Journal of Psychopathology and Clinical Science

© 2023 American Psychological Association
ISSN: 2769-7541

2023, Vol. 132, No. 2, 145–155
https://doi.org/10.1037/abn0000812

Recurrence of Depression Can Be Foreseen by Monitoring Mental States With Statistical Process Control

Evelien Snippe[1], Arnout C. Smit[1, 2], Peter Kuppens[3], Huibert Burger[4], and Eva Ceulemans[3]

[1] Department of Psychiatry, Interdisciplinary Center Psychopathology and Emotion Regulation, University Medical Center Groningen, University of Groningen

[2] Clinical Psychology, Faculty of Behavioral and Movement Sciences, VU Amsterdam

[3] Faculty of Psychology and Educational Sciences, Quantitative Psychology and Individual Differences, KU Leuven

[4] Department of General Practice and Elderly Care Medicine, University Medical Center Groningen, University of Groningen

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Detecting early signs of recurrence of psychopathology is key for prevention and treatment. Personalized risk assessment is especially relevant for formerly depressed patients, for whom recurrence is common. We aimed to examine whether recurrence of depression can be accurately foreseen by applying Exponentially Weighted Moving Average (EWMA) statistical process control charts to Ecological Momentary Assessment (EMA) data. Participants were formerly depressed patients ($n = 41$) in remission who (gradually) discontinued antidepressants. Participants completed five smartphone-based EMA questionnaires a day for 4 months. EWMA control charts were used to prospectively detect structural mean shifts in high and low arousal negative affect (NA), high and low arousal positive affect (PA), and repetitive negative thinking in each individual. A significant increase in repetitive negative thinking (worry, negative thoughts about the self) was the most sensitive early sign of recurrence: this was detected in 18 out of 22 patients (82%) before recurrence and in 8 out of 19 patients (42%) who stayed in remission. A significant increase in NA high arousal (stress, irritation, restlessness) was the most specific early sign of recurrence: this was detected in 10 out of 22 patients (45%) before recurrence and in 2 out of 19 patients (11%) who stayed in remission. These mean changes were detected at least a month before recurrence in the majority of the participants. The outcomes were robust across EWMA parameter choices, but not when using fewer observations per day. The findings demonstrate the value of monitoring EMA data with EWMA charts for detecting prodromal symptoms of depression in real-time.

General Scientific Summary

The findings of this study suggest that significant rises in repetitive negative thinking and high arousal negative emotions are early signs of recurrence of depression. Such increases can be detected reliably in real-time using person-specific statistical process control charts in formerly depressed patients.

Evelien Snippe https://orcid.org/0000-0002-3003-7475
Arnout C. Smit https://orcid.org/0000-0001-9465-8687
Peter Kuppens https://orcid.org/0000-0002-2363-2356
Huibert Burger https://orcid.org/0000-0001-6484-0349
Eva Ceulemans https://orcid.org/0000-0002-7611-4683

This project has received funding from the Netherlands Organisation for Health Research and Development (ZonMw Off Road; Project 451001029 (40-08125-98-18041) to E. Snippe), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovative programme (ERC-CoG-2015; No 681466 to M. Wichers), and the Research Council KU Leuven (Project C14/19 to E. Ceulemans and P. Kuppens). The ZonMw commission provided feedback on conceptualization of the study. The TRANS-ID Tapering study (NL58469.042.16) was approved by the Medical Ethical Committee of the University Medical Center Groningen. A detailed study protocol including the used study materials is available on OSF (see https://osf.io/zbwkp/). The analysis plan of the current study was preregistered on OSF (see https://osf.io/sbyq7/). The data are not freely available in a public repository due to restrictions related to the European Law regarding General Data Protection Regulation (GDPR). Collaboration on TRANS-ID Tapering data is possible and can be requested via h.riese@umcg.nl. The study analysis code will be made available upon request (eveliensnippe@gmail.com). The study results were presented at the Society for Ambulatory Assessment conference (2022), at the spring conference of the Dutch society for Psychiatry (2022), at the University of Trier (2021), and at the Belgian-Dutch ESM Expert Network Meeting (2022). The authors thank the participants for their time and effort, P. Harder for assistance with the recruitment of participants, M. Messchendorp and R. de Vries for data collection, Marieke Wichers, Harriëtte Riese, Marieke A. Helmich, and Yoram K. Kunkels for co-designing the TRANS-ID Tapering study, and Timon Elmer for assistance with data visualization.

We have no conflict of interest to disclose.

Evelien Snippe served as lead for Conceptualization, Formal analysis, Methodology, Project administration, Visualization, Writing–original draft and served in a supporting role for Data curation, Investigation, Software. Arnout C. Smit: Conceptualization, Software, Data curation, Investigation, Methodology, Writing–review and editing. Peter Kuppens: Writing–review and editing. Huibert Burger: Methodology, Writing–review and editing. Eva Ceulemans: Conceptualization, Methodology, Supervision, Writing–review and editing.

Correspondence concerning this article should be addressed to Evelien Snippe, Interdisciplinary Center Psychopathology and Emotion Regulation (ICPE), University Medical Center Groningen, Hanzeplein 1, PO Box 30.001, 9700 RB Groningen, The Netherlands. Email: eveliensnippe@gmail.com

Keywords: experience sampling methodology, replicated single-subject design, idiographic, prediction, prognostic model

Supplemental materials: https://doi.org/10.1037/abn0000812.supp

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Mental disorders present a huge cost for the individual and society. If one were able to foresee episodes of psychopathology before they occur, this would be of great clinical benefit for prevention and treatment. This quest for personalized early detection is especially relevant for disorders with a high risk of recurrence, such as depression (Richards, 2011).

Research so far has not provided methods that can determine whether a specific patient's depressive symptoms are about to return. Research has predominantly focused on group-level studies instead of studies at the individual level (Barlow & Nock, 2009; Hamaker, 2012; Piccirillo & Rodebaugh, 2019). For example, studies showed that the risk of depressive relapse is higher in patients with more residual symptoms (Berwian et al., 2017; Buckman et al., 2018). However, such studies do not provide insight in the risk of recurrence in specific individuals and ignore the likely possibility that an individual's risk may change over time. Methods to detect early signs and symptoms that precede the onset of a fully manifested depressive episode (i.e., prodromal symptoms; Benasi et al., 2021; Fava & Tossani, 2007) may help to determine whether an individual's risk of an upcoming depression is increasing.

We propose a method to detect early markers of recurrence of depression in real-time using the Exponentially Weighted Moving Average (EWMA) control chart (Montgomery, 2012). EWMA charts are a type of statistical process control charts with which small structural mean changes can be detected in time-series data. Statistical process control is the industry standard to determine whether and when the quality or performance of products goes significantly out-of-control, allowing for timely intervention. For example, EWMA charts can be used to monitor parts of the motor of an airplane and to raise an alert when the parts start to perform poorly, in order to prevent the whole motor from breaking down. By analogy, EWMA charts could be a promising way to foresee recurrence of depression (i.e., the motor break down) by monitoring change in components of depression (i.e., the parts of the motor) and raise an alert when a significant increase in these components is detected. This may allow the detection of early signs of recurrence (i.e., prodromal symptoms).

Previous between-person research showed that a broad range of symptoms predict future depression, such as higher levels of anxiety (Hettema et al., 2006; N. C. Jacobson & Newman, 2017), decreased experience of positive affect (Geschwind et al., 2010; Raes et al., 2012), increased experience of negative affect (Minaeva et al., 2021; Neubauer et al., 2022; Vaessen et al., 2017), and repetitive negative thinking (Spinhoven et al., 2018). Among the most commonly reported prodromal symptoms of depression are anxiety related symptoms, such as irritability, sleep problems, tension, and somatic complaints, as well as symptoms of exhaustion, such as fatigue and reduced energy (Benasi et al., 2021). It has yet to be examined whether these symptoms also significantly increase within-person preceding recurrence of depression, and thus, whether they may serve as early indicators of recurrence.

To detect within-person changes in these plausible prodromal symptoms, EWMA charts could be applied to long-term Ecological Momentary Assessment (EMA) data (Shiffman et al., 2008). EMA allows one to assess affective and cognitive states reflecting plausible prodromal symptoms, such as irritability or worrying, multiple times a day in the flow of their daily life (Myin-Germeys et al., 2009). Applying EWMA charts to long-term EMA data yields an idiographic method that can determine whether and when a specific person's affective and cognitive states start to deteriorate in real-time and thereby possibly capturing early signs of recurrence of depression. This is illustrated in Figure 1, which depicts the application of an EWMA control chart to EMA data of high arousal Negative Affect (NA) in a single person from the current study. In this individual, a significant increase in the EWMA of high arousal NA was detected more than 6 weeks before this patient experienced recurrence of depression.

In a first proof of principle study, we applied EWMA charts to long-term EMA data on the negative affective state "feeling restless" of seven formerly depressed patients (Smit et al., 2019). In two patients, a significant within-person increase in the mean of restlessness was detected with EWMA charts 2 months before recurrence of depressive symptoms. In five patients without a recurrence of symptoms, no such an increase was found. In a next confirmatory study on the Transitions In Depression (TRANS-ID) Tapering data (Smit et al., 2020), significant increases in restlessness were detected using EWMA charts in 68.3% of patients who transitioned to higher levels of depressive symptoms, whereas no such increases were detected in 73.7% of the patients who stayed in remission (Smit & Snippe, 2022).

Even though these results are promising, several questions are yet to be answered. First, only the state "restlessness" was examined despite the fact that there are other candidate prodromal symptoms, as outlined before. Thus, a more comprehensive approach would be to monitor change in multiple candidate affective and cognitive states using EWMA charts.

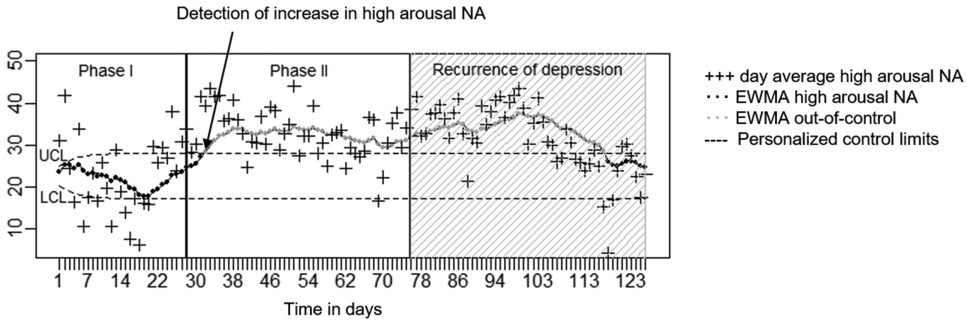
Second, there might be individual differences in the type of cognitive and affective changes that precede recurrence of depression, as suggested by current personalized views on psychopathology as well as empirical evidence regarding psychological processes (e.g., Barlow & Nock, 2009; Hamaker, 2012; Piccirillo & Rodebaugh, 2019; Snippe et al., 2015). It is unclear whether specific affective and cognitive states deteriorate in the period preceding recurrence across individuals. Another possibility is that depression can be better foreseen when considering heterogeneity, e.g., by examining whether at least one or more mental states deteriorate, regardless of the type of mental state.

Third, it is not clear what the effect is of the EWMA tuning parameters on the robustness of the outcomes of the EWMA charts. Previous research shows that variations in how time-series analyses are performed can have major effects on the outcomes (e.g., Bastiaansen et al., 2020). The predictive accuracy of the EWMA charts may depend on the width of the person-specific control limits, as has been shown in other fields of health care (Baldewijns et al., 2016; Ilies et al., 2020).

Fourth, for implementation in clinical practice, it is important to examine whether the burden of monitoring using EMA can be

Figure 1

EWMA Control Chart of High Arousal Negative Affect (NA) of One Study Participant



Note. High arousal Negative Affect (NA) was measured five times a day over 4 months with the items "feeling restless", "feeling irritated", and "feeling stressed". EWMA = Exponentially Weighted Moving Average of the day average of high arousal NA. The assessments during the first 28 days (phase I) were used to compute this person's mean and standard deviation of high arousal NA. These parameters were used to compute a threshold, i.e., an Upper Control Limit (UCL), above which high arousal NA was considered to have increased significantly (i.e., being out-of-control). In phase II, we continuously examined if and when the EWMA of high arousal NA exceeded this threshold. The EWMA exceeded the UCL for the first time on day 33; this is regarded as the moment a significant increase in high arousal NA was detected. Recurrence of depression (on day 76) was determined based on a reliable and meaningful increase in depressive symptoms using weekly depressive symptoms questionnaires and qualitative patient reports (see "Methods" section).

reduced. Previous studies showed that a higher frequency of assessments a day can be perceived as burdensome for participants (Bos et al., 2020; Palmier-Claus et al., 2011) and is associated with lower compliance (Vachon et al., 2019). Thus, it is important to establish how many assessments a day are needed to accurately foresee recurrence.

The overall aim is to empirically establish the EWMA control chart as a method to accurately and timely signal recurrence of depression. We use data from the TRANS-ID Tapering study including formerly depressed patients in remission who gradually discontinued their antidepressant medication (Smit et al., 2020). This particular sample is studied, as many patients who taper antidepressants will experience a depressive symptom recurrence (Borges et al., 2014). A unique feature of the TRANS-ID Tapering study is that participants provided 4 months of EMA data and that half of the participants experienced a recurrence of depressive symptoms during or shortly after these 4 months. These features allowed us to examine how often and how early significant increases in different affective and cognitive states occur before recurrence of depression, as compared to patients who stayed in remission. We examine change in high arousal negative affect (NA), low arousal NA, high arousal positive affect (PA), low arousal PA, and repetitive negative thinking, as these affective and cognitive states include commonly reported prodromal symptoms. Second, we will examine whether depression can be signaled accurately when monitoring whether one, two, or three mental states go out-of-control, independent of the type of mental state. Third, we will study the influence of the EWMA tuning parameters on the results. Fourth, we will explore whether burden for participants can be reduced by investigating the impact of using single items and using one or three instead of five assessments a day on the results. This is an exploratory study with no specific hypotheses on which mental states, items, tuning parameters, or number of assessments, would signal recurrence of depression best.

Method

Procedures and Participants

The current study was performed using the Transitions In Depression (TRANS-ID) Tapering dataset. A detailed study protocol is available on OSF (https://osf.io/zbwkp/). The study was advertised between June 2017 and October 2018 in newspapers, through social media and at a Dutch Pharmacy (de Regenboog Apotheek) that offers tapering strips that allow gradual tapering of antidepressant medication (Groot & van Os, 2018).[1] Potential participants could indicate their interest in study participation on a website. Thereafter, we screened patients for eligibility during a screening phone call and a baseline interview.

The study population were formerly depressed patients in remission who made a shared decision with their mental health care provider to discontinue antidepressant medication. The recommendation in the Netherlands is to continue antidepressants for a period of minimally 6–12 months after remission of depression before considering antidepressant tapering (Vinkers et al., 2021). Inclusion criteria were fulfilling the Diagnostic and Statistical Manual of Mental Disorders, 4th edition criteria of a past depressive episode, assessed with a Dutch brief version of the Schedules for Clinical Assessment in Neuropsychiatry (mini-SCAN, Nienhuis et al., 2010) by a trained interviewer, age ≥18, currently using an antidepressant, and a shared decision with a medical health professional to discontinue antidepressant medication. Patients were excluded when they fulfilled the criteria for a current depressive episode, a bipolar disorder or a psychotic disorder (assessed with the mini-SCAN), when they reported having a personality disorder, or when they were planning

[1]Tapering strips consist of daily pouches with antidepressant medication. Each pouch contains a slightly lower dose than the pouch of the day before (Groot & van Os, 2018).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

to start another type of antidepressant medication. Of the 69 participants who enrolled in the TRANS-ID Tapering study, 28 participants were excluded from the analyses in the current study for the following reasons: ≤1 month of at least 80% EMA data ($n = 13$), presence, absence, or timing of recurrence of depressive symptoms was unclear ($n = 5$), ≤6 weeks of EMA data before recurrence ($n = 10$). The final sample consisted of 22 participants with recurrence of depressive symptoms and 19 participants who stayed in remission ($n = 41$).

All participants provided written informed consent. The Medical Ethical Committee of the University Medical Center Groningen reviewed and approved the study procedures. All study procedures complied with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Participants performed ambulatory assessments during the last month of antidepressant tapering and the 3 months following antidepressant discontinuation. Ambulatory assessments consisted of smartphone-based Ecological Momentary Assessment (EMA) at five fixed time points a day (3 hr apart), continuous actigraphy, and heart rate measurements twice a day. Participants completed weekly questionnaires, were called approximately every 3 weeks, and participated in a semi-structured evaluation interview after 4 and 6 months. The EMA questionnaires were administered in Dutch (see Smit et al., 2020). We compensated participants financially if they completed at least 80% of the questionnaires. The 41 included participants completed a median of 540 EMA questionnaires (range 220–609). Median compliance was 87.6%, corresponding to a median of 4.4 out of 5 observations per day.

Mental States

We used EMA to assess high arousal high arousal negative affect (NA high), low arousal negative affect (NA low), high arousal positive affect (PA high), low arousal positive affect (PA low), and repetitive negative thinking. The choice for the EMA items and the grouping of items was based on the circumplex model of affect (Barrett & Russell, 1998), that distinguishes different emotional experiences using a valence (pleasantness) and activation (arousal) dimension of affect. High arousal NA was measured with the average of the items "I feel restless," "I feel stressed" and "I feel irritated." Low arousal NA was measured with the average of the items "I feel down," "I feel listless" and "I feel tired." All the affect items are answered on a VAS scale ranging from 0 (*not at all*) to 100 (*very much*). High arousal PA was measured as the average of the EMA items "I feel cheerful" and "I feel energetic." Low arousal PA was assessed by computing the average of the EMA items "I feel content" and "I feel at ease." Repetitive negative thinking was measured by averaging the two items "I am worrying," rated on a VAS scale ranging from 0 (*not at all*) to 100 (*very much*), and "my thoughts about myself are currently…" (rated on VAS scale ranging from 0 (*very negative*) to 100 (*very positive*). These two items reflect two processes encompassed by the concept "repetitive negative thinking," namely worry and habitual negative self-referential thinking (Mahoney et al., 2012). The median individual-level Cronbach's α across all individuals was .74 for high arousal NA, .64 for low arousal NA, .70 for high arousal PA, .67 for low arousal PA, and .52 for repetitive negative thinking. Measurement of repetitive negative thinking with only two items may explain the lower Cronbach's α.

Recurrence of Depressive Symptoms

Presence or absence of recurrence of depressive symptoms was determined based on quantitative criteria and a qualitative criterion (Smit et al., 2022). The quantitative criteria were: having a reliable depressive symptom increase and a persistent depressive symptom increase based on the weekly assessment of the Symptom Checklist-90 (SCL-90) depression subscale (Derogatis & Cleary, 1977). The Reliable Change Index (RCI) (N. S. Jacobson & Truax, 1991) was used to test whether there was reliable increase in depressive symptoms (≥8.5 points) on the SCL-90 depression subscale compared to the first 2 weeks of the study. The persistence criterion was fulfilled when this reliable change (≥8.5 points) persisted for at least three consecutive weeks and either treatment was started or increased or tapering was interrupted. The qualitative criterion was a consensus rating of three independent raters, one psychiatrist and two psychologists, of presence of a clinically meaningful increase in depressive symptoms based on open questions of the weekly questionnaires and the post-study evaluation interviews. The moment of onset of recurrence was defined as the first week with a reliable increase in scores on the SCL-90 depression subscale. If there was more than one significant reliable increase, we used the week that was closest to the date on which the participant first experienced a meaningful increase in depressive symptoms. For the individuals included in the current study, recurrence of depressive symptoms took place at least 6 weeks after start of the study. For six participants, recurrence took place after the 4-month EMA period (at most 6 weeks after the EMA period).

Statistical Analyses

Pre-Processing

To reduce autocorrelation in the EMA data, a day average (using max. five observations a day) was computed for each mental state (Schat et al., 2021). After that, day averages >3 *SD*s from the mean of the first 28 days (phase I) were Winsorized, with the intent to reduce the influence of errors in answering the EMA questions.

Exponentially Weighted Moving Average (EWMA) Control Charts

With EWMA charts, we examined whether and when significant within-person rises in each of the affective and cognitive states were detected in each of the participants. The EWMA control charts are graphs with time on the *x*-axis and the EWMA scores on the *y*-axis. These EWMA scores are a weighted average of the day average of the mental state up to a certain time point, where recent observations receive more weight. The EWMA of each mental state of each day average was computed as follows:

$$\mathrm{EWMA}_t = \lambda Y_t + (1 - \lambda)\mathrm{EWMA}_{t-1}$$

Y_t represents the day average of a mental state at time t, and lambda (λ) determines how much weight is given to the current observation. The higher λ, the more weight is assigned to recent observations. As the default, we chose a lambda of .1 in line with our earlier EWMA

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

analyses of simulated and empirical EMA data on mental states (Schat et al., 2021; Smit & Snippe, 2022).

The control chart also contains an Upper Control Limit (UCL) and a Lower Control Limit (LCL), based on the person-specific mean μ_1 and standard deviation σ_1 of the EWMA scores during the baseline period (i.e., phase I). We used the first 4 weeks (28 days) as phase I as this is the period before antidepressants were fully discontinued for most patients. During a subsequent period (phase II), we continuously examined whether the monitored EWMA exceed the UCL or the LCL control limits and could be flagged out-of-control. We defined out-of-control as low arousal NA, high arousal NA or repetitive negative thinking exceeding the UCL and low arousal PA or high arousal PA exceeding the LCL.

The UCL and LCL were computed as follows:

$$\mathrm{UCL} = \mu_1 + L\sigma_1\sqrt{\frac{\lambda}{(2-\lambda)}[1-(1-\lambda)^{2t}]} \quad \text{and}$$

$$\mathrm{LCL} = \mu_1 - L\sigma_1\sqrt{\frac{\lambda}{(2-\lambda)}[1-(1-\lambda)^{2t}]}$$

where L determines the width of the control limits (Montgomery, 2012). The appropriate value of L depends on λ and on the desired in-control Average Run Length (ARL_0). The ARL_0 is the expected number of days before a false positive occurs (i.e., an EWMA score is flagged as out-of-control when the process is in fact in-control). As a default parameter choice, we used an ARL_0 of 370 days, as this is the most commonly used ARL_0 (Montgomery, 2012), which corresponds with an L of 2.62 for a λ of .10 (when testing one-sided).

Analytical Approach for Each of the Four Research Questions

To answer the four research questions outlined in the introduction, we conducted the following analyses. First, we examined how often each of the five mental states (high arousal PA, low arousal PA, high arousal NA, low arousal NA, repetitive negative thinking) were flagged as out-of-control based on EWMA charts using the default parameter choices ($\lambda = .10$, $ARL_0 = 370$). We summarize these results using the performance measures reported in the next section. We also investigated how many days before the symptom recurrence the best performing mental states were flagged as out-of-control.

Second, we examined whether recurrence could be signaled more accurately by monitoring multiple mental states at the same time by providing an alert when at least one, two, or three of any of the mental states were flagged as out-of-control, independent of the type of mental state (using the default tuning choices).

Third, we examined the influence of the tuning parameters lambda (λ) and the in-control Average Run Length (ARL_0) on signaling recurrence by running the EWMA analyses of the five affective and cognitive states with four different values of λ and four different values of ARL_0 (sixteen different tuning choices). We used relatively small values of λ (i.e., .01, .05, .10, .15), as we were interested in detecting small changes in mental states (Montgomery, 2012). We investigated the following values of ARL_0: 185, 370, 740, and 1,480. We used the R package spc (Knoth, 2020) to determine L for different values of λ and ARL_0 (tested one-sided, see Supplemental Table S1).

Fourth, we explored whether participant burden could be reduced by using a single item or by reducing the number of assessments a day. We investigated which single items of the best performing affective and cognitive states signaled recurrence of depression best by re-running the EWMA charts on these single items, using the default parameter choices ($\lambda = .10$ and $ARL_0 = 370$), and using the sixteen different tuning choices as sensitivity analyses (see Supplemental). To examine the influence of the number of assessments a day, the EWMA charts were re-run with only three observations per day (omitting the second and fourth observation within each day), or with only the morning observation, or with only the evening observation. Analyses were run using the default parameter choices and using the sixteen different tuning choices (see Supplemental).

Performance Measures to Quantify Accuracy in Signaling Recurrence of Depression

We evaluated performance of the EWMA charts in signaling recurrence of depression primarily computing the following metrics across all participants: sensitivity, specificity, and the Matthew's correlation coefficient (MCC). Sensitivity is the number true positives (TP)/the number of participants with a recurrence, where the number TP is the number of participants for which the mental state was flagged as out-of-control before recurrence of depression happened. Specificity is the number true negatives (TN)/the number of participants without a recurrence, where the number TN is the number of participants for which the mental state was not flagged out of control in participants without a recurrence. The MCC is a correlation-type measure of accuracy. It quantifies the correlation between the predicted status (i.e., alert of a recurrence or not, based on whether a mental state goes out-of-control or not) and observed status of the participant (i.e., actual recurrence of depressive symptoms or not). $\mathrm{MCC} = (\mathrm{TP} * \mathrm{TN}) - (\mathrm{FP} * \mathrm{FN})/\sqrt{(\mathrm{TP}+\mathrm{FP}) * (\mathrm{TP}+\mathrm{FN}) * (\mathrm{TN}+\mathrm{FP}) * (\mathrm{TN}+\mathrm{FN})}$, where FP = false positives, and FN = false negatives. Values of the MCC range between −1 (the predicted status never matches the observed one) and +1 (the status of all participants was correctly predicted).

Secondarily, we examined for how many participants an alert based on EWMA charts was correct, by computing the Positive Predictive Value ($\mathrm{PPV} = \mathrm{TP}/\mathrm{TP}+\mathrm{FP}$). Similarly, we examined for how many participants an absence of an alert was correct, by means of the Negative Predictive Value ($\mathrm{NPV} = \mathrm{TN}/\mathrm{TN}+\mathrm{FN}$).

We also applied decision analysis to examine whether providing an alert of increased risk of recurrence based on EWMA charts is more beneficial in terms of net benefit than alerting all participants or none of them (Vickers & Elkin, 2006; Vickers et al., 2016). Details on these analyses as well as the results are included in the Supplement.

Transparency and Openness

A study protocol including the used study materials is available on OSF (see https://osf.io/zbwkp/). The analysis plan of the current study was preregistered (see https://osf.io/sbyq7/). The data are not freely available in a public repository due to restrictions related to the European Law regarding General Data Protection Regulation (GDPR). Collaboration on TRANS-ID Tapering data is possible and can be requested. The study analysis code will be made available upon request. Data were analyzed using R, version 4.2.0 (R Core Team, 2022). Analyses were performed using the ewma function

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

of the R package qcc (Scrucca, 2004). The R package spc (Knoth, 2020) was used to determine L for different values of λ and ARL_0. The R package ggplot2 was used for data visualization (Wickham, 2016).

Results

Sample Characteristics

The majority of the sample was female, highly educated, in a romantic relationship, and used tapering strips to taper antidepressant medication (see Table 1).

The sample seemed to be representative of the total sample as characteristics of participants who were included in the analyses did not differ from characteristics of dropouts and participants excluded from the analyses (see Table S2 in supplements). Only use of Venlafaxine was more common in the included sample (39.0%) than in the excluded sample (18.5%).

Which Mental States Signal Recurrence of Depressive Symptoms Best and How Early?

An upcoming recurrence of depressive symptoms could be signaled relatively accurately based on EWMA charts of each of the mental states (see Table 2). Significant increases in repetitive negative thinking and high arousal NA most accurately signaled recurrence according to the MCC (see Table 2). An increase in repetitive negative thinking was detected before recurrence of depressive symptoms in 18 of the 22 participants (82%) with recurrence of and in 8 of the 19 participants (42%) who remained in remission, implying high sensitivity and moderate specificity (see Table 2). Alerts based on significant increases in repetitive negative thinking would be correct in 69% of the cases (PPV) and 27% of the cases who would not receive an alert did in fact experience a recurrence (1-NPV, see Table 2). An increase in high arousal NA was detected before recurrence of depressive symptoms in 10 of the 22 participants (45%) with recurrence and in 2 of the 19 participants (11%) who remained in remission, implying high specificity and moderate sensitivity. Alerts based on increases in high arousal NA would be correct in 83% of the cases and 41% of participants who would not receive an alert did experience a recurrence.

An upcoming recurrence could be signaled between 1 and 150 days in advance (see Supplemental Figures S1 and S2). An increase in repetitive negative thinking (the most sensitive mental state) was detected at least 1 week before recurrence in 94% of the cases and at least 1 month before recurrence in 72% of the cases, using the default parameter choices. An increase in high arousal NA (the most specific mental state) was detected at least 1 week before recurrence in 90% of the cases and at least 1 month before recurrence in 80% of the cases. The detection times were comparable for the other mental states.

Can Recurrence Be Signaled Equally or More Accurately by Monitoring Whether One, Two, or Three Mental States Go Out-of-Control?

Compared to repetitive negative thinking and high arousal NA, the accuracy could not be improved by providing an alert of an upcoming recurrence when at least one (MCC = .36) or two (MCC = .31) or 3 (MCC = .36) mental states went out-of-control, using the default tuning choices. Providing an alert when at least one of the mental states went out-of-control resulted in high sensitivity (95%) at the cost of low specificity (32%). Providing an alert when at least two or three of the mental states went out-of-control

Table 1
Sample Characteristics at Baseline

Baseline characteristics	Total sample ($n = 41$) % (n) or M (SD)
Age	46.5 (13.0)
Female	83% (34)
Education	
Primary school or lower vocational school	0% (0)
Secondary school or advanced vocational school	12% (5)
Higher education or University	88% (36)
Hours of paid work per week	18.4 (15.0)
In a romantic relationship	81% (33)
Type of antidepressant medication	
Venlafaxine	39% (16)
Paroxetine	15% (6)
Sertraline	10% (4)
Fluoxetine	10% (4)
Citalopram	17% (7)
Escitalopram	5% (2)
Clomipramine	2% (1)
Bupropion	2% (1)
Primary reason tapering antidepressants	
My depressive mood has improved (enough) by now	38% (13)[a]
I experience side effects of antidepressants	29% (10)[a]
I want to rely on my own strength (not on antidepressants)	32% (11)[a]
Used tapering strips	81% (33)
Receives psychological treatment	35% (12)[a]
Tried tapering before	65% (22)[a]
N discontinuation attempts	1.2 (1.3)[a]
Years since starting current antidepressant	7.0 (6.8)[a]
Baseline SCL-90 depression	28.6 (9.2)

Note. [a] Based on $n = 34$ out of 41 as a subset of participants did not fill out all questionnaires. N = number, SCL-90 depression = Symptom Checklist-90 depression subscale.

Table 2
Accuracy of EWMA Charts in Signaling Recurrence of Depression

Mental states	MCC	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Composite mental states					
NA high	.38	45	89	83	59
NA low	.20	45	74	67	54
PA high	.22	59	63	65	57
PA low	.31	73	58	67	65
Rep. neg. think.	.41	82	58	69	73
Items of NA high					
Restless	.33	59	74	72	61
Irritated	.34	55	79	75	60
Stressed	.10	41	69	60	50
Items of Rep. neg. think.					
Worry	.36	73	63	70	67
Neg. thoughts self	.41	82	58	69	73

Note. EWMA = Exponentially Weighted Moving Average, NA high = high arousal negative affect, NA low = low arousal negative affect, PA high = high arousal Positive Affect, PA low = low arousal positive affect, Rep. neg. think. = Repetitive negative Thinking, Neg. thoughts self = negative thoughts about the self, MCC = Matthew's Correlation Coefficient.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

resulted in a more adequate balance between sensitivity (77% and 73% respectively) and specificity (53% and 63% respectively).

What is the Influence of the Tuning Parameters of the EWMA Charts on the Results?

Other tuning choices (i.e., values of λ and ARL_0) did not lead to major changes in the results for the best performing states, only one to three more patients were correctly classified (see Figure 2). Other tuning choices did lead to major improvements in the performance of high arousal PA; five more patients were correctly classified, the sensitivity increased to 86% and the MCC to .47 when using an λ of .01 instead of .10.

Overall, lower ARL_0 values slightly increased sensitivity and higher ARL_0 values slightly increased specificity. The impact of λ depended on the valence of the affective and cognitive states: depressive symptoms were signaled most accurately using the lowest values of λ (.01) for high arousal PA and low arousal PA, and using higher values of λ (.10 or .15) for repetitive negative thinking, high arousal NA, and low arousal NA (see MCC in Figure 2).

What is the Influence of Using Single Items and Using Fewer Assessments on the Results?

Using the default parameter choices, none of the single items performed better than the combined mental states (see Table 1 and Supplemental Figures S4 and S5). Of the items of the best performing mental states, "negative thoughts about the self" had exactly the same accuracy, sensitivity, and specificity as repetitive negative thinking. Significant increases in the items "restless," "irritated," and "worry" signaled recurrence of depression comparably well to the best performing mental states, although the balance in sensitivity and specificity was different which resulted in either lower specificity or sensitivity than the best performing mental states.

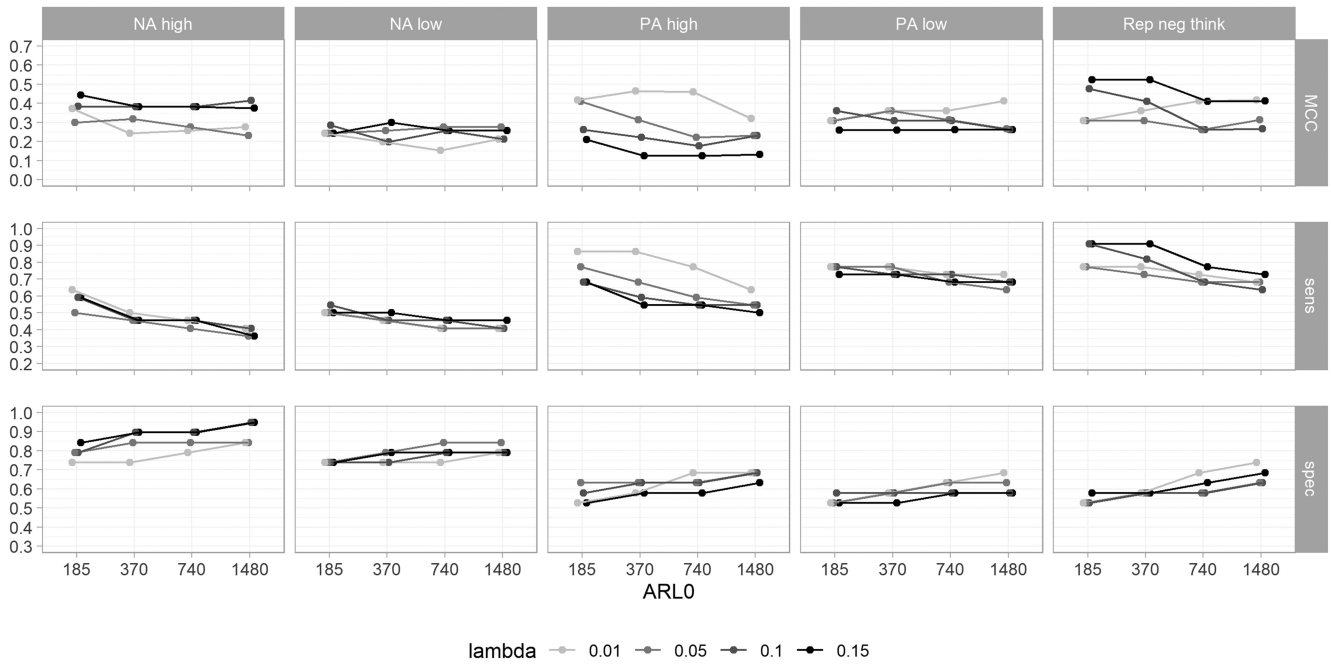
The number of observations a day did not strongly affect the sensitivity and the MCC of repetitive negative thinking, using the default parameter choices. However, across tuning choices, the sensitivity and the MCC was highest when using five observations per day (see Supplemental Figures S6, S7, & S8). The specificity and the MCC of high arousal NA did decrease linearly with decreasing number of observations a day, both using the default parameter choices (see Figure 3) and across tuning choices (see Supplemental Figures S6, S7, & S8).

Discussion

Speculations about the potential of Ecological Momentary Assessment (EMA) for predicting future psychopathology have been around for several years (e.g., Andreescu et al., 2019; Dogan et al., 2017; Wichers, 2014). This study lives up to that promise by showing that EWMA control charts applied to EMA data of affective and cognitive states can aid personalized early detection of recurrence of depression in real-time. Mean changes in EMA data of cognitive and affective states were detected with EWMA charts more often before recurrence of depression than in patients who stayed in remission. Recurrences were signaled based on EWMA charts more than a month in advance in over two thirds of the participants. Thus, an upcoming recurrence of depression could be

Figure 2

Predictive Performance of the EWMA Charts Across Mental States and Tuning Choices

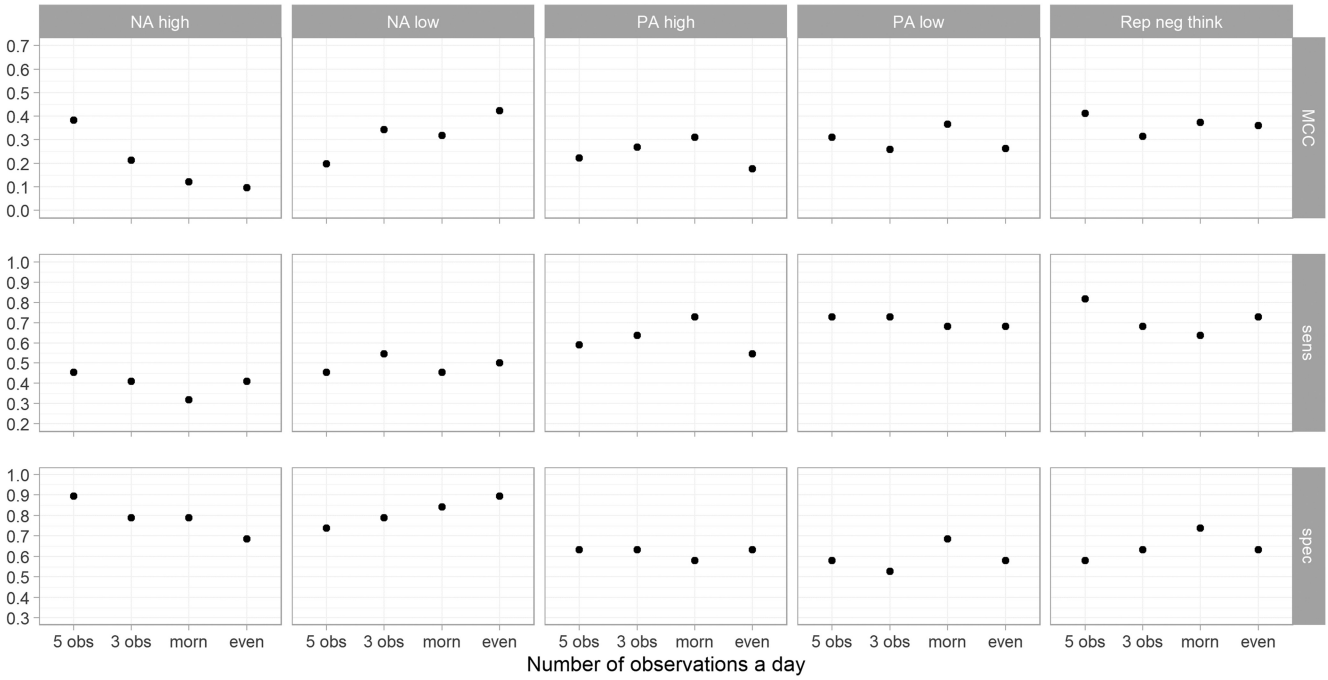


Note. EWMA = Exponentially Weighted Moving Average, ARL0 = in-control Average Run Length, MCC = Matthew's correlation coefficient, sens = sensitivity, spec = specificity, Negative Affect (NA) high = high arousal NA, NA low = low arousal NA, Positive Affect (PA) high = high arousal PA, PA low = low arousal PA, Rep neg think = Repetitive negative thinking.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Figure 3

Predictive Performance of the EWMA Charts Across Mental States and Number of Observations a Day



Note. EWMA = Exponentially Weighted Moving Average, MCC = Matthew's correlation coefficient, sens = sensitivity, spec = specificity, Negative Affect (NA) high = high arousal NA, NA low = low arousal NA, Positive Affect (PA) high = high arousal PA, PA low = low arousal PA, Rep neg think = Repetitive negative thinking, 5 obs = 5 observations a day, 3 obs = 3 observations a day, morn = morning observation, even = evening observation.

detected earlier using EWMA charts applied to EMA data than using weekly symptom questionnaires and patient reports of symptom increases. The accuracy of foreseeing recurrence based on the EWMA control chart was stable across different tuning choices (i.e., different values of ARL_0 and λ).

The results showed that significant increases in repetitive negative thinking (worrying and negative thoughts about the self) and high arousal NA (feeling restless, feeling irritated, and feeling stressed) were the most accurate early signs of recurrence of depression. Most of the recurrences (82%) could be detected in advance based on EWMA charts of repetitive negative thinking. Yet, providing alerts based on increases in repetitive negative thinking would also result in 31% false positives. When aiming to be surer of risk of recurrence of depressive symptoms, providing alerts based on increases in high arousal NA would be most suitable since this would result in few false positives (17%). However, this would come at the cost of detecting only about half of the recurrences.

Our results suggest that significant increases in repetitive negative thinking are very common in the prodromal phase of depression and that significant increases in high arousal NA occur almost exclusively in patients with a recurrence of depression. Whereas previous research showed that patients who have higher levels of anxiety or repetitive negative thinking than patients with lower levels are more vulnerable for developing depression in the future (N. C. Jacobson & Newman, 2017; Spinhoven et al., 2018), we show that significant within-person increases in symptoms of anxiety (i.e., high arousal NA) and repetitive negative thinking are common early signs of an upcoming depression, suggesting that ergodicity holds in this case. Our findings align with previous research showing that symptoms of anxiety or hyper arousal are among the most commonly reported prodromal symptoms of depression (Benasi et al., 2021).

Although increases in repetitive negative thinking and high arousal NA were not detected in all participants, indicating individual differences in the presence of these prodromal symptoms of depression, an increase in repetitive negative thinking and high arousal were detected more often in patients with a recurrence than in patients without a recurrence. Moreover, recurrences could not be signaled more accurately by providing an alert when at least one, two, or three mental states went out-of-control, independent of the type of mental state.

When researchers or clinicians decide to reduce participant burden by lowering the number of observations per day, it is likely that the accuracy of the EWMA charts in signaling recurrence of depression will be slightly lower. Recurrence of depression could be signaled better using five assessments a day than using three or one assessment a day across tuning choices, which is in line with the findings of a recent simulation study (Schat et al., 2021). One could consider reducing participant burden by only using the single item reflecting "negative thoughts about the self," as this item had the same accuracy in signaling an upcoming recurrence as the composite measure repetitive negative thinking. Furthermore, the accuracy of "restless," "irritated," "worry" was comparable to the accuracy of the composite mental states, but both the sensitivity and specificity of the indicators became less extreme. Thus, the choice for monitoring repetitive negative thinking, high arousal

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

NA or specific single items may depend on whether patients and clinicians aim for higher sensitivity, higher specificity, or a better balance between sensitivity and specificity.

The results of the current study are promising when compared to previous studies aiming to predict recurrence of depression. The sensitivity and specificity found in the current study were comparable or slightly higher than found in studies determining general risk of recurrence based on between-subject data (Judd et al., 2016; Klein et al., 2018; Lorimer et al., 2021; van Loo et al., 2020). Accuracy and sensitivity of the EWMA charts were higher than the best early warning signals of recurrence found in other studies using the TRANS-ID Tapering data. These studies examined within-person rises in complexity measures (e.g., autocorrelation and variance) in PA and NA (Smit et al., under review), actigraphy data (Kunkels et al., in press), and heart rate data (George et al., under review). Moreover, these and other previous studies examining early warning signals of depression (Cabrieto et al., 2019; Wichers et al., 2016, 2020), used all the data up till the moment of recurrence, which makes these methods not suitable for use in clinical practice.

EWMA charts are suitable for use in clinical practice as they can be used to detect early signs of recurrence prospectively in real-time, after each EMA observation. Patients have voiced a desire for a tool that can help them notice when symptoms start to return (Bos et al., 2019), as they often fail to recognize this (Epstein et al., 2010; Sherwood et al., 2007). Monitoring change in repetitive negative thinking and high arousal NA using EWMA control charts may form the basis of such a depression barometer, which may be especially valuable for patients who are tapering their antidepressants. EWMA charts may have empowering effects on patients, may help patients understand and manage their own health, and may aid shared decision making (Thor et al., 2007). Using EWMA charts, patients could be provided insight in the type of symptoms that are returning, which may provide additional clues on what type of intervention is needed. The fact that depression was foreseen at least a month before recurrence in most participants allows for timely intervention. The EWMA method comes with the advantage that patients can be alerted that their risk of depression has increased over time. This may help patients and clinicians to decide if action is needed at this very moment.

Future research needs to provide insight in how many patients are willing to engage in intensive long-term monitoring. A considerable proportion of the eligible participants declined to participate because of the expected burden of the ambulatory assessments. Yet, study participation was feasible for those who participated as evidenced by high compliance (88%) to EMA and relatively low dropout rate (19%). This indicates that long-term monitoring is feasible for most patients who are motivated to do so, but also that not all patients are willing to try long-term monitoring.

A limitation of the current study is the number of monitored participants. Although the sample size was quite large for a replicated single-subject study, it was still modest and consequently the precision of the estimates of the accuracy of the EWMA charts is limited. Second, the results of the current study may only generalize to patients who taper their antidepressants and have similar characteristics as the current sample. Another limitation of the study is that we measured some of the mental states with only two items to limit participant burden. Therefore, not all aspects of repetitive negative thinking, such as repetitive thinking about past events and uncontrollability of thoughts, were measured (Ehring & Watkins, 2008). The median within-person internal consistency of our measure of repetitive negative thinking was also relatively low (0.52). Though this may not have been very problematic as increases in the average of the two items measuring repetitive negative thinking were an accurate early sign of recurrence, measuring repetitive negative thinking using more items may further improve the reliability of this signal. Furthermore, several common prodromal symptoms, such as sleep problems, somatic complaints, and feeling anxious (Benasi et al., 2021), were not included in the present study. Finally, the presented EWMA method may only be clinically valuable for those patients who are willing to engage in self-monitoring using intensive EMA over a prolonged period.

In conclusion, the present study suggests that early detection of psychopathology is possible by monitoring affective and cognitive states intensively and providing alerts based on EWMA control charts. Alerting patients of early signs of recurrence may help to intervene in time before depression fully manifests.

References

Andreescu, C., Ajilore, O., Aizenstein, H. J., Albert, K., Butters, M. A., Landman, B. A., Karim, H. T., Krafty, R., & Taylor, W. D. (2019). Disruption of neural homeostasis as a model of relapse and recurrence in late-life depression. *The American Journal of Geriatric Psychiatry*, *27*(12), 1316–1330. https://doi.org/10.1016/j.jagp.2019.07.016

Baldewijns, G., Luca, S., Vanrumste, B., & Croonenborghs, T. (2016). Developing a system that can automatically detect health changes using transfer times of older adults. *BMC Medical Research Methodology*, *16*(1), Article 23. https://doi.org/10.1186/s12874-016-0124-4

Barlow, D. H., & Nock, M. K. (2009). Why can't we be more idiographic in our research? *Perspectives on Psychological Science*, *4*(1), 19–21. https://doi.org/10.1111/j.1745-6924.2009.01088.x

Barrett, L. F., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*(4), 967–984. https://doi.org/10.1037/0022-3514.74.4.967

Bastiaansen, J. A., Kunkels, Y. K., Blaauw, F. J., Boker, S. M., Ceulemans, E., Chen, M., Chow, S. M., de Jonge, P., Emerencia, A. C., Epskamp, S., Fisher, A. J., Hamaker, E. L., Kuppens, P., Lutz, W., Meyer, M. J., Moulder, R., Oravecz, Z., Riese, H., Rubel, J., … Bringmann, L. F. (2020). Time to get personal? The impact of researchers choices on the selection of treatment targets using the experience sampling methodology. *Journal of Psychosomatic Research*, *137*, Article 110211. https://doi.org/10.1016/j.jpsychores.2020.110211

Benasi, G., Fava, G. A., & Guidi, J. (2021). Prodromal symptoms in depression: A systematic review. *Psychotherapy and Psychosomatics*, *90*(6), 365–372. https://doi.org/10.1159/000517953

Berwian, I. M., Walter, H., Seifritz, E., & Huys, Q. J. M. (2017). Predicting relapse after antidepressant withdrawal—A systematic review. *Psychological Medicine*, *47*(3), 426–437. https://doi.org/10.1017/S0033291716002580

Borges, S., Chen, Y. F., Laughren, T. P., Temple, R., Patel, H. D., David, P. A., Mathis, M., Unger, E., Yang, P., & Khin, N. A. (2014). Review of maintenance trials for major depressive disorder: A 25-year perspective from the US Food and Drug Administration. *Journal of Clinical Psychiatry*, *75*(3), 205–214. https://doi.org/10.4088/JCP.13r08722

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

Bos, F. M., Snippe, E., Bruggeman, R., Wichers, M., & van der Krieke, L. (2019). Insights of patients and clinicians on the promise of the experience sampling method for psychiatric care. *Psychiatric Services*, *70*(11), 983–991. https://doi.org/10.1176/appi.ps.201900050

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Buckman, J. E. J., Underwood, A., Clarke, K., Saunders, R., Hollon, S. D., Fearon, P., & Pilling, S. (2018). Risk factors for relapse and recurrence of depression in adults and how they operate: A four-phase systematic review and meta-synthesis. *Clinical Psychology Review*, *64*, 13–38. https://doi.org/10.1016/j.cpr.2018.07.005

Cabrieto, J., Adolf, J., Tuerlinckx, F., Kuppens, P., & Ceulemans, E. (2019). An objective, comprehensive and flexible statistical framework for detecting early warning signs of mental health problems. *Psychotherapy and Psychosomatics*, *88*(3), 184–186. https://doi.org/10.1159/000494356

Derogatis, L. R., & Cleary, P. A. (1977). Confirmation of the dimensional structure of the SCL-90: A study in construct validation. *Journal of Clinical Psychology*, *33*(4), 981–989. https://doi.org/10.1002/1097-4679(197710)33:4 < 981::AID-JCLP2270330412 > 3.0.CO;2-0

Dogan, E., Sander, C., Wagner, X., Hegerl, U., & Kohls, E. (2017). Smartphone-based monitoring of objective and subjective data in affective disorders: Where are we and where are we going? Systematic review. *Journal of Medical Internet Research*, *19*(7), Article e262. https://doi.org/10.2196/jmir.7006

Ehring, T., & Watkins, E. R. (2008). Repetitive negative thinking as a transdiagnostic process. *International Journal of Cognitive Therapy*, *1*(3), 192–205. https://doi.org/10.1521/ijct.2008.1.3.192

Epstein, R. M., Duberstein, P. R., Feldman, M. D., Rochlen, A. B., Bell, R. A., Kravitz, R. L., Cipri, C., Becker, J. D., Bamonti, P. M., & Paterniti, D. A. (2010). "I didn't know what was wrong:" How people with undiagnosed depression recognize, name and explain their distress. *Journal of General Internal Medicine*, *25*(9), 954–961. https://doi.org/10.1007/s11606-010-1367-0

Fava, G. A., & Tossani, E. (2007). Prodromal stage of major depression. *Early Intervention in Psychiatry*, *1*(1), 9–18. https://doi.org/10.1111/j.1751-7893.2007.00005.x

George, S. V., Kunkels, Y. K., Smit, A. C., Wichers, M., Snippe, E., van Roon, A. M., & Riese, H. (under review). Using complexity of cardiac dynamics as a predictor of recurrence of depression in individuals tapering their antidepressants use. https://doi.org/10.31234/osf.io/gk8rd

Geschwind, N., Peeters, F., Jacobs, N., Delespaul, P., Derom, C., Thiery, E., van Os, J., & Wichers, M. (2010). Meeting risk with resilience: High daily life reward experience preserves mental health. *Acta Psychiatrica Scandinavica*, *122*(2), 129–138. https://doi.org/10.1111/j.1600-0447.2009.01525.x

Groot, P. C., & van Os, J. (2018). Antidepressant tapering strips to help people come off medication more safely. *Psychosis-Psychological Social and Integrative Approaches*, *10*(2), 142–145. https://doi.org/10.1080/17522439.2018.1469163

Hamaker, E. L. (2012). Why researchers should think 'within-person': A pragmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 43–61). Guilford Press.

Hettema, J. M., Kuhn, K. W., Prescott, C. A., & Kendler, K. S. (2006). The impact of generalized anxiety disorder and stressful life events on risk for major depressive episodes. *Psychological Medicine*, *36*(6), 789–795. https://doi.org/10.1017/S0033291706007367

Ilies, I., Anderson, D. J., Salem, J., Baker, A. W., Jacobsen, M., & Benneyan, J. C. (2020). Large-scale empirical optimisation of statistical control charts to detect clinically relevant increases in surgical site infection rates. *BMJ Quality & Safety*, *29*(6), 472–481. https://doi.org/10.1136/bmjqs-2018-008976

Jacobson, N. C., & Newman, M. G. (2017). Anxiety and depression as bidirectional risk factors for one another: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *143*(11), 1155–1200. https://doi.org/10.1037/bul0000111

Jacobson, N. S., & Truax, P. (1991). Clinical-significance—A statistical approach to defining meaningful change in psychotherapy-research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Judd, L. L., Schettler, P. J., & Rush, A. J. (2016). A brief clinical tool to estimate individual patients' risk of depressive relapse following remission: Proof of concept. *American Journal of Psychiatry*, *173*(11), 1140–1146. https://doi.org/10.1176/appi.ajp.2016.15111462

Klein, N. S., Holtman, G. A., Bockting, C. L. H., Heymans, M. W., & Burger, H. (2018). Development and validation of a clinical prediction tool to estimate the individual risk of depressive relapse or recurrence in individuals with recurrent depression. *Journal of Psychiatric Research*, *104*, 1–7. https://doi.org/10.1016/j.jpsychires.2018.06.006

Knoth, S. (2020). *Statistical process control—Calculation of ARL and other control chart performance measures*. https://cran.r-project.org/package=spc

Kunkels, Y. K., Smit, A. C., Minaeva, O., Snippe, E., George, S. V., van Roon, A. M., & Riese, H. (in press). Risk Ahead: Actigraphy-based early-warning signals of increases in depressive symptoms during antidepressant discontinuation. *Clinical Psychological Science*.

Lorimer, B., Delgadillo, J., Kellett, S., & Lawrence, J. (2021). Dynamic prediction and identification of cases at risk of relapse following completion of low-intensity cognitive behavioural therapy. *Psychotherapy Research*, *31*(1), 19–32. https://doi.org/10.1080/10503307.2020.1733127

Mahoney, A. E. J., McEvoy, P. M., & Moulds, M. L. (2012). Psychometric properties of the Repetitive Thinking Questionnaire in a clinical sample. *Journal of Anxiety Disorders*, *26*(2), 359–367. https://doi.org/10.1016/j.janxdis.2011.12.003

Minaeva, O., George, S. V., Kuranova, A., Jacobs, N., Thiery, E., Derom, C., Wichers, M., Riese, H., & Booij, S. H. (2021). Overnight affective dynamics and sleep characteristics as predictors of depression and its development in women. *Sleep*, *44*(10), Article zsab129. https://doi.org/10.1093/sleep/zsab129

Montgomery, D. C. (2012). *Statistical quality control: A modern introduction* (7th ed.). Wiley.

Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & van Os, J. (2009). Experience sampling research in psychopathology: Opening the black box of daily life. *Psychological Medicine*, *39*(9), 1533–1547. https://doi.org/10.1017/S0033291708004947

Neubauer, A. B., Brose, A., & Schmiedek, F. (2022). How within-person effects shape between-person differences: A multilevel structural equation modeling perspective. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000481

Nienhuis, F. J., van de Willige, G., Rijnders, C. A., de Jonge, P., & Wiersma, D. (2010). Validity of a short clinical interview for psychiatric diagnosis: The mini-SCAN. *British Journal of Psychiatry*, *196*(1), 64–68. https://doi.org/10.1192/bjp.bp.109.066563

Palmier-Claus, J. E., Myin-Germeys, I., Barkus, E., Bentley, L., Udachina, A., Delespaul, P. A., Lewis, S. W., & Dunn, G. (2011). Experience sampling research in individuals with mental illness: Reflections and guidance. *Acta Psychiatrica Scandinavica*, *123*(1), 12–20. https://doi.org/10.1111/j.1600-0447.2010.01596.x

Piccirillo, M. L., & Rodebaugh, T. L. (2019). Foundations of idiographic methods in psychology and applications for psychotherapy. *Clinical Psychology Review*, *71*, 90–100. https://doi.org/10.1016/j.cpr.2019.01.002

Raes, F., Smets, J., Nelis, S., & Schoofs, H. (2012). Dampening of positive affect prospectively predicts depressive symptoms in non-clinical samples. *Cognition and Emotion*, *26*(1), 75–82. https://doi.org/10.1080/02699931.2011.555474

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Richards, D. (2011). Prevalence and clinical course of depression: A review. *Clinical Psychology Review*, *31*(7), 1117–1125. https://doi.org/10.1016/j.cpr.2011.07.004

Schat, E., Tuerlinckx, F., Smit, A. C., De Ketelaere, B., & Ceulemans, E. (2021). Detecting mean changes in experience sampling data in real-time: A comparison of univariate and multivariate statistical process control

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

methods. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000447

Scrucca, L. (2004). Qcc: An R package for quality control charting and statistical process control. *R News*, *4*(1), 11–17.

Sherwood, C., Salkovskis, P. M., & Rimes, K. A. (2007). Help-seeking for depression: The role of beliefs, attitudes and mood. *Behavioural and Cognitive Psychotherapy*, *35*(5), 541–554. https://doi.org/10.1017/S1352465807003815

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

Smit, A. C., Helmich, M. A., Bringmann, L. F., Oldehinkel, A. J., Wichers, M., & Snippe, E. (under review). Critical slowing down in momentary affect as early warning signal of impending transitions in depression. 10.31234/osf.io/u6vdg.

Smit, A. C., Helmich, M. A., Snippe, E., & Wichers, M. (2020). *Transitions in Depression (TRANS-ID) Tapering: Study protocol for a repeated intensive longitudinal n=1 study design to search for personalized early warning signals of critical transitions towards depression*. 2020(03-25-2021). https://osf.io/zbwkp/

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

Smit, A. C., Snippe, E., Bringmann, L. F., Hoenders, H. J. R., & Wichers, M. (2022). Transitions in depression: If, how, and when depressive symptoms return during and after discontinuing antidepressants. *Quality of Life Research*. https://doi.org/10.1007/s11136-022-03301-0

Smit, A. C., Snippe, E., & Wichers, M. (2019). Increasing restlessness signals impending increase in depressive symptoms more than 2 months before it happens in individual patients. *Psychotherapy and Psychosomatics*, *88*(4), 249–251. https://doi.org/10.1159/000500594

Snippe, E., Bos, E. H., van der Ploeg, K. M., Sanderman, R., Fleer, J., & Schroevers, M. J. (2015). Time-series analysis of daily changes in mindfulness, repetitive thinking, and depressive symptoms during mindfulness-based treatment. *Mindfulness*, *6*(5), 1053–1062. https://doi.org/10.1007/s12671-014-0354-7

Spinhoven, P., van Hemert, A. M., & Penninx, B. W. (2018). Repetitive negative thinking as a predictor of depression and anxiety: A longitudinal cohort study. *Journal of Affective Disorders*, *241*, 216–225. https://doi.org/10.1016/j.jad.2018.08.037

Thor, J., Lundberg, J., Ask, J., Olsson, J., Carli, C., Harenstam, K. P., & Brommels, M. (2007). Application of statistical process control in healthcare improvement: Systematic review. *Quality and Safety in Health Care*, *16*(5), 387–399. https://doi.org/10.1136/qshc.2006.022194

Vachon, H., Viechtbauer, W., Rintala, A., & Myin-Germeys, I. (2019). Compliance and retention with the experience sampling method over the continuum of severe mental disorders: Meta-analysis and recommendations. *Journal of Medical Internet Research*, *21*(12), Article e14475. https://doi.org/10.2196/14475

Vaessen, T., van Nierop, M., Decoster, J., Delespaul, P., Derom, C., de Hert, M., Jacobs, N., Menne-Lothmann, C., Rutten, B., Thiery, E., van Os, J., van Winkel, R., Wichers, M., & Myin-Germeys, I. (2017). Is sensitivity to daily stress predictive of onset or persistence of psychopathology? *European Psychiatry*, *45*, 167–173. https://doi.org/10.1016/j.eurpsy.2017.07.002

van Loo, H. M., Bigdeli, T. B., Milaneschi, Y., Aggen, S. H., & Kendler, K. S. (2020). Data mining algorithm predicts a range of adverse outcomes in major depression. *Journal of Affective Disorders*, *276*, 945–953. https://doi.org/10.1016/j.jad.2020.07.098

Vickers, A. J., & Elkin, E. B. (2006). Decision curve analysis: A novel method for evaluating prediction models. *Medical Decision Making*, *26*(6), 565–574. https://doi.org/10.1177/0272989X06295361

Vickers, A. J., Van Calster, B., & Steyerberg, E. W. (2016). Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *Bmj-British Medical Journal*, *352*, i6. https://doi.org/10.1136/bmj.i6

Vinkers, C. H., De Wit, L. E., Maarsingh, O. R., Ligthart, S. A., & Ruhe, H. G. (2021). Rational prescription of antidepressants for depression. *Nederlands Tijdschrift Voor Geneeskunde*, *165*, D5589.

Wichers, M. (2014). The dynamic nature of depression: A new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine*, *44*(7), 1349–1360. https://doi.org/10.1017/S0033291713001979

Wichers, M., Groot, P. C., Psychosystems, E.S.M. Group, & E.W.S. Group. (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

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer. ISBN 978-3-319-24277-4. https://ggplot2.tidyverse.org

Received March 6, 2022
Revision received November 30, 2022
Accepted December 1, 2022 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.