Psychological Medicine

cambridge.org/psm

Limited evidence of autocorrelation signaling upcoming affective episodes: a 12-month e-diary study in patients with bipolar disorder

V. M. Ludwig[1], I. Reinhard[2], E. Mühlbauer[1], H. Hill[1,3], W. E. Severus[1,4], M. Bauer[1], P. Ritter[1] and U. W. Ebner-Priemer[3,5]

[1]Department of Psychiatry and Psychotherapy, University Hospital Carl Gustav Carus, Technical University Dresden, Dresden, Germany; [2]Department of Biostatistics, Central Institute of Mental Health, University of Heidelberg, Medical Faculty Mannheim, Mannheim, Germany; [3]Mental mHealth Lab, Institute of Sport and Sport Sciences, Karlsruhe Institute of Technology, Karlsruhe, Germany; [4]Asklepios Klinik Nord-Ochsenzoll, Hamburg, Germany and [5]Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, University of Heidelberg, Medical Faculty Mannheim, Mannheim, Germany

Original Article

Cite this article: Ludwig VM, Reinhard I, Mühlbauer E, Hill H, Severus WE, Bauer M, Ritter P, Ebner-Priemer UW (2024). Limited evidence of autocorrelation signaling upcoming affective episodes: a 12-month e-diary study in patients with bipolar disorder. *Psychological Medicine* **54**, 1844–1852. https://doi.org/10.1017/S0033291723003811

Received: 26 April 2023
Revised: 3 December 2023
Accepted: 14 December 2023
First published online: 29 January 2024

Keywords:
autocorrelation; bipolar disorder; depression; e-diaries; mania

Corresponding author:
V. M. Ludwig;
Email: Vera.Ludwig@ukdd.de

Abstract

Background. Increased autocorrelation (AR) of system-specific measures has been suggested as a predictor for critical transitions in complex systems. Increased AR of mood scores has been reported to anticipate depressive episodes in major depressive disorder, while other studies found AR increases to be associated with depressive episodes themselves. Data on AR in patients with bipolar disorders (BD) is limited and inconclusive.

Methods. Patients with BD reported their current mood via daily e-diaries for 12 months. Current affective status (euthymic, prodromal, depressed, (hypo)manic) was assessed in 26 bi-weekly expert interviews. Exploratory analyses tested whether self-reported current mood and AR of the same item could differentiate between prodromal phases or affective episodes and euthymia.

Results. A total of 29 depressive and 20 (hypo)manic episodes were observed in 29 participants with BD. Self-reported current mood was significantly decreased during the two weeks prior to a depressive episode (early prodromal, late prodromal), but not changed prior to manic episodes. The AR was neither a significant predictor for the early or late prodromal phase of depression nor for the early prodromal phase of (hypo)mania. Decreased AR was found in the late prodromal phase of (hypo)mania. Increased AR was mainly found during depressive episodes.

Conclusions. AR changes might not be better at predicting depressive episodes than simple self-report measures on current mood in patients with BD. Increased AR was mostly found during depressive episodes. Potentially, changes in AR might anticipate (hypo)manic episodes.

Introduction

Bipolar disorders (BD) are mental disorders characterized by episodes of (hypo)mania and depression with an estimated global lifetime prevalence of 0.4–1.1% (Mcintyre et al., 2020). Due to their onset in early adulthood and chronic course, they constitute a substantial burden of disease (GBD 2019 Mental Disorders Collaborators, 2022). One main goal of BD treatment is to prevent the occurrence of new affective episodes and extend the time spent in remission by recognizing early signs of recurrence (Morriss et al., 2007; Perry, Tarrier, Morriss, McCarthy, & Limb, 1999).

A suitable methodological approach to studying early signs of recurring affective episodes is the collection of real-time self-report data via e-diaries (Bauer et al., 2004; Bos et al., 2022). Key advantages of e-diaries (also known as *experience sampling method* [ESM; Csikszentmihalyi & Larson, 1987], *ecological momentary assessment* [EMA; Stone & Shiffman, 1994], or ambulatory assessment [Trull & Ebner-Priemer, 2009]) are the near real-time assessment of dynamical constructs, thus minimizing retrospective biases, and enhancing ecological validity (Trull & Ebner-Priemer, 2009).

In recent years, concepts of dynamical systems theory, originally rooted in mathematics and physics, have been applied to psychopathology providing a theoretical framework to conceptualize sudden shifts or transitions to alternate states within complex systems (Van De Leemput et al., 2014). Critical transitions, like the switch to an alternate mood state, e.g. a depressive episode, are hypothesized to be preceded by so-called *early warning signals* (EWS). These EWS are based on the idea of '*critical slowing down*' – the observation that a

system is slower to recover from small external perturbations when it is close to a critical transition than when a tipping point is further away (Scheffer et al., 2012).

These different speeds of recovery after perturbations can be operationalized by analyzing autocorrelation (AR; the correlation of a variable at time t with itself at $t-1$) of system-specific measures (e.g. current mood). Increased AR then indicates a higher concordance of the current state with the state before and thus a less dynamic, slower system, as seen close before critical transitions. Thus, increased AR acts as early warning signal prior to shifts between alternate states like euthymia to depression (Dablander, Pichler, Cika, & Bacilieri, 2022). Importantly, once a new stable state is reached after the transition, the system is hypothesized to be at full speed again, so recovery from external or internal perturbances is quicker than shortly before the transition, evidenced by the AR returning to 'normal'.

Surprisingly, the mathematical index of AR is also used to describe another psychopathological phenomenon: *emotional inertia*. Emotional inertia can be defined as resistance to emotional change (Kuppens, Allen, & Sheeber, 2010). Increased AR equals high emotional inertia and represents lowered emotional responsiveness to different stimuli, a state that reportedly correlates with depressive episodes (Houben, Van Den Noortgate, & Kuppens, 2015).

In sum, AR has been hypothesized to be: (a) increased *before* critical transitions to new states, such as depressive episodes, and back to normal once a new state is reached within the dynamical systems theory, as well as (b) increased *during* depressive episodes, in the context of the emotional inertia concept.

Early warning signals in unipolar depression

A seminal study found increased AR at baseline to be related to increased depressive symptoms at follow-up (Van De Leemput et al., 2014), thereby introducing the concept of critical slowing down to psychopathology. However, no temporal patterns for AR change before or during affective episodes in individual illness trajectories were assessed. These shortcomings were addressed by subsequent studies that featured longer assessment periods and looked at time series before actual transitions on an individual level. A single-case study (Wichers & Groot, 2016) and a pilot study (Wichers, Smit, & Snippe, 2020) each detected one transition to a depressive episode in individuals with major depressive disorder (MDD). They reported increased AR of EMA items one to two months prior to the transition with a steep decline in AR after the transition. Another study followed 41 depressed patients with known MDD that were starting treatment and observed nine transitions to euthymia (Helmich et al., 2023). Increased AR was reportedly more frequent in individuals with a critical transition to remission than in those without. However, these results varied considerably across individuals and EMA items. In sum, while across these three studies a total of 11 clinically relevant transitions in MDD patients were observed, the results varied considerably depending on the individuals and EMA items and thus require cautious interpretation.

Early warning signs in bipolar disorder

Only few studies empirically investigated EWS using EMA in patients with BD. First, a study comparing daily affect ratings in 32 BD I patients in remission and 36 healthy controls reported that increased AR in two of the ten EMA affect items predicted worsening of depressive symptoms in the BD group (Curtiss, Fulford, Hofmann, & Gershon, 2019). A severe limitation is that no affective episodes occurred during the study period and results only refer to subthreshold depressive symptom changes. Kunkels et al., studied 15 BD I patients for six months and observed eight mood transitions ($n=5$ depressive, $n=3$ manic) during the study period. Results showed that changes in AR were able to significantly predict mood episodes; however, the effect was not in the hypothesized direction. AR shifts in *either* direction (increased or decreased) preceded mood episodes of both polarities (Kunkels et al., 2021). Bos et al. reported results on 20 BD I/II patients that completed four months of EMA (Bos et al., 2022). Fifteen transitions to affective episodes were observed within the study period. All mood episodes were preceded by at least one EWS (rise in AR or standard deviation) in at least one EMA item. Notably, of the 17 analyzed EMA items, only two proved relatively robust, the rest were characterized by large inter- and intraindividual variations. The clinical utility of EWS was determined by calculating positive predictive values (the chance of an upcoming episode after an EWS), which were slightly increased, as well as negative predictive values (the chance to stay euthymic in the absence of EWS), which were not above chance.

Emotional inertia and affective symptoms

The reported associations between high levels of inertia and depressive symptoms (Koval, Sütterlin, & Kuppens, 2016; Kuppens et al., 2010; Nelson, Klumparendt, Doebler, & Ehring, 2020; Panaite, Rottenberg, & Bylsma, 2020), were confirmed in a meta-analysis (Houben et al., 2015). To the best of our knowledge, so far only one research group has investigated AR as a proxy for emotional inertia within bipolar spectrum symptomatology and did not find consistent or promising results (Sperry & Kwapil, 2019, 2022; Sperry, Walsh, & Kwapil, 2020). While these studies feature very short assessment periods with only 7–14 days of EMA collection, more importantly they were conducted on student samples, not BD patients. Thus, no actual affective episodes were observed within these studies and their results cannot assume any statements on the patterns of emotional inertia during affective episodes in BD.

Aims of this study

To investigate the temporal patterns of AR changes before and after the onset of affective episodes (depressive or [hypo-]manic), we studied a clinical sample of BD patients. In an exploratory approach, we investigated (a) whether the AR is heightened in the weeks before affective episodes, establishing its utility as EWS in BD, (b) whether the AR is heightened during affective episodes, making it an attribute of mood episodes rather than a prodromal sign, and (c) if the AR of self-reported mood contributes beyond the level of self-reported mood to differentiate whether there is added value in the AR as a predictor.

Methods

Study protocol and assessments

Data were collected as part of the BipoSense study (Ebner-Priemer et al., 2020). Participants were monitored over a 12-month period, completing bi-weekly interviews, daily

https://doi.org/10.1017/S0033291723003811 Published online by Cambridge University Press

end-of-day e-diary ratings, and continuous digital phenotyping (not part of the current manuscript). A flow chart of the patient flow through the study can be found in the online Supplementary materials (Fig. S1).

Participants

Participants were recruited through a specialized outpatient clinic at the University Hospital Dresden and through articles in print and online media. This was a voluntary study offer. The inclusion criteria were: (a) diagnosed with BD Type I or II, in remission at the time of enrollment (Young Mania Rating Scale [YMRS] score $\leqslant 12$ and Montgomery-Asberg Depression Rating Scale [MADRS] score $\leqslant 12$); (b) to have experienced at least three affective episodes within the past five years, at least one of them hypomanic or manic; (c) $\geqslant$ 18 years; (d) willing to use a smartphone. Exclusion criteria were (a) current substance abuse (except for tobacco and caffeine); (b) comorbid diagnosis of borderline personality disorder or antisocial personality disorder; (c) dementia or other organic brain diseases; (d) instable or insufficiently treated other physical illnesses; (e) clinically relevant cardiovascular, neoplastic, kidney or liver disease. The study was carried out in accordance with the Declaration of Helsinki and approved by the local ethics committee at the medical faculty of the Technical University of Dresden (EK-Nr.: 26012014). All participants gave written informed consent before being included in the study.

Procedures

After controlling for in- and exclusion criteria, participants were informed about the study procedures. For data collection, the movisensXS app (movisens GmbH, Karlsruhe, Germany) was used. Participants had the choice to have the app installed on their private phones or to receive a study smartphone that was to replace their private phone for the duration of the study. Furthermore, psychopathological assessments (see measures) were conducted at the first visit. Henceforth, the psychopathological status was assessed every 14 days by a trained clinical psychologist, alternating between in-person and telephone appointments. In addition, every evening the movisensXS app prompted participants to fill out their end-of-day e-diary, repeating the prompt hourly up to three times when the participants did not react. The e-diary had to be completed the same day. Participants received a compensation of 35 Euros/month. After the completion of the 12-month study period, the app was deinstalled, patients were debriefed and, if desired, received feedback on their personal data.

Measures

Diagnostic interviews

A clinically trained psychologist interviewed participants bi-weekly for 12 months, resulting in a total of 26 assessments per participant. During the interviews, current affective episodes were determined for the previous two weeks with the SCID-I section A for affective episodes according to DSM-5 (First, Williams, Karg, & Spitzer, 2015). Detailed information on (hypo)manic and depressive symptoms was also rated at each assessment, but are not part of the current manuscript German versions of the YMRS (Young, Biggs, Ziegler, & Meyer, 1978), the Bech-Rafaelsen Mania Rating Scale (BRMRS; Bech, Bolwig, Kramp, and Rafaelsen, 1979), and the MADRS (Montgomery & Asberg, 1979).

Self-reported mood

Participants answered daily end-of-day diary questions on their smartphone, which were adapted from ChronoRecord, a computer-based, thoroughly validated system tracking mood, medication, and sleep in patients with BD (Bauer et al., 2004, 2008). The relevant item for this study is a single question, asking participants about their current mood on a scale ranging from 'depressed' to 'even-tempered' to 'elevated'. Importantly, this represents a bipolar item in the sense that two different extremes of mood (manic and depressed) are assessed at once. A screenshot of the item in the study app can be found in the online Supplementary materials Fig. S2 Participants answered this item on a continuous visual analog scale ranging from 0 to 100. The 365 repeated measures of this item were used as the basis for analyzing the AR. No minimum number of e-diaries or interviews was required; all participants were included regardless of missing values.

Data preparation

Autocorrelation

To make sure patterns in self-reported mood were not due to changes in absolute levels of affect, we detrended the data by using the kernel smoothing function *locpoly* contained in the R package *KernSmooth* 2.23-20 (Wand & Jones, 1995) in R-4.1.2. Then we subtracted the resulting estimated values from the observed values to obtain the detrended values. These deviations were used in further analyses. To calculate individual daily values of AR, we used a moving window technique, applying a window of 14 days. For every day of the study period, the AR of the detrended self-reported mood items of the last 14 days was calculated for each participant. We did not impute missing values. If mood was missing at day t or day $t-1$ then the autocorrelation for these days was set to missing. Moreover, if within the 14-day window there were less than seven pairs of values available for the calculation of the AR, we set the value missing likewise. As a sensitivity analysis, we performed all calculations based on a 30-day moving window with a minimum of 20 available pairs of data for the AR. The individual daily AR values were used to assess the concurrent associations with disorder status on that specific day.

Expert-rated disorder status

To further differentiate temporal patterns before and during phase shifts, we introduced the variable 'disorder status'. This variable discriminates between the weeks prior to affective episodes and differentiates the subsequent depressive or (hypo) manic weeks. The 7 days 2 weeks before episode onset (days 14 to 8 prior to the start of the affective episode) are labelled 'early prodromal' and the week directly before episode onset (days 7 to 1) 'late prodromal'. Furthermore, we differentiated the first week of the affective episode from the second and then clustered all ongoing depressive or (hypo)manic weeks $\geqslant$ 3 weeks together. Lastly, the rest of the study days were labelled as euthymic. This means euthymic days during the two weeks prior to affective episodes were not labelled euthymic, but early and late prodromal weeks. In four cases, there were between 2 and 7 missing days right before the coded onset of an affective episode, caused by slight delays in the bi-weekly interviews. In those cases, we treated

https://doi.org/10.1017/S0033291723003811 Published online by Cambridge University Press

the first coded day of the episode as the first actual day and counted the missing days as prodromal days.

Analyses and statistical models

To account for the nesting structure of each data point within a participant, we estimated multilevel (mixed) models with fixed and random effects. Generalized linear mixed models, in particular multilevel logit models, were fitted for disorder status as binary dependent variable (differential status before/during affective episode = yes *v.* euthymic status = no) with AR and/or self-reported mood as predictors (fixed main effects). Our generalized linear mixed models (here: mixed logit models) separate for the disorder status category k (e.g. early prodromal days before a depressive episode) *v.* euthymic days and look like this:

$$\text{logit}(p_k)_{ij} = \log\left(\frac{p_k}{1-p_k}\right) = \beta_{00} + \beta_{0k} \times X\,(\text{Autocorrelation})_{ij} + u_{0j}$$

where $p_k = P(Y = k)$ denotes the probability to be in phase k *v.* euthymic for a person j at time i, u_{0j} denotes the random intercept, and the beta coefficients represent the intercept and the effect of AR. In simple terms: the AR of self-reported mood was used as a predictor for a binary outcome of disorder status (comparing euthymic days to early-/late-prodromal days, days of depressive or (hypo)manic episodes). Consequently, these models were run to test if AR as the predictor was able to differentiate euthymic *v.* prodromal or episodic days respectively. These models also work with the odds ratio of being in a specific phase *v.* being euthymic.

Moreover, we used general linear mixed models (multilevel linear models) to evaluate the differences between the categories of disorder status (as an independent factor or fixed effect respectively) and the *dependent variable* AR. Our linear mixed model is as follows:

$$Y(\text{Autocorrelation})_{ij} = \beta_{00} + \sum_{k1}^{\beta} \times I(\text{phase category} = k)_{ij} + u_{0j} + r_{ij}$$

where Y_{ij} represents the outcome 'autocorrelation' of person j at time i, k denotes the category of disorder status of a person j at time i. Beta coefficients represent the intercept and the effects of the disorder status category, r_{ij} represents the residuals at level 1, and we included a random intercept u_{0j}. Simply put, this additional step reversed the models using disorder status (early-/late-prodromal days, weeks of depressive or (hypo)manic episodes) as the independent variable and AR as an outcome. These models were run to explore the size of AR differences between different disorder status categories. The data were analyzed and visualized with SAS (Version 9.4, SAS Institute Inc. Cary, NC, USA), SPSS (IBM SPSS for Windows 27.0; SPSS Inc., Chicago, IL, USA), and R (Version 4.1.2; R Core Team, 2021).

Results

Sample characteristics and compliance

The final sample consisted of 29 participants ($n = 16$ female, $n = 13$ male) with an average age of 43.97 years (S.D. = 11.90, range 25–70 years). $N = 17$ participants were diagnosed with a BD type I and $n = 12$ participants fulfilled the criteria for BD type II. Over the course of their lives, participants reported an average of 7.07 depressive episodes (S.D. = 5.61, range 2–30), 2.97 hypomanic (S.D. = 3.78, range 0–15), and 2.76 manic episodes (S.D. 3.48, range 0–10). Furthermore, they had experienced an average of 3.59 hospital admissions (S.D. = 3.70, range 0–15) due to their BD. Overall, the data consists of 10 587 study days ($M = 365.07$ per patient; range 308–398 days) and participants completed 726 (97%) of the bi-weekly interviews. $N = 9433$ (89%) of the daily e-diaries were completed.

Affective episodes

We observed 29 depressive and 20 (hypo)manic episodes during the study period[†1]. We observed 182 early prodromal and 186 late prodromal days before depressive episodes and 207 days of 1st week, 193 days of the 2nd week, and 359 days of ongoing (⩾3 weeks) depressive episodes in 16 patients. For (hypo)manic episodes, 11 patients contributed 126 early prodromal and 134 late prodromal days as well as 140 days of 1st week, 139 days of 2nd week and 108 days of ongoing (⩾3 weeks) (hypo)manic episodes. A total of 7928 euthymic days and 885 missing days resulted in the total of 10 587 observed patient days. For statistical reasons, we discarded the category of ongoing (⩾3 weeks) (hypo) manic episodes for the analyses, as there were too few observed days.

Using autocorrelation as a predictor for expert-rated disorder status

Multilevel logit models revealed no significant effect of AR as a predictor for prodromal depressive phases (early-prodromal days *v.* euthymic days: $p = 0.107$; late-prodromal days *v.* euthymic days: $p = 0.364$). AR was not able to differentiate prodromal depressive from euthymic days. In contrast, there were significant effects of the predictor AR on the 1st week, the 2nd week, and ongoing depressive weeks ($p = 0.003$; $p = 0.026$; $p < 0.0001$) compared to euthymic days.

For (hypo)manic episodes, there was no significant effect of the predictor AR on early-prodromal days *v.* euthymic days ($p = 0.355$), but a significant effect on late-prodromal days *v.* euthymic days ($p = 0.001$). Furthermore, the predictor AR did not reveal significant effects for the 1st – or 2nd week of (hypo) mania *v.* euthymic days ($p = 0.068$, $p = 0.697$). Put simply, the AR was only able to differentiate between late prodromal (hypo)manic days and euthymic days, but not between euthymic days and early prodromal or (hypo)manic days.

Exploring the size of AR differences between different disorder statuses, we ran reversed models with disorder status (early-/late-prodromal, weeks of depressive or (hypo)manic episodes) as factor and AR as an outcome. Those linear mixed models revealed the estimates depicted in Table 1. Descriptively the AR estimate starts out at 0.11 during euthymic days. Approaching a depressive episode, an increase is visible during early-prodromal (0.16) and late-prodromal days (0.14), which is further pronounced during the depressive phase (1st-week = 0.20, 2nd week = 0.17, and ongoing depressive weeks = 0.26). Prior to (hypo)manic episodes, there is a decrease of the AR estimates in early-prodromal (0.09) and late-prodromal days (0.02), which recedes during 1st- and

[†]The notes appear after the main text.

https://doi.org/10.1017/S0033291723003811 Published online by Cambridge University Press

Table 1. Linear mixed models with differences in autocorrelation between levels of expert-rated disorder status

Disorder status		Estimates of mean AR	Standard error (S.E.)	Post-hoc tests v. euthymic, p value
	Euthymic days	0.11	0.02	
Depressive episodes	Early prodromal	0.16	0.03	**0.049**
	Late prodromal	0.14	0.03	0.283
	1st week depressive episode	0.20	0.03	**0.001**
	2nd week depressive episode	0.17	0.03	**0.032**
	Ongoing depressive weeks	0.26	0.03	**<0.0001**
(Hypo)manic episodes	Early prodromal	0.09	0.04	0.441
	Late prodromal	0.02	0.04	**0.002**
	1st week (hypo)manic episode	0.06	0.04	0.133
	2nd week (hypo)manic episode	0.13	0.04	0.423

Results for linear mixed models with factor disorder status and outcome autocorrelation (AR). Early prodromal = days 14–8 before episode onset; late prodromal = days 7–1 before episode onset. Post-hoc tests were conducted between each disorder state and euthymic days without α-error correction. Significant p values are presented in bold.

2nd week (hypo)mania (0.06 and 0.13). Outcomes of the explorative post-hoc tests, without any α-error corrections, are shown in Table 1.

Those linear mixed models with AR as outcome mirror, of course, the results from the generalized multilevel logit models, as they were just reversed order analyses. However, they help to better understand that heightened AR is especially evident during the depressive weeks and less during prodromal days. Contrarily, decreased – and not heightened – AR is visible before the onset of (hypo)manic episodes and during the first week of (hypo)mania, but is back to normal already during the second (hypo)manic week.

Self-reported current mood as a predictor for expert-rated disorder status

For depressive episodes, multilevel logit models revealed no significant effect of self-reported mood as a predictor on expert-rated disorder status for early prodromal days ($p = 0.117$), but showed significant effects on all other disorder stages (late-prodromal $p = 0.001$, 1st $p < 0.0001$, 2nd $p < 0.0001$, and ongoing depressive weeks $p < 0.0001$). Similarly, self-reported mood as a predictor did not reveal significant effects on early-prodromal or late prodromal weeks before (hypo)manic episodes ($p = 0.117$, $p = 0.507$), but on the 1st and 2nd week of (hypo)mania ($p < 0.0001$, $p < 0.0001$).

To better understand the differences in self-reported mood between the stages of disorder status, we ran reversed models with disorder status (early-/late-prodromal, weeks of depressive or (hypo)manic episodes) as a factor and self-reported mood as outcome. Those linear mixed models revealed the estimates depicted in Table 2. Descriptively, self-reported mood (with a possible range from 0 to 100) starts out at 49.37 during euthymic days, then shows a steady pronounced decline towards depressive episodes. From early-prodromal (47.30) to late-prodromal (45.76), to first (42.12), and second week of a depressive episode (38.16) with the lowest estimate for ongoing depressive weeks (37.21). Approaching (hypo)manic episodes, a smaller change can be observed. Early- and late prodromal weeks are minimally increased compared to euthymic days (50.77, 49.94), with slightly higher values during 1st and 2nd week of (hypo)mania (53.03, 55.30). Results of the explorative post-hoc tests, comparing each disorder status with euthymic days, shown in Table 2, mirror the results of the multilevel logit models.

Combined analyses of autocorrelation and self-reported current mood as predictors for expert-rated disorder status

When comparing the estimates from Table 1 (AR) and Table 2 (self-reported current mood), one might suspect self-reported mood to be a stronger predictor for disorder status than AR of self-reported mood. To gain deeper insight, we ran further multilevel logit models with AR and self-reported mood as concomitant predictors, with results shown in Table 3.

The results of the combined model (Table 3) were quite similar to the two single predictor models (AR; self-reported mood). This denotes that both parameters, AR and mean self-reported mood, contributed independently in predicting the expert-rated disorder status and therefore represent different entities.

Discussion

Using AR of self-reported current mood as EWS for affective episodes was of limited use in our data set, as the AR was neither a significant predictor for the early or late prodromal phase of depression, nor for the early prodromal phase of (hypo)mania. Altered AR was mainly found during depressive episodes (1st week, 2nd week, ongoing depressive weeks) and shortly before (hypo)manic episodes (late prodromal). However, self-reported current mood had already significantly decreased two weeks before a depressive episode (early prodromal, late prodromal), but not significantly changed before (hypo)manic episodes.

At first glance, our results seem to contradict previous studies that describe AR increases as EWS before depressive episodes in unipolar depression (Wichers et al., 2020; Wichers & Groot, 2016). Considering depressive episodes, our results correspond to the emotional inertia theory: during prodromal days, AR does not differ significantly from euthymic days; however, during depressive episodes AR is significantly increased, as suggested for the concept of emotional inertia (Houben et al., 2015; Kuppens et al., 2010). However, regarding data on patients with BD, our results mirror and expand some aspects of previous research

https://doi.org/10.1017/S0033291723003811 Published online by Cambridge University Press

Table 2. Linear mixed models with expert-rated disorder status predicting mean current self-reported mood

Disorder status		Estimates of mean self-reported mood	Standard Error	Post-hoc tests v. euthymic, *p* value
	Euthymic days	49.37	1.01	
Depressive episodes	Early prodromal	47.30	1.28	**0.011**
	Late prodromal	45.76	1.28	**<0.0001**
	1st week depressive episode	42.12	1.26	**<0.0001**
	2nd week depressive episode	38.16	1.28	**<0.0001**
	Ongoing depressive weeks	37.21	1.22	**<0.0001**
(Hypo)manic episodes	Early prodromal	50.77	1.39	0.149
	Late prodromal	49.94	1.35	0.537
	1st week (hypo)manic episode	53.03	1.35	**<0.0001**
	2nd week (hypo)manic episode	55.30	1.35	**<0.0001**

Results for the linear mixed models with disorder status as factor and self-reported mood (assessed via e-diary) as outcome. Early prodromal = days 14–8 before episode onset; late prodromal = days 7–1 before episode onset. Post-hoc tests were conducted between each disorder status and euthymic days without α-error correction. Significant *p* values are presented in bold.

from the dynamical systems perspective. Kunkels et al. (2021) reported that AR changes in *any* direction were detected before transitions to affective episodes in patients with BD (e.g. two of the three observed manic episodes showed significantly decreased AR, and one significantly increased AR prior to the episode). This is in line with our finding of significantly decreased AR the week before a (hypo)manic episode onset. Furthermore, Bos et al. found that the positive predictive value of EWS improved more before manic than before depressive episodes, suggesting EWS, such as AR, might be more useful for the prediction of manic than for the prediction of depressive episodes (Bos et al., 2022). Similarly, we found AR to only be able to discriminate between the week prior to (hypo)manic episodes and euthymic days, but not the weeks prior to depressive episodes and euthymic days. Thus, the utility of AR as EWS before mood shifts might be more promising for (hypo)manic than depressive episodes in BD.

Table 3. Predicting expert-rated disorder status with autocorrelation and self-reported current mood concomitantly

Disorder status		Predicting disorder status with AR and self-reported mood as predictors, *p* values	
		AR	self-reported mood
Depressive episodes	Early prodromal	0.111	0.161
	Late prodromal	0.428	**<0.0001**
	1st week depressive episode	**0.010**	**<0.0001**
	2nd week depressive episode	0.066	**<0.0001**
	Ongoing depressive weeks	**<0.0001**	**<0.0001**
(Hypo)manic episodes	Early prodromal	0.367	0.544
	Late prodromal	**0.001**	0.809
	1st week (hypo) manic episode	0.076	**0.001**
	2nd week (hypo) manic episode	0.715	**<0.0001**

Results of multilevel logit models, predicting disorder status with autocorrelation (AR) and current self-reported mood (assessed via e-diaries) as predictors. Early prodromal = days 14–8 before episode onset; late prodromal = days 7–1 before episode onset. All disorder stages are compared to euthymic days. No α-error correction was applied. Significant *p* values are presented in bold.

Results for self-reported current mood via e-diaries showed that mood changes towards the depressed pole were well captured by this item. The self-ratings decreased already two weeks before onset of the episodes, but showed their most pronounced dip during depressive episodes. The weeks prior to depressive episodes were already significantly different from euthymic days, as were all depressed days. Changes towards (hypo)mania were numerically smaller and only differed significantly from euthymic days during the (hypo)manic episodes, not before. Generally, these results are not surprising, as ChronoRecord is a well-validated and internationally used tool that shows a high concurrent validity with depressive, hypomanic, and manic episodes, as assessed by the HAMD and YMRS (Bauer et al., 2004, 2008). Clinically, these results might indicate that it is more difficult for patients to spot upcoming (hypo)manic episodes than upcoming depressive episodes. Thus, adequate identification of personal EWS, potentially supported through professional supervision, might be especially crucial before (hypo)manic episodes.

When testing the predictive properties of both, self-reported current mood and AR of the same item in the same statistical model, findings remain rather similar compared to the single-predictor analyses. For instance, both self-reported mood as well as it is AR, significantly predict ongoing depressive weeks. This contradicts studies suggesting group differences between instability indices (such as emotional inertia) vanish when mean affect is included in those models (Bos, Jonge, & Cox, 2019; Dejonckheere et al., 2019). In our model, both predictors (mood and AR of mood) explain different parts of the variation in disorder status. However, these results also suggest that mean levels of affect might be at least as informative as complex measures, like AR,

https://doi.org/10.1017/S0033291723003811 Published online by Cambridge University Press

which has been reported previously. Some studies have even found that changes in mean levels of affect measures have a higher accuracy in predicting recurrence of depression than rises in AR (Smit et al., 2022; Smit & Snippe, 2022; Snippe, Smit, Kuppens, Burger, & Ceulemans, 2023).

Although not formally tested, eye-balling Table 3 favors self-reported mood as a prodromal sign for depressive episodes, whereas a reduced AR might prove valuable as a prodromal sign for (hypo)manic episodes. In clinical terms, this might suggest that patients may initially perceive mood instability before (hypo)manic episodes and observe a stabilization of their mood after the first week of hypomania. However, this should be further tested.

Strengths and limitations

While our sample size is small, the number of prospectively encountered episodes and the precision of episode detection are compelling. In total, there were 10 587 observed patient days with very high compliance rates for bi-weekly clinical interviews and daily e-diaries. Thus, the number of affective episodes captured during the study period markedly exceeds the number of affective episodes reported in previous studies on patients with BD (29 depressive and 20 [hypo]manic episodes). While previous AR studies mostly, but not always (e.g. Curtiss et al., 2023), assessed a set of EMA items several times per day, we only analyzed a single item that was measured once a day. However, this simpler end-of-day design may have helped the compliance rates during the long 12-month study period, resulting in the large number of observed affective episodes. Despite the relatively longer duration between adjacent observations when compared to other EMA studies, our results align rather well with those of previous studies. Specifically, although EWS was sometimes observed more frequently before transitions than in patients without transitions, not all, and mostly not even the majority of transitions were preceded by EWS (Bos et al., 2022; Helmich et al., 2023; Kunkels et al., 2023; Schreuder et al., 2022; Smit et al., 2022).

Importantly, we did not apply a correction for multiple testing to our analyses, as we were following an exploratory approach. However, even an application of the conservative Bonferroni correction for nine post-hoc tests per analysis, multiplying each p value with nine, would not relevantly change the main results or the conclusions we draw from them. Moreover, in this study we clustered hypomanic and manic episodes together due to statistical reasons. Future studies should investigate them separately, especially as the potential for EWS might be bigger for hypomanic or manic episodes than depressive episodes in BD. Furthermore, future studies might benefit from an a priori power analysis, which was not conducted in this study.

Some studies specifically looked at *increases* in AR as EWS of upcoming episodes, rather than absolute levels of AR (Smit et al., 2022), employing Mann-Kendal Tau's as a measure for rising AR. As online Supplementary analyses, we partly mimicked this approach and integrated the Mann–Kendall Tau's of AR in our multilevel models. However, findings, as reported in the online Supplementary Material Tables S1 and S2, did not support the hypothesis of critical slowing down either.

Lastly, an inherent issue is the lag of the AR measure. Due to, in our case, the 14-day window necessary to compute the AR, as a certain number of data points is necessary to get a good estimation of AR, temporal precision is lower than desired. E.g., in the first week of a depressive episode, at most half of the raw self-reported current mood items stem from actual depressed days, the rest are still values measured during prodromal days. While previous studies used even longer moving windows, e.g., 30 days, to calculate the AR, we shortened it to 14 days after sensitivity analyses revealed no substantial changes in the results as compared to a 30-day moving window.

Conclusion

Using an exploratory approach, we investigated the use of AR as an EWS for affective episodes in BD. Our results substantially add to the existing research on AR as EWS in BD by contributing longitudinal data with excellent temporal precision and an above-average number of affective episodes during the study period. The results of our analyses suggest AR might not be superior at predicting depressive episodes in patients with BD than simple self-report measures (e.g. asking for their current mood). The data suggest there might be potential in AR anticipating (hypo) manic episodes, which might be especially important as these were less easily noticed by patients themselves. As the analyses were exploratory, further studies are needed.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291723003811

Acknowledgements. None.

Funding statement. The study was funded by a personal university budget of UE-P (KIT).

Competing interests. UE-P reports consultancy for Boehringer-Ingelheim and speaker honorarium from Angelini Pharma. MB reports personal fees from Aristo Pharma GmbH, Hexal AG, Jannsen-Cilag GmbH, Neuraxpharm Arzneimittel GmbH, Novartis Pharma GmbH, Shire International GmbH, Sunovion Pharmaceuticals Inc., and Takeda Pharmaceutical Company. PR reports personal fees from Servier. The authors assert that the mentioned companies had no influence over the content of this article. VML, IR, EM, HH, and WES declare no competing interests.

Notes

1 As the main goal of our analysis was to investigate prodromal days and EWS, we dropped one criterion of the DSM-system. The DSM-5 states that an interval of at least two months needs to separate two episodes of the same polarity in order for them to be considered a recurrence and not part of the same episode. To make optimal use of our data, we counted two periods fulfilling the criteria for affective episodes as disconnected, as soon as 14 days of full symptom remission separated them. Accordingly, the number of affective episodes is not identical to the number of episodes previously reported for this data set (Ebner-Priemer et al., 2020), where we complied with all DSM-5 criteria.

References

Bauer, M., Grof, P., Gyulai, L., Rasgon, N., Glenn, T., & Whybrow, P. C. (2004). Using technology to improve longitudinal studies: Self-reporting with chrono record in bipolar disorder. *Bipolar Disorders*, *6*(1), 67–74. https://doi.org/10.1046/j.1399-5618.2003.00085.x

Bauer, M., Wilson, T., Neuhaus, K., Sasse, J., Pfennig, A., Lewitzka, U., … Whybrow, P. C. (2008). Self-reporting software for bipolar disorder: Validation of ChronoRecord by patients with mania. *Psychiatry Research*, *159*(3), 359–366. https://doi.org/10.1016/j.psychres.2007.04.013

Bech, P., Bolwig, T. G., Kramp, P., & Rafaelsen, O. J. (1979). The bech-rafaelsen mania scale and the hamilton depression scale. *Acta Psychiatrica Scandinavica*, *59*(4), 420–430. https://doi.org/10.1111/j.1600-0447.1979.tb04484.x

Bos, E. H., Jonge, P., & Cox, R. F. A. (2019). Affective variability in depression: Revisiting the inertia–instability paradox. *British Journal of Psychology*, *110* (4), 814–827. https://doi.org/10.1111/bjop.12372

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

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *The Journal of Nervous and Mental Disease*, *175*(9), 526–536. https://doi.org/10.1097/00005053-198709000-00004

Curtiss, J., Fulford, D., Hofmann, S. G., & Gershon, A. (2019). Network dynamics of positive and negative affect in bipolar disorder. *Journal of Affective Disorders*, *249*, 270–277. https://doi.org/10.1016/j.jad.2019.02.017

Curtiss, J. E., Mischoulon, D., Fisher, L. B., Cusin, C., Fedor, S., Picard, R. W., & Pedrelli, P. (2023). Rising early warning signals in affect associated with future changes in depression: A dynamical systems approach. *Psychological Medicine*, *53*(7), 3124–3132. https://doi.org/10.1017/S0033291721005183

Dablander, F., Pichler, A., Cika, A., & Bacilieri, A. (2022). Anticipating critical transitions in psychological systems using early warning signals: Theoretical and practical considerations. *Psychological Methods*. https://doi.org/10.1037/met0000450

Dejonckheere, E., Mestdagh, M., Houben, M., Rutten, I., Sels, L., Kuppens, P., & Tuerlinckx, F. (2019). Complex affect dynamics add limited information to the prediction of psychological well-being. *Nature Human Behaviour*, *3*, 478–491. Nature Publishing Group. https://doi.org/10.1038/s41562-019-0555-0

Ebner-Priemer, U. W., Mühlbauer, E., Neubauer, A. B., Hill, H., Beier, F., Santangelo, P. S., … Severus, E. (2020). Digital phenotyping: Towards replicable findings with comprehensive assessments and integrative models in bipolar disorders. *International Journal of Bipolar Disorders*, *8*(1), 1–9. https://doi.org/10.1186/s40345-020-00210-4

First, M. B., Williams, J. B., Karg, R. S., & Spitzer, R. L. (2015). *User's guide for the structured clinical interview for DSM-5 disorders, research version* (*SCID-5-RV*). Arlington: American Psychiatric Association.

GBD 2019 Mental Disorders Collaborators. (2022). Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. *The Lancet Psychiatry*, *9*(2), 137–150. https://doi.org/10.1016/S2215-0366(21)00395-3

Helmich, M. A., Smit, A. C., Bringmann, L. F., Schreuder, M. J., Oldehinkel, A. J., Wichers, M., … Snippe, E. (2023). Detecting impending symptom transitions using early-warning signals in individuals receiving treatment for depression. *Clinical Psychological Science*, *11*(6), 994–1010. https://doi.org/10.1177/21677026221137006

Houben, M., Van Den Noortgate, W., & Kuppens, P. (2015). The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological Bulletin*, *141*(4), 901–930. https://doi.org/10.1037/a0038822

Koval, P., Sütterlin, S., & Kuppens, P. (2016). Emotional inertia is associated with lower well-being when controlling for differences in emotional context. *Frontiers in Psychology*, *6*(Jan). https://doi.org/10.3389/fpsyg.2015.01997

Kunkels, Y. K., Riese, H., Knapen, S. E., Riemersma - van der Lek, R. F., George, S. V., van Roon, A. M., … Wichers, M. (2021). Efficacy of early warning signals and spectral periodicity for predicting transitions in bipolar patients: An actigraphy study. *Translational Psychiatry*, *11*(1), 350–360. https://doi.org/10.1038/s41398-021-01465-w

Kunkels, Y. K., Smit, A. C., Minaeva, O., Snippe, E., George, S. V., van Roon, A. M., … Riese, H. (2023). Risk ahead: Actigraphy-based early-warning signals of increases in depressive symptoms during antidepressant discontinuation. *Clinical Psychological Science*, *11*(5), 942–953. https://doi.org/10.1177/21677026221148101

Kuppens, P., Allen, N. B., & Sheeber, L. B. (2010). Emotional inertia and psychological maladjustment. *Psychological Science*, *21*(7), 984–991. https://doi.org/10.1177/0956797610372634

Mcintyre, R. S., Berk, M., Brietzke, E., Goldstein, B. I., López-Jaramillo, C., Kessing, L. V., … Mansur, R. B. (2020). Bipolar disorders. In www.thelancet.com (Vol. 396). Retrieved from www.thelancet.com

Montgomery, S. A., & Asberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, *134*(4), 382–389. https://doi.org/10.1192/bjp.134.4.382

Morriss, R., Faizal, M. A., Jones, A. P., Williamson, P. R., Bolton, C. A., & McCarthy, J. P. (2007). Interventions for helping people recognise early signs of recurrence in bipolar disorder. *Cochrane Database of Systematic Reviews*, *2010*(1). https://doi.org/10.1002/14651858.CD004854.pub2

Nelson, J., Klumparendt, A., Doebler, P., & Ehring, T. (2020). Everyday emotional dynamics in major depression. *Emotion (Washington, D.C.)*, *20*(2), 179–191. https://doi.org/10.1037/emo0000541

Panaite, V., Rottenberg, J., & Bylsma, L. M. (2020). Daily affective dynamics predict depression symptom trajectories among adults with major and minor depression. *Affective Science*, *1*(3), 186–198. https://doi.org/10.1007/s42761-020-00014-w

Perry, A., Tarrier, N., Morriss, R., McCarthy, E., & Limb, K. (1999). Randomised controlled trial of efficacy of teaching patients with bipolar disorder to identify early symptoms of relapse and obtain treatment. *British Medical Journal*, *318*(7177), 149–153. https://doi.org/10.1136/bmj.318.7177.149

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Retrieved March 1, 2023, from https://www.R-project.org/

Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W., Dakos, V., … Vandermeer, J. (2012). Anticipating critical transitions. *Science (New York, N.Y.)*, *338*(6105), 344–348. https://doi.org/10.1126/science.1225244

Schreuder, M. J., Hartman, C. A., Groen, R. N., Smit, A. C., Wichers, M., & Wigman, J. T. W. (2022). Anticipating transitions in mental health in at-risk youths: A 6-month daily diary study into early-warning signals. *Clinical Psychological Science*, *11*(6), 1026–1043. https://doi.org/10.1177/21677026221103138

Smit, A. C., Helmich, M. A., Bringmann, L. F., Oldehinkel, A. J., Wichers, M., & Snippe, E. (2022). Critical slowing down in momentary affect as early warning signal of impending transitions in depression. *PsyArXiv Preprints*.

Smit, A. C., & Snippe, E. (2022). Real-time monitoring of increases in restlessness to assess idiographic risk of recurrence of depressive symptoms. *Psychological Medicine*, *53*, 1–10. https://doi.org/10.1017/S0033291722002069

Snippe, E., Smit, A. C., Kuppens, P., Burger, H., & Ceulemans, E. (2023). Recurrence of depression can be foreseen by monitoring mental states with statistical process control. *Journal of Psychopathology and Clinical Science*, *132*(2), 145–155. https://doi.org/10.1037/abn0000812

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

Stone, A. A., & Shiffman, S. (1994). Ecological Momentary Assessment (Ema) in behavioral medicine. *Annals of Behavioral Medicine*, *16*(3), 199–202. https://doi.org/10.1093/abm/16.3.199

Trull, T. J., & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: Introduction to the special section. *Psychological Assessment*, *21*(4), 457–462. https://doi.org/10.1037/a0017653

Van De Leemput, I. A., Wichers, M., Cramer, A. O. J., Borsboom, D., Tuerlinckx, F., Kuppens, P., … Scheffer, M. (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences of the United States of America*, *111* (1), 87–92. https://doi.org/10.1073/pnas.1312114110

Wand, M. P., & Jones, M. C. (1995). *Kernel smoothing*. Boston, MA: Springer US. https://doi.org/10.1007/978-1-4899-4493-1

https://doi.org/10.1017/S0033291723003811 Published online by Cambridge University Press

Wichers, M., & Groot, P. C. (2016). Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*, *85* (2), 114–116. https://doi.org/10.1159/000441458

Wichers, M., Smit, A. C., & Snippe, E. (2020). Early warning signals based on momentary affect dynamics can expose nearby transitions in depression: A confirmatory single-subject time-series study. *Journal for Person-Oriented Research*, *6*(1), 1–15. https://doi.org/10.17505/jpor.2020.22042

Young, R. C., Biggs, J. T., Ziegler, V. E., & Meyer, D. A. (1978). A rating scale for mania: Reliability, validity and sensitivity. *British Journal of Psychiatry*, *133*(11), 429–435. https://doi.org/10.1192/bjp.133.5.429