Psychological Medicine

cambridge.org/psm

Rising early warning signals in affect associated with future changes in depression: a dynamical systems approach

Joshua E. Curtiss[1,2], David Mischoulon[1,2], Lauren B. Fisher[1,2], Cristina Cusin[1,2], Szymon Fedor[3], Rosalind W. Picard[3] and Paola Pedrelli[1,2]

[1]Depression Clinical and Research Program at Massachusetts General Hospital, Boston, MA, USA; [2]Harvard Medical School, Boston, MA, USA and [3]The Media Lab, Massachusetts Institute of Technology, Cambridge, MA, USA

Original Article

Cite this article: Curtiss JE, Mischoulon D, Fisher LB, Cusin C, Fedor S, Picard RW, Pedrelli P (2023). Rising early warning signals in affect associated with future changes in depression: a dynamical systems approach. *Psychological Medicine* **53**, 3124–3132. https://doi.org/10.1017/S0033291721005183

Received: 31 May 2021
Revised: 21 November 2021
Accepted: 29 November 2021
First published online: 23 December 2021

Key words:
Depression; time series; early warning signals; dynamics

Author for correspondence:
Joshua E. Curtiss,
E-mail: jcurtiss@mgh.harvard.edu

Abstract

Background. Predicting future states of psychopathology such as depressive episodes has been a hallmark initiative in mental health research. Dynamical systems theory has proposed that rises in certain 'early warning signals' (EWSs) in time-series data (e.g. auto-correlation, temporal variance, network connectivity) may precede impending changes in disorder severity. The current study investigates whether *rises* in these EWSs over time are associated with future changes in disorder severity among a group of patients with major depressive disorder (MDD).
Methods. Thirty-one patients with MDD completed the study, which consisted of daily smartphone-delivered surveys over 8 weeks. Daily positive and negative affect were collected for the time-series analyses. A rolling window approach was used to determine whether rises in auto-correlation of total affect, temporal standard deviation of total affect, and overall network connectivity in individual affect items were predictive of increases in depression symptoms.
Results. Results suggested that rises in auto-correlation were significantly associated with worsening in depression symptoms ($r = 0.41$, $p = 0.02$). Results indicated that neither rises in temporal standard deviation ($r = -0.23$, $p = 0.23$) nor in network connectivity ($r = -0.12$, $p = 0.59$) were associated with changes in depression symptoms.
Conclusions. This study more rigorously examines whether rises in EWSs were associated with future depression symptoms in a larger group of patients with MDD. Results indicated that rises in auto-correlation were the only EWS that was associated with worsening future changes in depression.

Introduction

Major depressive disorder (MDD) has a 16% lifetime prevalence rate and poses a substantial societal health burden as one of the leading causes of worldwide disability (Kessler et al., 2009). Moreover, depressive episodes have a high risk of recurrence, often eventuating in chronic suffering (Kendler, Thornton, & Gardner, 2000). Prediction of future states of psychopathology such as depressive episodes has been a hallmark initiative in mental health research (Bernardini et al., 2017). Although several decades of research have been devoted to elucidating risk factors for the onset of MDD, results have been limited to the identification of generic clinical and demographic predictors. Several methodological objections have been levied against the utility of such nomothetic risk factors in predicting future states of psychopathology. The generalizability of nomothetic research findings to any given individual is questionable (Fisher & Boswell, 2016), and static predictors fail to capture their highly dynamic and changeable nature of psychopathology over time (Nelson, McGorry, Wichers, Wigman, & Hartmann, 2017).

One potential reason for limited progress with predicting future depression may be our understanding of the disorder. Traditional conceptualizations postulate that a latent disease mechanism underlies the constellation of symptoms commonly associated with depression (e.g. low mood, anhedonia, sleep disturbances, etc.). However, recent perspectives on psychopathology acknowledge that mental disorders are complex networks of features that dynamically unfold and evolve over time (Hofmann, Curtiss, & McNally, 2016; Nelson et al., 2017; Scheffer et al., 2009, 2012; van de Leemput et al., 2014; Wichers et al., 2016). Prior researchers have posited that such a dynamical systems framework of mental disorders can facilitate precision medicine objectives of predicting disorder onset and identifying individually tailored treatments (Hayes et al., 2019; Hofmann, Curtiss, & Hayes, 2020). The dense time-series data necessitated by dynamical systems research afford more idiographic insights into prediction at the individual level.

Indeed, other research has leveraged longitudinal or idiographic time-series data in combination with various analytic methods such as growth curve modeling using multilevel modes (Wakefield, Delgadillo, Kellett, White, & Hepple, 2021), growth mixture modeling (Davies et al., 2020; Gueorguieva, Chekroud, & Krystal, 2017), machine learning modeling using generative embedding (Frässle et al., 2020), and best-fitting defined (linear, log-linear, 1-step) trajectories modeled for each patient to identify sudden gains (Helmich et al., 2020). One important limitation of prior research examining trajectories of depression over time is that such analytic approaches fail to capture the dynamic nature of depression using intensive time-series data. The aforementioned dynamical systems approach may offer a more theoretically and statistically innovative modeling framework that exploits dynamic time-series properties to predict depression outcomes (Hofmann et al., 2016; Nelson et al., 2017; Scheffer et al., 2012).

A dynamical systems approach to depression represents mental health as a complex system that occupies different attractor states of equilibrium (e.g. pathology or psychological health). When an attractor state exhibits stability, it tends to occupy a given state (e.g. psychological health) and be resilient to external influences that attempt to push it out of its state. However, when an attractor state lacks such stability, a small disturbance to the system (e.g. a life stressor) will more likely initiate a sudden shift in the system, causing it to occupy a new state of equilibrium (e.g. depression) (Cramer et al., 2016; Scheffer et al., 2009, 2012).

Rather than being gradual and linear, such transitions are often characterized by sudden changes and increased variability in system behavior, which is also known as critical fluctuations (Hayes, Laurenceau, Feldman, Strauss, & Cardaciotto, 2007). Evidence has supported the notion that improvement in response to psychotherapy occurs as discontinuous changes, the occurrence of which is preceded by critical instabilities such as critical fluctuations and other dynamic time-series indices (Schiepek, Tominschek, & Heinzel, 2014). Furthermore, recent emphasis has been placed on the identification of control parameters that cause changes between phases (Olthof, Hasselman, Oude Maatman, Bosman, & Lichtwarck-Aschoff, 2021). That is, what embedding and embodied constraints influence the system? For instance, Olthof et al. (2021) have proposed that idiographic processes that alter stress levels and stress vulnerability may be a candidate control parameter that dictates transitions into psychopathology.

According to dynamical systems theory, critical transitions from one state to another might be preceded by critical slowing down, which refers to an increasingly slow return to equilibrium after the system is impacted by small perturbations (Cramer et al., 2016; Hofmann et al., 2016; Scheffer et al., 2009, 2012). If the phenomenon were to be translated to mental health, then perhaps a person who is at risk for imminently developing depression may take longer and longer to return to a state of mental health after experiencing a small perturbation (e.g. life stressor). This slowness to recover may increase the likelihood that any given stressor will disturb the system enough to reach a tipping point, thereby precipitating a depressive episode.

Dynamical systems theory has proposed that certain 'early warning signals' (EWSs) are indicative of critical slowing down and impending shifts between states (Wichers et al., 2016). Specifically, such EWSs include time-series indices such as increasing temporal auto-correlation (i.e. the correlation between a variable at t and itself at t–1), increasing temporal variance, and increasing network connectivity (i.e. how strongly variables in a time-series are associated) (Liu, Chen, Aihara, & Chen, 2015).

Although critical slowing down and dynamical systems theory has been applied to more mature fields such as ecology and climate change (Scheffer et al., 2009), only recently have the concepts been articulated in mental health research, and there have been difficulties in identifying reliable clinical value in dynamical systems theory. For instance, Dablander, Pichler, Cika, and Bacilieri (2020) discuss how the ability of critical slowing down to predict depression transitions can be complicated by issues such as the fact that EWSs are sensitive to noise, critical slowing down can precede smooth transitions between stable states, and not all variables in a system will unanimously express critical slowing down prior to a sudden change.

Several studies have already examined depression from a dynamical systems perspective, accumulating tentative evidence that EWSs such as temporal auto-correlation tend to be higher among individuals who experience future increases in depression (Curtiss, Fulford, Hofmann, & Gershon, 2019; van de Leemput et al., 2014). Although the evidence from these studies is consistent with general tenets of dynamical systems theory, certain aspects have yet to be rigorously validated. Specifically, prior studies have only demonstrated that auto-correlation in certain variables is associated with depression changes when computing auto-correlation across all time-points in the entire time-series dataset. Dynamic systems theory more specifically postulates that *rises* in EWSs across time and within a specific individual provide evidence of critical slowing down (Hofmann et al., 2016; Scheffer et al., 2009). Methodologically, this would require a rolling window approach to time-series data analyses, whereby EWS metrics would be calculated within a single person sequentially several times over a given time window size (e.g. for a window size of 10, a time-series metric would be calculated from t_1 to t_{10}, t_2 to t_{11}, t_3 to t_{12}, etc., until the end of the time-series data). By adopting this approach, more direct evidence can be obtained to determine whether rises in EWSs can predict future changes in depression.

Wichers, Smit, and Snippe (2020) employed this rolling window approach in a single subject, substantiating that certain EWSs (i.e. auto-correlation and network connectivity) significantly increased over time before the patient experienced an abrupt exacerbation in depression symptoms. Other studies have suggested that rising increases in EWSs precede meaningful change in psychological treatments, such as increasing entropy values in child-therapist communication patterns preceding symptom reduction in anxiety (Lichtwarck-Aschoff, Hasselman, Cox, Pepler, & Granic, 2012), critical fluctuations (higher dynamic complexity) preceding symptom reduction in obsessive-compulsive disorder (Schiepek et al., 2014), increasing critical fluctuations (higher dynamic complexity) in psychotherapeutic processes predicting sudden gains and losses in psychotherapy for depressed patients (Olthof et al., 2020b), and increasing dynamic complexity and destabilization in therapy process variables preceding better treatment outcomes in depressed patients (Olthof et al., 2020a).

In the current study, a rolling window approach to examining EWSs will be pursued in a sample of patients with MDD. It may be that rises in certain EWSs are more associated with symptom change than others, suggesting predictive specificity. The current study extends previous research by providing a more robust examination of whether rising EWSs are related to future depression states. Although Wichers et al. (2020) found that all three EWS metrics were increasing before the onset of a depressive episode in a single patient, these metrics may not be predictive in general beyond that individual patient. In line with prominent models of emotional psychopathology positing dysregulation of negative affect and deficits in positive affect as core mechanisms underlying disorders such as depression (Hofmann, Sawyer, Fang, &

Asnaani, 2012), affect was the primary time-series variable used given its theoretical relevance. Thus, the primary objective of the current study is to determine which rising EWSs (i.e. temporal auto-correlation, temporal variance, or network connectivity) in time-series affect data are most reliably related to future exacerbations in symptoms in a larger sample of patients with MDD.

Methods

Participants and procedure

In the current study, 41 participants [74% female, mean age was 33.7 (S.D. = 14)] with MDD were initially enrolled, and 31 completed the study. Thus, all analyses are conducted on this final set of completers. Participants identified as the following: White = 22 (71%), Hispanic/Latino = 4 (23%), Asian = 5 (16%), Haitian/Black/African-American = 4 (12%), American Indian/Alaskan = 1 (3%), mixed-race = 2 (6%), and other = 1 (3%).

The study consisted of six in-person visits, daily smartphone-delivered surveys, and passive assessment over 8 weeks. At the first visit, we obtained informed consent and conducted the clinician-rated screening assessment to ascertain diagnoses and depression symptom severity. At the second visit (i.e. baseline visit), participants were assisted in downloading the monitoring app onto their smartphones, equipping them with wristband sensors, and administering clinical assessments. Clinical assessments were administered at the remaining four visits, which occurred once every 2 weeks over 8 weeks.

For study inclusion, participants were required to be diagnosed with current MDD (per the DSM-IV; APA, 2000) and have a Hamilton Depression Rating Scale (HDRS-28; Hamilton, 1960) score >18. Exclusionary criteria included substance use disorder within the past 3 months, psychosis, mania, and acute suicide risk. Full study details about participants and procedures have been described elsewhere (Pedrelli et al., 2020).

Measures

Positive and Negative Affect Schedule (PANAS)

The PANAS is a 20-item instrument that assesses positive and negative affect (Watson, Clark, & Tellegen, 1988). This measure was administered daily via patients' smartphones during the entire 8-week period. Both the total affect score (i.e. the sum score of all the affect variables) and individual affect items were used for the data analyses. This approach is consistent with Wichers, Groot, Psychosystems, and Group (2016), which used a total affect score for temporal auto-correlation and variance and individual items for networks.

Specifically, the networks were calculated using a 10-item subset of the PANAS with non-redundant items that measure well-characterized domains of positive and negative affect (Thompson, 2007). Of the 10 items, five indicate positive affect (i.e. 'active', 'alert', 'inspired', 'determined', and 'attentive'), and five measure negative affect (i.e. 'ashamed', 'upset', 'hostile', 'nervous', and 'afraid'). Prior research has demonstrated that these items evidenced good reliability and validity (Thompson, 2007).

Quick Inventory of Depression Symptomatology-Self Report (QIDS-SR)

The QIDS-SR is a 16-item self-report instrument that assesses overall depression severity, and higher scores indicate worse depressive symptoms (Rush et al., 2003). In the current study, the dependent variable was change in depressive symptoms over the course of the study expressed by the QIDS-SR change score (i.e. final endpoint QIDS-SR–baseline QIDS-SR). Thus, positive scores indicate worsening of depression symptoms across time, whereas negative scores indicate improvement in depression.

Clinical Global Impression Scale Improvement (CGI-I)

The CGI-I scale measures symptom improvement on a 1 (very much improved) to 7 (very much worse) scale (Guy, 1976). A score of 4 indicates no clinically meaningful improvement. For the context of the current study, the CGI-I was converted to a multi-categorical variable such that scores of 6 and above indicated clinically meaningful worsening of symptoms, scores between 3 and 5 indicated no substantial change, and scores of 2 and below indicated clinically meaningful improvement.

Statistical analyses

Prior to conducting any time-series analyses, each individual's time-series data were detrended using differencing, which computes differences between consecutive scores. This transformation procedure mitigates the influence of mean trends, an assumption underlying subsequent analyses (Dakos et al., 2012). Trends in absolute mean levels of affect may influence the window-based estimates of auto-correlation, standard deviation, and network connectivity. Detrending has been undertaken in other early warning studies for similar purposes (Wichers et al., 2020).

For the current study, the overarching approach to the time-series analyses consisted of an overlapping rolling window approach. That is, each EWS metric was computed for each person's time-series data in a smaller and fixed subset of time-points (i.e. a window), and they were recalculated for each sequential window until the last time-point was reached (e.g. for a window size of 18, a time-series metric would be calculated from t_1 to t_{18}, t_2 to t_{19}, t_3 to t_{20}, etc., until the end of the time-series data). Specifically, a window size of 18 time-points (i.e. 18 days) was selected to optimize the balance between having an adequate number of rolling window samples and time-points within each window, which were overlapping by 17 days. Given the number of complete time-points of data for each participant, which ranged from 19 days to full available data across the assessment period, window sizes of 15–25 days were tested to determine which window size would result in optimizing the number of participants who have a large enough number of rolling windows samples to be used for subsequent correlation analyses (i.e. have 5 or more rolling window samples). Window sizes of 15–18 days resulted in the exclusion of one participant due to an insufficient number of rolling window samples (i.e. <5). Window sizes larger than 18 days resulted in the loss of more participants. Thus, a final window size of 18 days was used because it was the largest window size associated with minimal loss of participants.

In other words, making the window size much larger resulted in having a smaller number of windows, which in turn would have undermined our ability to detect associations between each metric and the time index for each window (e.g. first window has time index of one, second window has time index of two, etc.). A much smaller window size would have reduced our power to estimate the EWS metrics.

Influenced by the work of Wichers et al. (2016), three standard metrics of critical slowing down were computed by examining EWSs: temporal auto-correlation, variability, and internode network connectivity. Temporal auto-correlation of the total PANAS score was computed using the *acf* function in R with a lag of 1 (i.e. within each time window the total PANAS at time t was correlated with itself at t–1). Variability was calculated using the standard deviation of the total PANAS score time-series within each window. Finally, network connectivity was estimated by way of vector auto-regression (VAR) analyses of each of the 10 non-redundant PANAS items. Specifically, overall connectivity was calculated by taking the absolute sum of all the β coefficients of the VAR analyses for each person. VAR analyses were conducted using the VAR function of the *vars* package in R (Pfaff, 2008).

Two approaches were used to determine whether statistically significant changes in EWS metrics occurring over time were associated with symptom outcomes. First, to assess whether changes over time in EWSs was associated with categorical disorder changes in general, χ^2 tests were used to assess the relationship between categorical symptoms outcomes as measured by the CGI-I (i.e. improvement, worsening, and no meaningful change) and whether or not there was a significant correlation between the EWS metric and time index (i.e. significant *v.* non-significant).

Second, to probe the exact nature of change, a more dimensional approach was adopted using changes in the QIDS-SR measure. Correlations between EWS metrics and their rolling window time index (i.e. the first window has a time index of 1, the second window has an index of 2, etc.) were conducted for each individual to evaluate whether such metrics significantly increase over each time window. Specifically, Kendall's τ parameter was estimated given that it specifically models rank ordering. Subsequently, participants' individual Kendall τ correlations between the EWSs and time indices were associated with changes in depression between follow-up and baseline, which would test whether individuals with rising EWSs are more likely to have future changes in depression severity. As a check to ensure depression measures were not better predicted by trivial features of the time-series affect data, the correlations between non-trended means of the affect data and time index were used to predict depression changes. The α was set to 0.05. Furthermore, for each EWS, we examined the number of participants whose metrics increased (i.e. significant and positive correlation between EWS metrics and their rolling window time index), decreased (i.e. significant and negative correlation between EWS metrics and their rolling window time index), and did not change over time (i.e. no significant correlation between EWS metrics and their rolling window time index).

For all analyses, all available time-series data were used, and incomplete time-points were removed. Available PANAS data ranged from some participants having all available data across the full 8-week time period to one participant having only 19 days of data. As indicated in the Results section, we excluded participants without sufficient time-points for an 18-day window size used for the rolling window analyses. Specifically, there were a number of participants who had three or fewer windows and were thus excluded. All analyses were conducted in R.

Results

Temporal auto-correlation

After applying the rolling window procedure to determine individual auto-correlations of total affect for each window within each individual, one participant's data were not usable due to not having enough time windows (i.e. only one time window). Of the 30 remaining participants, nine exhibited rising auto-correlations over time, seven had decreasing auto-correlations, and 14 had auto-correlations that did not significantly change over time. From the categorical perspective, symptom change category as measured by the CGI-I was not associated with statistically significant changes in auto-correlation over time ($\chi^2 = 7.47$, df = 2, $p = 0.03$). To probe the exact nature of this relationship, changes in QIDS-SR were associated with the correlations between auto-correlation and rolling time index. Overall, rises in temporal auto-correlation were significantly associated with worsening in depression symptoms ($r = 0.41$, $p = 0.02$) (Fig. 1). That is, participants who experienced rising auto-correlations over time in the total PANAS score were significantly more likely to experience deterioration of depression symptoms measured by the QIDS-SR.

Standard deviation

Likewise, 30 participants had sufficient data to extract individual standard deviation values using a rolling window approach (i.e. the same participant as mentioned above was excluded for having only one time window). Eight participants exhibited rising standard deviations over time, 14 exhibited decreasing standard deviations over time, and eight had standard deviations that did not significantly change over time. From the categorical perspective, symptom change category as measured by the CGI-I was not associated with statistically significant changes in standard deviation over time ($\chi^2 = 0.56$, df = 2, $p = 0.75$). To probe the exact nature of this relationship, changes in QIDS-SR were associated with the correlations between standard deviation and rolling time index. Temporal standard deviation of total affect was not significantly associated with worsening in depression symptoms using the QIDS-SR ($r = -0.23$, $p = 0.23$) (Fig. 2).

Network connectivity

After applying the rolling window procedure to calculate individual VAR networks containing the 10 non-redundant individual PANAS items, 10 participants' data were not usable due to insufficient time windows to calculate correlations over time (i.e. they had three or fewer time windows). Of the 21 remaining participants, three exhibited rising network connectivity over time, three exhibited decreasing network connectivity over time, and 15 had connectivity that did not significantly change over time. From the categorical perspective, symptom change category as measured by the CGI-I was not associated with statistically significant changes in network connectivity over time ($\chi^2 = 2.36$, df = 2, $p = 0.31$). From the dimensional perspective, changes in QIDS-SR were associated with the correlations between network connectivity and rolling time index. Temporal rises in network connectivity were not associated with changes in depression symptoms using the QIDS-SR ($r = -0.12$, $p = 0.59$) (Fig. 3).

Overall, across all three time-series metrics of interest (i.e. auto-correlation, standard deviation, and network connectivity), there were only three patients who experienced the same direction of change over time. Specifically, two patients experienced a decrease in all three metrics over time, whereas one patient experienced an increase in all three metrics over time.

Mean

The rolling window procedure to non-detrended data to calculate affect means for each window. From the categorical perspective, symptom change category as measured by the CGI-I was not associated with statistically significant changes in mean over

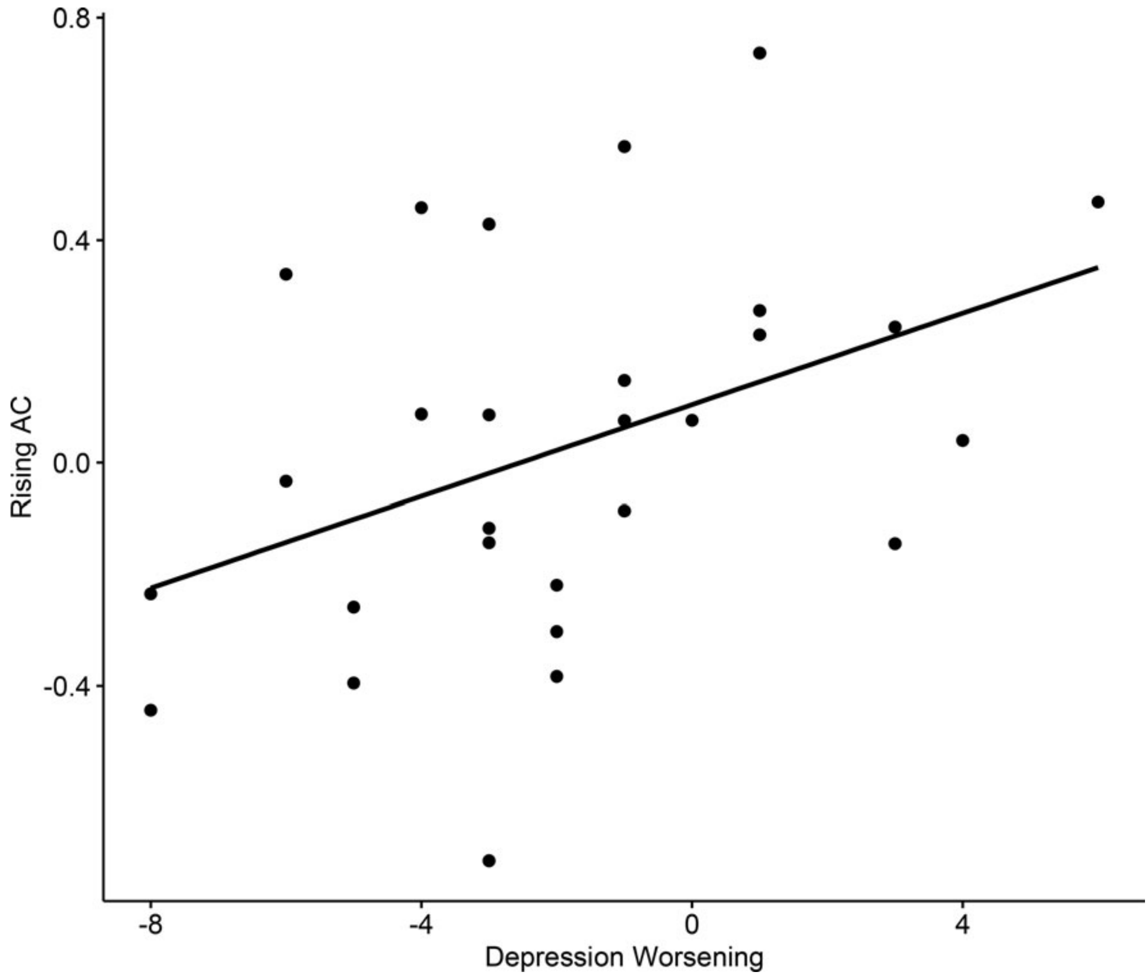

Fig. 1. Rising auto-correlation predicting depression changes. *Note:* Rising AC = Kendall's τ correlation between auto-correlation in each time window and the time window index. Rising temporal auto-correlation over time was associated with depression worsening, suggesting rising temporal auto-correlation predicts exacerbations in depression. Depression worsening is the difference between the baseline and post-assessment QIDS-SR score such that positive scores indicate depression deterioration and negative scores indicate improvement.

time ($\chi^2 = 0.53$, df = 2, $p = 0.77$). Moreover, changes in the mean of total affect over time were not significantly associated with worsening in depression symptoms using the QIDS-SR ($r = -0.28$, $p = 0.14$).

Discussion

In recent years, there has been a burgeoning interest in understanding depression from a dynamical systems perspective (Hayes et al., 2007; Hofmann et al., 2016; Schiepek et al., 2014; Wichers et al., 2016). Promising prior research has leveraged EWSs in symptom and affect time-series data to predict future depression severity (Curtiss et al., 2019; van de Leemput et al., 2014; Wichers et al., 2016). However, the current study more rigorously examined whether *rises* in EWSs over time were associated with future depression symptoms in a larger group of patients with MDD. Identifying more robust EWSs as predictors of depression outcome can better facilitate precision medicine approaches to depression. Results indicated that rises in temporal auto-correlation of overall affect over time are associated with worsening future changes in depression. Yet, neither rises in variance nor rises in network connectivity of overall affect over time predicted changes in depression. Of course, null results do not demonstrate that variance and network connectivity are unrelated to changes in depression; instead, there may not have been sufficient data to establish the presence of a relationship.

The findings on rising auto-correlation accord with prior literature and provide novel insight into potential dynamic time-series mechanisms underlying changes in depression. In prior studies that have examined dynamic properties across an entire time-series dataset, temporal auto-correlation was the most prominent EWS metric associated with future changes in depression (e.g. Curtiss *et al.*, 2019; van de Leemput *et al.*, 2014; Wichers *et al.*, 2016). Moreover, in the only studies examining *rising* EWSs, rises in temporal auto-correlation preceded exacerbations in depression for a single patient (Wichers et al., 2016; Wichers et al., 2020). Thus, results of the current study appear to corroborate the potential utility of auto-correlation as an important EWS metric. That notwithstanding, there are several other early earning signal metrics than the ones considered in the current study that would benefit from additional examination (e.g. dynamic complexity; Olthof et al., 2020b). A precision medicine approach to predicting worsening depression may be bolstered by examining rising auto-correlation as a relevant EWS of worsening depression. A corollary of our results is that improvements in depression are preceded by decreasing auto-correlation over

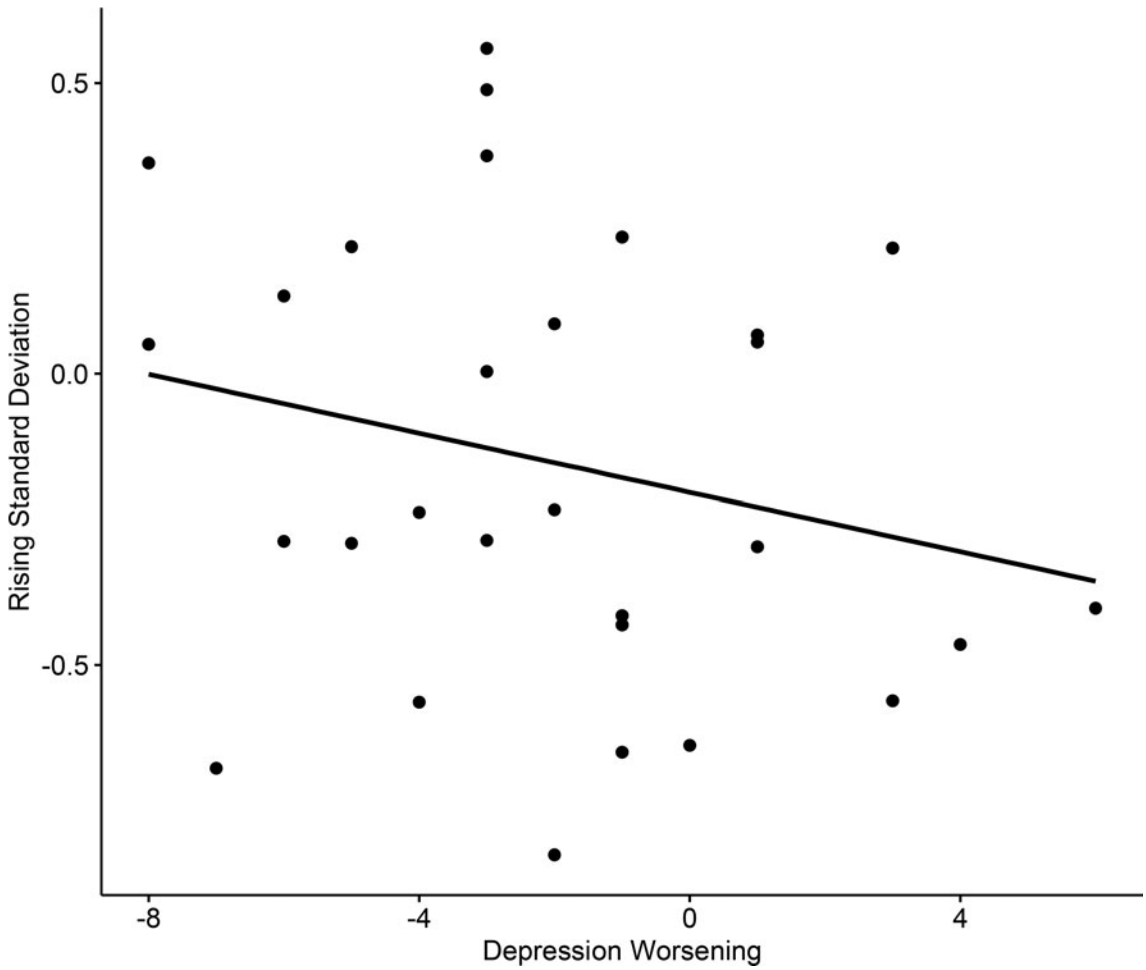

Fig. 2. Rising standard deviation predicting depression changes. *Note*: Rising standard deviation = Kendall's τ correlation between standard deviation in each time window and the time window index. Depression worsening is the difference between the baseline and post-assessment QIDS-SR score such that positive scores indicate depression deterioration and negative scores indicate improvement.

time, suggesting that both rises and decreases in auto-correlation have utility as signals of different depression outcomes. This might suggest that temporal auto-correlation is not serving as an EWS of critical slowing down in the traditional sense. If critical slowing down were occurring as specified according to the strict tenets of dynamical systems theory, then rising auto-correlations would likely precede both deterioration and improvement. In ecology, for instance, decreases in auto-correlation and skewness of water quality data preceded sudden transitions to unhealthy lake conditions, which is more related to the phenomenon of flickering rather than critical slowing down (Wang et al., 2012). Related to the current study, perhaps the decreasing auto-correlation preceding improvements in depression may bear a stronger resemblance to flickering, which compels a system to alternate back and forth between different states in response to relatively large impacts (Wang et al., 2012).

We did not replicate the finding that rises in variance of total affect and network connectivity predict changes in depression (Wichers et al., 2016, 2020). A primary difference between the current study and the two prior studies that examined the same rising EWSs is that the latter examined these metrics only in single patients. The current study used a full sample of depressed patients to determine how robust and consistent all three EWS metrics were in predicting future changes in depression. Indeed, for some patients in the current study, support was found for rises in standard deviation ($n = 8$) and network connectivity ($n = 3$), yet this pattern was not robust enough across the population for these metrics to be more general predictors of changes in depression symptoms. Although there was specificity in which rising EWSs more reliably predict changes in depression in the current study, this does not necessarily indicate that there is always specificity in which time-series dynamics predict depression changes. For instance, it could be the case that the current study was underpowered to detect the other EWSs, or that critical slowing down, as evidenced by rises in all three metrics, does not hold for depression in the robust way dictated by dynamical systems theory.

Collectively, findings from this study and prior research motivate a more nuanced framework for understanding the role of EWSs in predicting depression outcomes. A common conclusion is that the predictive utility of *rising* EWSs is supported, which is an important finding bearing on the theoretical tenets of dynamical systems theory (Hofmann et al., 2016; Scheffer et al., 2009). Moreover, when EWSs are studied beyond a single subject, rising temporal auto-correlation appears to be the most reliable metric significantly emerging in multiple people. Thus, auto-correlation may be a more generic and robust EWS. However, it is important to note that idiographic research supports the predictive utility of other rising EWSs for individual patients. It may be the case that time-series metrics that are predictive for one patient may not

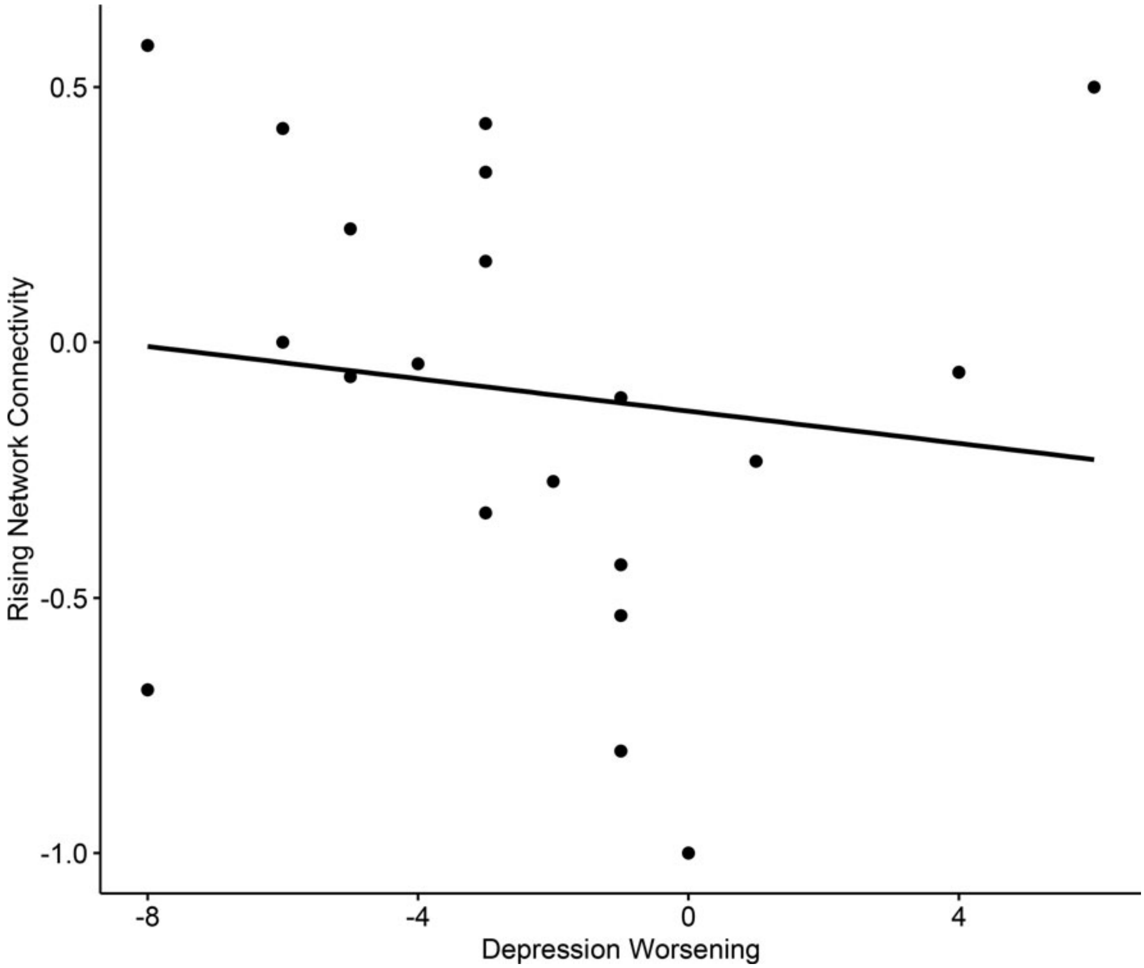

Fig. 3. Rising network connectivity predicting depression changes. *Note*: Rising network connectivity = Kendall's τ correlation between network connectivity in each time window and the time window index. Depression worsening is the difference between the baseline and post-assessment QIDS-SR score such that positive scores indicate depression deterioration and negative scores indicate improvement.

necessarily bear similar utility for another patient. Although examining the predictive utility of every single time-series metric for each patient is neither feasible nor appropriate, consideration of individual differences in theory-driven dynamic metrics to a certain extent could still be useful.

This study corroborates rising auto-correlation as a potential pathway forward in augmenting our ability to predict meaningful outcomes relevant to depression. In precision medicine, it is important to identify which features more reliably predict clinical outcomes across groups of people and which features are idiographically useful, arising in one patient but not another. For time-series data to better inform precision medicine approaches to depression, more research is required to better understand under what circumstances are certain time-series metrics predictive of depression change. If time-series analysis procedures can be developed to identify which unique dynamic metrics are most predictive of depression outcomes for individual patients, then this may be to bolster precision medicine science for depression.

Another important issue worthy of consideration is a better conceptualization of phase changes in mental disorders. Dynamical systems theory posits that EWSs should precede phase changes (i.e. the qualitative transition from one state to another). In mental health research, a phase transition would represent a qualitative change between different attractor states of health or pathology. Nonetheless, it is not necessarily clear which attractor state a system will enter given the appearance of EWSs. Indeed prior studies have found that EWSs are related to both improvement and deterioration in depression (Olthof et al., 2020b; van de Leemput et al., 2014). In the current study, there was some evidence that significantly changing auto-correlations over time were associated with clinically significant categorical changes in symptoms. Moreover, it appeared that rising levels of temporal auto-correlations over time preceded deterioration in depression rather than improvement. This may be the case because this study only involved assessment and did not provide any sort of treatment. Future work into dynamical system approaches to mental health disorders might benefit from more explicit considerations of why certain transitions occur over others.

Relatedly it would be profitable for future dynamical systems research in depression to examine what drivers are underlying changes. For instance, Dablander et al. (2020) mention that in the case of depression this would require assessing underlying variables (e.g. stress) that may drive critical transitions in the symptom variables. There may be multiple drivers underlying transitions, and identification of appropriate drivers requires an adequate theoretical and conceptual foundation of the dynamic behavior of depression. Furthermore, it is important to understand the nature of the relationship between the driver variables and symptom variables. For instance, EWSs preceding critical transitions assume a nonlinear

https://doi.org/10.1017/S0033291721005183 Published online by Cambridge University Press

relationship (i.e. small changes in a driver ideally should be associated with large changes in the state variable of interest, such as depression severity) (Dablander et al., 2020). It is important that future research acknowledge these issues in the context of mental health disorders.

Our findings must be interpreted in light of certain limitations. First, affect measurements were only assessed daily, in which one-half of the questions were measured earlier in the day and the other half later in the day. Thus, our methodology bears a stronger resemblance to a daily diary study than a more intensive ecological momentary assessment study, in which more rich time-series data would be collected. Second, although the rolling windows approach was an innovative methodology for the time-series data analysis, the size of the moving window (i.e. 18 days) was smaller than would be desired. This window size was selected to optimize the balance between having adequate rolling window samples and time-points within each window. Although small window sizes might be advantageous for just-in-time clinical interventions, such a window size might make it more difficult to detect meaningful effects with the VAR analyses used to compute network connectivity. Third, future studies would benefit from measuring disorder severity more frequently and for greater periods of time to better detect exactly when EWSs precede exacerbations and improvements in depression. Longer term monitoring of changes in depression, including both recovery and relapse, can facilitate our understanding of whether the auto-correlation, variance, and network connectivity re-occur in a person-dependent way over time. Fourth, the current study examined symptom changes and clinical change across the entirety of the study. Thus, it remains difficult to determine whether changes in the time-series metrics of affect specifically preceded a large change in symptoms. The current study does not have sufficient measurement occasions to rigorously examine whether EWSs are markedly identifiable right before a sudden transition. Future research would require more frequent assessment over longer periods of time to identify the extent to which EWSs immediately precede critical transitions. These limitations notwithstanding, the current study provides further evidence that rises in certain EWSs (i.e. temporal auto-correlation) appear to be relatively reliable predictors of depression change across a larger sample of depressed patients, affording insights into how dynamical systems procedures can inform a precision medicine framework for depression.

Financial support. This study was supported by the MGH-MIT Strategic Partnership Grand Challenge Grant, by the R01MH118274 grant from the National Mental Health Institute, by European Union's FP7 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 327702, and by Abdul Latif Jameel Clinic for Machine Learning in Health.

Conflict of interest. RP served as cofounder and chairman of the board for Empatica, which manufactured the wearable sensors used to collect a subset of the data used in the study. She owned stock in Empatica and served as part-time consultant and chief scientist for them. PP also received royalties from MIT for patents. She was an inventor related to wearable technology; however, none of these are directly related to this work. DM has received research support from Nordic Naturals and heckel medizintechnik GmbH. He has received honoraria for speaking from the Massachusetts General Hospital Psychiatry Academy, Harvard Blog, and PeerPoint Medical Education Institute, LLC. He also works with the MGH Clinical Trials Network and Institute (CTNI), which has received research funding from multiple pharmaceutical companies and NIMH. JEC receives royalties from self-help books; however, none of these books are related to the current work. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed). Washington, DC: American Psychiatric Association.

Bernardini, F., Attademo, L., Cleary, S. D., Luther, C., Shim, R., Quartesan, R., & Compton, M. T. (2017). Risk prediction models in psychiatry: Toward a new frontier for the prevention of mental illnesses. *Journal of Clinical Psychiatry*, *78*, 572–583.

Cramer, A. O., Van Borkulo, C. D., Giltay, E. J., Van Der Maas, H. L., Kendler, K. S., Scheffer, M., & Borsboom, D. (2016). Major depression as a complex dynamic system. *PLoS ONE*, *11*(12), e0167490.

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.

Dablander, F., Pichler, A., Cika, A., & Bacilieri, A. (2020). Anticipating critical transitions in psychological systems using early warning signals: Theoretical and practical considerations. *Psycharxiv*, 1–20.

Dakos, V., Carpenter, S. R., Brock, W. A., Ellison, A. M., Guttal, V., Ives, A. R., … Scheffer, M. (2012). Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS ONE*, *7*(7), e41010.

Davies, S. E., Neufeld, S. A., van Sprang, E., Schweren, L., Keivit, R., Fonagy, P., … Goodyer, I. M. (2020). Trajectories of depression symptom change during and following treatment in adolescents with unipolar major depression. *Journal of Child Psychology and Psychiatry*, *61*, 565–574.

Fisher, A. J., & Boswell, J. F. (2016). Enhancing the personalization of psychotherapy with dynamic assessment and modeling. *Assessment*, *23*, 496–506.

Frässle, S., Marquand, A. F., Schmaal, L., Dinga, R., Veltman, D. J., Van der Wee, N. J., … Stephan, K. E. (2020). Predicting individual clinical trajectories of depression with generative embedding. *NeuroImage: Clinical*, *26*, 102213.

Gueorguieva, R., Chekroud, A. M., & Krystal, J. H. (2017). Trajectories of relapse in randomised, placebo-controlled trials of treatment discontinuation in major depressive disorder: An individual patient-level data meta-analysis. *The Lancet Psychiatry*, *4*, 230–237.

Guy, W. (1976). *ECDEU Assessment Manual for Psychopharmacology*. Rockville, MD: U.S. Dept. of Health, Education, and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, National Institute of Mental Health, Psychopharmacology Research Branch, Division of Extramural Research Programs in Rockville, MD, 217–222.

Hamilton, M. (1960). The Hamilton Depression Scale – accelerator or break on antidepressant drug discovery. *Psychiatry*, *23*, 56–62.

Hayes, A. M., Laurenceau, J. P., Feldman, G., Strauss, J. L., & Cardaciotto, L. (2007). Change is not always linear: The study of nonlinear and discontinuous patterns of change in psychotherapy. *Clinical Psychology Review*, *27*(6), 715–723.

Hayes, S. C., Hofmann, S. G., Stanton, C. E., Carpenter, J. K., Sanford, B. T., Curtiss, J. E., & Ciarrochi, J. (2019). The role of the individual in the coming era of process-based therapy. *Behaviour Research and Therapy*, *117*, 40–53.

Helmich, M. A., Wichers, M., Olthof, M., Strunk, G., Aas, B., Aichhorn, W., … Snippe, E. (2020). Sudden gains in day-to-day change: Revealing nonlinear patterns of individual improvement in depression. *Journal of Consulting and Clinical Psychology*, *88*(2), 119–127.

Hofmann, S. G., Curtiss, J., & McNally, R. J. (2016). A complex network perspective on clinical science. *Perspectives on Psychological Science*, *11*, 597–605.

Hofmann, S. G., Curtiss, J. E., & Hayes, S. C. (2020). Beyond linear mediation: Toward a dynamic network approach to study treatment processes. *Clinical Psychology Review*, *76*, 101824.

Hofmann, S. G., Sawyer, A. T., Fang, A., & Asnaani, A. (2012). Emotion dysregulation model of mood and anxiety disorders. *Depression and Anxiety*, *29*, 409–416.

Kendler, K. S., Thornton, L. M., & Gardner, C. O. (2000). Stressful life events and previous episodes in the etiology of major depression in women: An evaluation of the 'kindling' hypothesis. *American Journal of Psychiatry*, *157*, 1243–1251.

Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Chatterji, S., Lee, S., Ormel, J., … Wang, P. S. (2009). The global burden of mental disorders: An update from the WHO World Mental Health (WMH) Surveys. *Epidemiology and Psychiatric Sciences*, *18*, 23–33.

Lichtwarck-Aschoff, A., Hasselman, F., Cox, R., Pepler, D., & Granic, I. (2012). A characteristic destabilization profile in parent-child interactions associated with treatment efficacy for aggressive children. *Nonlinear Dynamics, Psychology, and Life Sciences*, *16*, 353–379.

Liu, R., Chen, P., Aihara, K., & Chen, L. (2015). Identifying early-warning signals of critical transitions with strong noise by dynamical network markers. *Scientific Reports*, *5*, 1–13.

Nelson, B., McGorry, P. D., Wichers, M., Wigman, J. T., & Hartmann, J. A. (2017). Moving from static to dynamic models of the onset of mental disorder: A review. *JAMA Psychiatry*, *74*, 528–534.

Olthof, M., Hasselman, F., Oude Maatman, F., Bosman, A. M. T., & Lichtwarck-Aschoff, A. (2021). Complexity theory of psychopathology [Preprint]. *PsyArXiv*. https://doi.org/10.31234/osf.io/f68ej.

Olthof, M., Hasselman, F., Strunk, G., Aas, B., Schiepek, G., & Lichtwarck-Aschoff, A. (2020a). Destabilization in self-ratings of the psychotherapeutic process is associated with better treatment outcome in patients with mood disorders. *Psychotherapy Research*, *30*, 520–531.

Olthof, M., Hasselman, F., Strunk, G., van Rooij, M., Aas, B., Helmich, M. A., … Lichtwarck-Aschoff, A. (2020b). Critical fluctuations as an early-warning signal for sudden gains and losses in patients receiving psychotherapy for mood disorders. *Clinical Psychological Science*, *8*, 25–35.

Pedrelli, P., Fedor, S., Ghandeharioun, A., Howe, E., Ionescu, D. F., Bhathena, D., … Picard, R. W. (2020). Monitoring changes in depression severity using wearable and mobile sensors. *Frontiers in Psychiatry*, *11*, 1413.

Pfaff, B. (2008). VAR, SVAR and SVEC models: Implementation within R package vars. *Journal of Statistical Software*, *27*, 1–32.

Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., … Keller, M. B. (2003). The 16-item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*, *54*, 573–583.

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., … Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, *461*, 53–59.

Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W., Dakos, V., … Pascual, M. (2012). Anticipating critical transitions. *Science (New York, N.Y.)*, *338*, 344–348.

Schiepek, G. K., Tominschek, I., & Heinzel, S. (2014). Self-organization in psychotherapy: Testing the synergetic model of change processes. *Frontiers in Psychology*, *5*, 1089.

Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). *Journal of Cross-Cultural Psychology*, *38*, 227–242.

van de Leemput, I. A., Wichers, M., Cramer, A. O., 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*, *111*, 87–92.

Wakefield, S., Delgadillo, J., Kellett, S., White, S., & Hepple, J. (2021). The effectiveness of brief cognitive analytic therapy for anxiety and depression: A quasi-experimental case–control study. *British Journal of Clinical Psychology*, *60*(2), 194–211.

Wang, R., Dearing, J. A., Langdon, P. G., Zhang, E., Yang, X., Dakos, V., & Scheffer, M. (2012). Flickering gives early warning signals of a critical transition to a eutrophic lake state. *Nature*, *492*(7429), 419–422.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070.

Wichers, M., Groot, P. C., Psychosystems, E. S. M., & Group, E. W. S. (2016). Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*, *85*, 114–116.

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*, 1–15.