	0.42	<0.001	1.52
		'Depressed'	0.21	0.003	1.23
	'Depressed'	'Mixed'	0.21	0.008	1.23
Panic disorder	'Euthymic'	'Mixed'	0.58	<0.001	1.79
		'Depressed'	0.62	<0.001	1.86
	'Depressed'	'Mixed'	-0.04	0.61	0.97

Hx, history; BD, bipolar disorder.

Note: Covariate effects reported are above and beyond the effect of latent class membership at a given time point on remaining in or transitioning to latent states at the subsequent time point (i.e. first order autoregressive effects). Associations between covariates and latent states were estimated in a multinomial logistic regression framework: 'reference' = reference group and 'comparison' = comparison group for a given logit. Significant effects (i.e. $p < 0.05$) are bolded.

Although the current study has a number of unique strengths (e.g. large naturalistic sample of individuals with BD, repeated clinician-administered assessment of depressive and manic symptoms using valid and reliable instruments, and sophisticated statistical methodology), it also suffers from notable weaknesses. First, although the current study defined latent mood states and transitions among mood states using longitudinal data, each contiguous pair of assessments was separated by 3 months. Second, a relatively limited subset of clinically relevant variables was examined in relation to the empirically defined mood states in order to reduce experimentwise alpha inflation. Third, to maximize our ability to identify latent mood states with potentially sparse patient assignment at any given time point, we did not evaluate the replicability of the obtained mood states within the STEP-BD data. Nor did we evaluate the construct validity of models containing four to six latent mixtures per time point. However, we provided definitions of these additional mixtures, in terms of YMRS and MADRS scores, in hopes that this information may provide an impetus for further research. Fourth, the current study was limited to outpatient individuals which naturally restricted the incidence of acute manic symptoms in the sample, as can be seen by the low reported base rate of acute manic episodes at baseline (Table 1) (Carragher *et al.*, 2013). As such, the absence of an obtained latent (hypo)manic state in the current study was likely largely due to methodological limitations. In addition to poor representation of individuals experiencing acute

mania in the STEP-BD study, the psychometric measure used to measure (hypo)manic symptoms in the study (i.e. YMRS) has been shown to suffer from notable psychometric limitations, including poor coverage of average or below-average outpatient levels of manic symptoms (Prisciandaro and Tolliver, 2016). In sum, additional supportive studies will be needed before the symptom thresholds derived from the current study can be used with confidence to define mood states in future research or clinical decision making.

These limitations notwithstanding, the results from the current study represent an important first step in defining, and characterizing the longitudinal course of, empirically derived mood states that can be used to form the foundation of objective, empirical attempts to define meaningful subtypes of affective illness defined by clinical course.

Acknowledgements. STEP-BD Data Use Certification for Public Release Dataset version 4.1 provided 27 April 2011 (Tolliver, Recipient PI). Data used in the preparation of this article were obtained from the limited access datasets distributed from the NIH-supported 'Systematic Treatment Enhancement Program for Bipolar Disorder' (STEP BD). This is a multisite, clinical trial studying the current treatments for BD, including medications and psychosocial therapies. This study was supported by NIMH contract no. N01MH80001 to Massachusetts General Hospital and the University of Pittsburgh. The ClinicalTrials.gov identifier is NCT00012558. This manuscript reflects the views of the authors and may not reflect the opinions or views of the STEP BD Study Investigators or the NIH.

Financial support. Dr Prisciandaro was funded by K23 AA020842.

Conflict of interest. None.

References

- Abreu T and Braganca M (2015) The bipolarity of light and dark: a review on Bipolar Disorder and circadian cycles. *Journal of Affective Disorders* **185**, 219–229.
- Akiskal HS (1983) Dysthymic and cyclothymic disorders: a paradigm for highrisk research. In Davis JM and Maas JW (eds), *The Affective Disorders*. Washington DC: American Psychiatric Press, pp. 211–232.
- American Psychiatric Association (1980) *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Association.
- American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Association.
- Angst J (1966) *Ätiologie und Nosologie endogener depressiver Psychosen. Eine genetische, soziologische und klinische Studie*. Berlin, Heidelberg, New York: Springer.
- Angst J, Cui L, Swendsen J, Rothen S, Cravchik A, Kessler RC and Merikangas KR (2010) Major depressive disorder with subthreshold bipolarity in the National Comorbidity Survey Replication. *American Journal of Psychiatry* **167**, 1194–1201.
- Angst J, Azorin JM, Bowden CL, Perugi G, Vieta E, Gamma A and Young AH (2011) Prevalence and characteristics of undiagnosed bipolar disorders in patients with a major depressive episode: the BRIDGE study. *Archives of General Psychiatry* **68**, 791–798.
- Carragher N, Weinstock LM and Strong D (2013) Psychometric evaluation of the DSM-IV criterion B mania symptoms in an Australian national sample. *Psychological Medicine* **43**, 433–443.
- Cochran AL, Mcinnis MG and Forger DB (2016) Data-driven classification of bipolar I disorder from longitudinal course of mood. *Translational Psychiatry* **6**, e912.
- Denicoff KD, Leverich GS, Nolen WA, Rush AJ, Mcelroy SL, Keck PE, Suppes T, Altshuler LL, Kupka R, Frye MA, Hatfield J, Brotman MA and Post RM (2000) Validation of the prospective NIMH-Life-Chart Method (NIMH-LCM-p) for longitudinal assessment of bipolar illness. *Psychological Medicine* **30**, 1391–1397.
- Falret JP (1851) De la folie circulaire ou forme de maladie mentale caractérisée par l'alternative régulière de la manie et de la mélancolie. *Bulletin de l'Académie Nationale de Médecine* **19**, 382–400.
- Fiedorowicz JG, Endicott J, Leon AC, Solomon DA, Keller MB and Coryell WH (2011) Subthreshold hypomanic symptoms in progression from unipolar major depression to bipolar disorder. *American Journal of Psychiatry* **168**, 40–48.
- Goldberg JF, Perlis RH, Bowden CL, Thase ME, Miklowitz DJ, Marangell LB, Calabrese JR, Nierenberg AA and Sachs GS (2009) Manic symptoms during depressive episodes in 1380 patients with bipolar disorder: findings from the STEP-BD. *American Journal of Psychiatry* **166**, 173–181.
- Goodwin FK and Jamison KR (2007) *Manic-Depressive Illness: Bipolar Disorders and Recurrent Depression*. Oxford, UK: Oxford University Press.
- Kraepelin E (1921) *Manic Depressive Insanity and Paranoia*. Edinburgh, ES: Livingstone.
- Langeheine R and van de Pol F (2002) Latent Markov chains. In Hagenaars JA and McCutcheon AL (eds), *Applied Latent Class Analysis*. Cambridge UK: Cambridge University Press, pp. 304–341.
- MacDonald IL and Zucchini W (1997) *Hidden Markov and Other Models for Discrete-valued Time Series*. Boca Raton, Florida: CRC Press.
- Marneros A and Angst J (2001) *Bipolar Disorders: 100 Years after Manic-Depressive Insanity*. Netherlands: Springer.
- Muthen LK and Muthen BO (2011) *Mplus User's Guide*, 6th Edn. Los Angeles, CA: Muthen & Muthen.
- Nitsu T, Fabbri C and Serretti A (2015) Predictors of switch from depression to mania in bipolar disorder. *Journal of Psychiatric Research* **66–67**, 45–53.
- Nylund KL, Asparouhov T and Muthén BO (2007) Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo Simulation Study. *Structural Equation Modeling: A Multidisciplinary Journal* **14**, 535–569.
- Ostacher MJ, Perlis RH, Nierenberg AA, Calabrese J, Stange JP, Salloum I, Weiss RD and Sachs GS (2010) Impact of substance use disorders on recovery from episodes of depression in bipolar disorder patients: prospective data from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *American Journal of Psychiatry* **167**, 289–297.
- Pallaskorpi S, Suominen K, Ketokivi M, Valtonen H, Arvilommi P, Mantere O, Leppamaki S and Isometsa E (2017) Incidence and predictors of suicide attempts in bipolar I and II disorders: a 5-year follow-up study. *Bipolar Disorders* **19**, 13–22.
- Perlis RH, Miyahara S, Marangell LB, Wisniewski SR, Ostacher M, Delbello MP, Bowden CL, Sachs GS and Nierenberg AA (2004) Long-term implications of early onset in bipolar disorder: data from the first 1000 participants in the systematic treatment enhancement program for bipolar disorder (STEP-BD). *Biological Psychiatry* **55**, 875–881.
- Perlis RH, Ostacher MJ, Goldberg JF, Miklowitz DJ, Friedman E, Calabrese J, Thase ME and Sachs GS (2010) Transition to mania during treatment of bipolar depression. *Neuropsychopharmacology* **35**, 2545–2552.
- Perugi G, Angst J, Azorin JM, Bowden CL, Mosolov S, Reis J, Vieta E and Young AH (2015) Mixed features in patients with a major depressive episode: the BRIDGE-II-MIX study. *Journal of Clinical Psychiatry* **76**, e351–e358.
- Prisciandaro JJ and Roberts JE (2005) A taxometric investigation of unipolar depression in the national comorbidity survey. *Journal of Abnormal Psychology* **114**, 718–728.
- Prisciandaro JJ and Roberts JE (2011) Evidence for the continuous latent structure of mania in the Epidemiologic Catchment Area from multiple latent structure and construct validation methodologies. *Psychological Medicine* **41**, 575–588.
- Prisciandaro JJ and Tolliver BK (2016) An item response theory evaluation of the young mania rating scale and the Montgomery-Asberg Depression Rating scale in the systematic treatment enhancement program for bipolar disorder (STEP-BD). *Journal of Affective Disorders* **205**, 73–80.
- Prisciandaro JJ, Desantis SM, Chiuzan C, Brown DG, Brady KT and Tolliver BK (2012) Impact of depressive symptoms on future alcohol use in patients with co-occurring bipolar disorder and alcohol dependence: a prospective analysis in an 8-week randomized controlled trial of acamprosate. *Alcoholism, Clinical and Experimental Research* **36**, 490–496.
- Sachs GS, Thase ME, Otto MW, Bauer M, Miklowitz D, Wisniewski SR, Lavori P, Lebowitz B, Rudorfer M, Frank E, Nierenberg AA, Fava M, Bowden C, Ketter T, Marangell L, Calabrese J, Kupfer D and Rosenbaum JF (2003) Rationale, design, and methods of the systematic treatment enhancement program for bipolar disorder (STEP-BD). *Biological Psychiatry* **53**, 1028–1042.
- Schneck CD, Miklowitz DJ, Miyahara S, Araga M, Wisniewski S, Gyulai L, Allen MH, Thase ME and Sachs GS (2008) The prospective course of rapid-cycling bipolar disorder: findings from the STEP-BD. *American Journal of Psychiatry* **165**, 370–377; quiz 410.
- Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R and Dunbar GC (1998) The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry* **59**(Suppl. 20), 22–33; quiz 34–57.
- Shirley KE, Small DS, Lynch KG, Maisto SA and Oslin DW (2010) Hidden Markov models for alcoholism treatment trial data. *The Annals of Applied Statistics* **4**, 366–395.
- Wall MM and Li R (2009) Multiple indicator hidden Markov model with an application to medical utilization data. *Statistics in Medicine* **28**, 293–310.
- Young RC, Biggs JT, Ziegler VE and Meyer DA (1978) A rating scale for mania: reliability, validity and sensitivity. *British Journal of Psychiatry* **133**, 429–435.
- Zimmermann P, Bruckl T, Nocon A, Pfister H, Lieb R, Wittchen HU, Holsboer F and Angst J (2009) Heterogeneity of DSM-IV major depressive disorder as a consequence of subthreshold bipolarity. *Archives of General Psychiatry* **66**, 1341–1352.