To cite this article:

Schat, E.*, Schreuder, M. J.*, & Ceulemans, E. (in press). Statistical process control for real-time monitoring in clinical psychology: State of the art and future research agenda. *Neuroscience Applied*. https://doi.org/10.1016/j.nsa.2026.106991

Statistical Process Control for Real-Time Monitoring in Clinical Psychology: State of the Art and Future Research Agenda

Evelien Schat[1†], Marieke J. Schreuder[1,2†], Eva Ceulemans[1]

[1] Faculty of Psychology and Educational Sciences, KU Leuven—University of Leuven, Leuven, Belgium

[2] Department of Developmental Psychology, Tilburg School Social and Behavioral Sciences, Tilburg University, Tilburg, The Netherlands

† These authors contributed equally to this work.

Correspondence should be addressed to Evelien Schat, Faculty of Psychology and Educational Sciences, KU Leuven, Tiensestraat 102 box 3713, B-3000 Leuven. Email: evelien.schat@kuleuven.be.

Abstract

Mental disorders place a substantial burden on individuals and society, underscoring the need for preventive interventions that can be delivered to the right person at the right time. This, in turn, requires real-time monitoring of person-specific indicators of risk for mental disorders and identifying when risk is increasing. Statistical process control (SPC) procedures, originally developed for monitoring industrial processes, offer a promising framework for addressing this need, allowing for the real-time detection of meaningful person-specific changes. Although SPC has only recently been introduced in clinical psychology, initial empirical applications demonstrate its potential for forecasting changes in mental health. In this perspective, we introduce the main concepts of SPC and describe two prominent SPC procedures (Shewhart and EWMA), outlining when each is most appropriate to use and summarizing current empirical findings in clinical psychology. We also discuss how SPC procedures compare to other change-detection methods. Finally, we discuss recent methodological advances in the use of SPC within clinical psychology and propose a research agenda that identifies key challenges for future work.

Keywords: Statistical process control, real-time monitoring, momentary interventions, experience sampling method

1. Introduction

Mental disorders place an enormous burden on individuals and society, underscoring the need for preventive interventions (Vigo et al., 2016). Such preventive interventions are likely most effective when provided to the right person, at the right time. To this end, it is necessary to monitor person-specific indicators of risk for mental disorders in real time. Until now, these risk indicators have primarily been tracked using experience sampling methods (ESM), the gold standard for studying how psychological constructs fluctuate in daily life (Myin-Germeys et al., 2009). In ESM studies, participants are asked to report on their momentary mental states using smartphones, multiple times a day for multiple days, while going about their daily lives. This generates dense, ecologically valid time-series data that allow researchers to examine how mental states change prior to the onset or recurrence of mental disorders (Dejonckheere et al., 2019; Houben et al., 2015). To leverage such insights for the real-time detection of increasing risk for mental disorders, it is necessary to adopt statistical techniques that yield person-specific and real-time predictions. A recent advancement in this direction concerns the introduction and refinement of statistical process control (SPC) procedures for monitoring ESM data (Schat et al., 2023; Smit et al., 2023). First empirical applications using ESM data—focusing on forecasting recurrent depression—have shown that real-time monitoring of mental states using SPC is promising and that meaningful changes in risk indicators could be detected up to a month prior to the recurrence of depression (Marieke J. Schreuder et al., 2024; Snippe et al., 2023). Beyond ESM, SPC has also been applied to mobile passive sensing data to predict depressive and (hypo)manic episodes in patients with bipolar disorder (Ludwig et al., 2025).

In this perspective, we first introduce the main principles of SPC and describe two prominent SPC procedures, outlining when each is most appropriate and summarizing how they have been applied to empirical psychological data. We then consider how SPC procedures compare to other change-detection methods. Finally, we discuss recent advances in the use of SPC within clinical psychology and propose a research agenda that highlights what we consider to be the most crucial challenges for future research.

2. What is SPC and why is it relevant for momentary risk detection

SPC is a family of methods designed for real-time monitoring of intensive longitudinal data. Originally developed in industry to ensure product quality and detect deviations in manufacturing processes (Montgomery, 2019; Shewhart, 1931), SPC has since been proven useful in various other domains (e.g., finance: Jumah et al., 2012; healthcare: Perez-Benitez et al., 2023; public transport: Zhang et al., 2022). Applying SPC in practice requires two phases of data collection. In phase I, baseline data are collected to capture the process's natural variability. Based on the mean and standard deviation of these data and a number of assumptions (see Section *3.1 Assumptions of SPC procedures*), upper and lower control limits[1] are calculated, representing the range of values that is considered normal. In phase II, monitoring begins and incoming data are compared against these control limits. As long as the monitored scores remain within the limits, the process is considered to be in control. When a score exceeds a control limit, this suggests that the process has gone out of control[2], indicating the need for intervention.

This idea has recently been translated to clinical psychology, where SPC has been proposed as a promising technique for anticipating the onset and recurrence of mental disorders (Schat et al., 2023; Smit et al., 2023). By monitoring an individual's mental states over time, SPC aims to detect meaningful changes that may signal increasing risk for worsening mental health. In other words, SPC procedures may help to infer whether a specific person, at a specific moment in time, reports emotions, cognitions, or behaviors that deviate from what is normal for that individual in healthy periods. Concretely, in phase I an individual's data are used to quantify the variability of their mental states and to set individual-specific control limits that indicate their normal range of scores. In phase II, incoming data are compared against these limits to determine whether and when the individual's mental states

[1] In some applications, only one control limit is used—either upper or lower—depending on the expected direction of change. For example, when monitoring high-arousal negative affect in order to foresee worsening depression, only an upper control limit is needed (Snippe et al., 2023). This is because risk for depression is reflected in increasing, rather than declining, high-arousal negative affect.

[2] In SPC, the term out-of-control is a technical label indicating that an observation exceeds the control limits. It should not be interpreted substantively or clinically. The term out of bounds has also been used to describe such observations (Ludwig et al., 2025).

deviate from this normal range (i.e., go out of control). Such deviations may be considered warning signs for impending mental health problems and indicate that intervention is needed.

2.1 Two key SPC procedures: Shewhart and EWMA

Two SPC procedures that may be particularly relevant for clinical psychology are the Shewhart procedure and the exponentially weighted moving average (EWMA) procedure. In this section, we describe these two procedures, outline for which inferential goals each is most appropriate to use and review applications to empirical psychological data.

2.1.1 Shewhart procedure

The Shewhart procedure (Shewhart, 1931) monitors observed scores directly, meaning that decisions about whether the process is out of control are based solely on the current observation. The control limits are calculated from phase I data as the phase I average ± L times the phase I standard deviation. The width parameter L is chosen based on the desired false positive rate (Montgomery, 2019; Schat et al., 2023), which describes the probability of detecting a warning sign even though the individual remains healthy. The Shewhart procedure is particularly useful for detecting larger, abrupt changes in mental health (Montgomery, 2019; Schat et al., 2023; Shewhart, 1931). Figure 1 shows an example of the Shewhart procedure: phase I data are used to calculate the control limits (Figure 1A), and in phase II the observed scores are monitored and compared against the limits. Figure 1B shows phase II data containing occasional outliers, all of which are detected by the Shewhart procedure. Figure 1C shows phase II data with a sustained mean change, for which Shewhart signals several out-of-control scores but does not continuously flag the process as out of control.

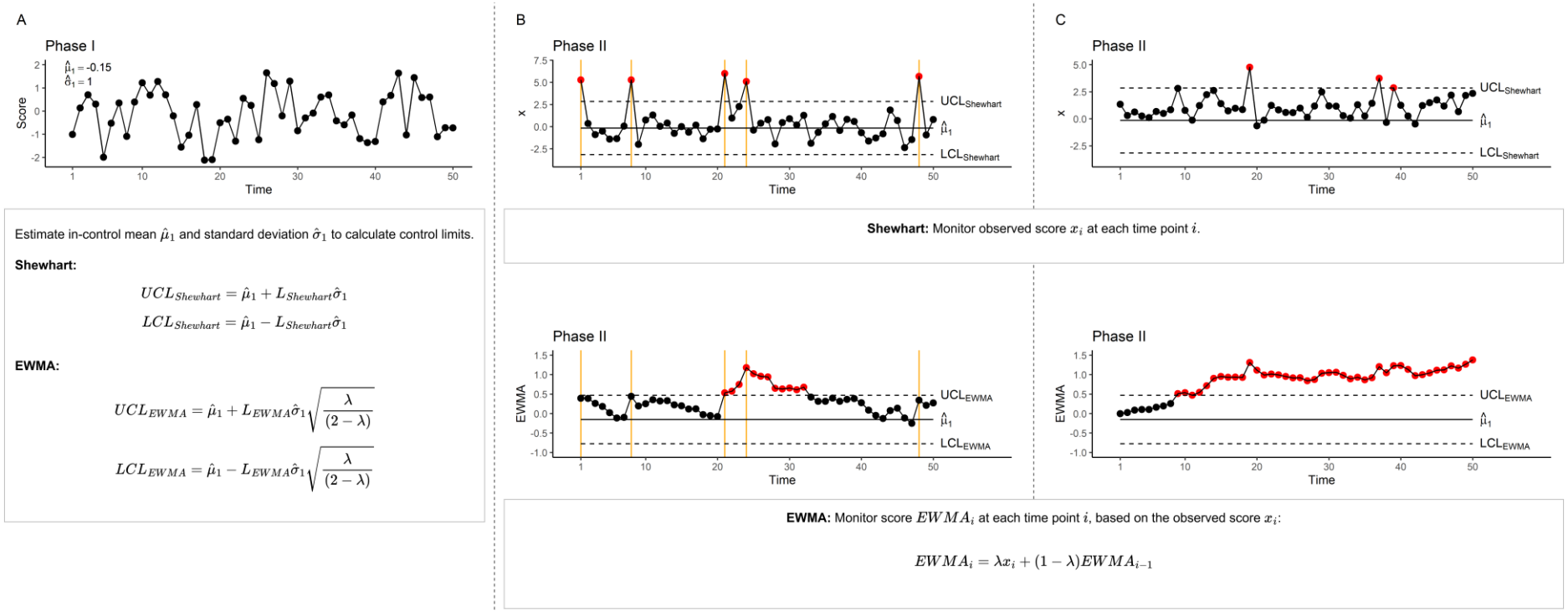
For processes prone to sudden changes that require immediate action, the Shewhart procedure is often most suitable. This procedure detects outliers in raw data, which may be relevant when monitoring for suicidal ideation, binge eating episodes, or panic attacks—events that can unfold rapidly and may require an immediate intervention (e.g., in Figure 1B an intervention would be

implemented immediately after the detected warning signs at time points 1, 8, 21, 24, and 48). Such immediate interventions could for instance follow the framework of Just-In-Time Adaptive Interventions (JITAIs), which promise to deliver the right type of support, for the right person, at the right moment in time (Nahum-Shani et al., 2018).

Although the Shewhart procedure has not yet been applied to detect changes in mental health, it holds substantial potential for application, as shown by studies on momentary interventions that have explored similar outlier-detection methods in ESM data. For instance, a study on an ecological momentary intervention for improving self-esteem in vulnerable youth triggered an intervention whenever youth reported high negative affect or stress, which was operationalized as a score above 4 (on a scale from 1 to 7, negative affect) or below 0 (on a scale from -3 to 3, stress; Reininghaus et al., 2024). Similarly, Saulnier and colleagues delivered an intervention when participants reported anxiety levels above a pre-defined threshold (Saulnier et al., 2024). The advantage of such generic, one-size-fits-all thresholds is that they do not require the collection of phase I data. The downside, however, is that they do not do justice to individual differences in the distribution of, for instance, negative effect, stress, or anxiety. To accommodate the latter, recent work has used quantiles of phase I data in order to detect outliers. For instance, interventions may be triggered when individuals report scores beyond the 20th and 80th quantiles detected in their phase I data (Schreuder et al., 2025). Given that these quantiles describe the scores that–for a particular person–may be considered extreme, the latter approach is more in line with the Shewhart procedure.

Figure 1

Example of the Shewhart and EWMA procedures



Note. (A) Phase I data used to calculate the upper (UCL) and lower (LCL) control limits. L determines the width of the control limits and is set such that, on average, one false positive occurs every 370 time points (i.e., $L_{Shewhart}$ = 3 and L_{EWMA} = 2.7 with λ = .10). The parameter λ denotes the weight that is given to incoming observations when computing the EWMA scores. (B) Shewhart (top) and EWMA (bottom) applied to phase II data containing occasional outliers (indicated by the orange vertical lines at time points 1, 8, 21, 24, and 48). (C) Shewhart (top) and EWMA (bottom) applied to phase II data, where a sustained mean change of 1σ was introduced at the start of phase II. The dashed horizontal lines represent the UCLs and LCLs, the solid horizontal line denotes the phase I mean $\hat{\mu}_1$. The red dots indicate the out-of-control scores.

2.1.2 EWMA procedure

The EWMA procedure (Roberts, 1959) does not track raw observations, but instead combines the current observation with all previous observations into a weighted sum that is monitored over time. Recent data receive more weight, while weights for more distant observations decrease exponentially over time. Here, similar to the Shewhart procedure, the upper and lower control limits depend on the phase I average, standard deviation, and the expected false positive rate (Montgomery, 2019; Schat et al., 2023). For EWMA, the calculation of these limits[3] additionally takes into account the weight given to current (λ) versus previous observations. EWMA is particularly useful for detecting smaller (and gradual), sustained changes in mental health (Montgomery, 2019; Roberts, 1959; Schat et al., 2023). An example of the EWMA procedure is shown in Figure 1: as with Shewhart, phase I data are used to calculate the control limits (Figure 1A). In phase II, EWMA scores are then computed and monitored (Figures 1B and 1C). In contrast to Shewhart, EWMA detects fewer isolated outliers (Figure 1B) but is more sensitive to sustained changes, and identifies the mean change at time point 9 and continuously signals the process as out of control thereafter (Figure 1C).

While the lingering out-of-control scores provided by EWMA might be favorable in case of reliable detection of upcoming change, it is less suitable for triggering interventions. Indeed, in case out-of-control scores would lead to interventions, EWMA may overburden participants due to excessive, repeated triggering of interventions. The EWMA procedure is thus most appropriate to use when a process unfolds more slowly and a slight delay in detecting impending change is acceptable (Carson & Yeh, 2008; Roberts, 1959). Such slowly evolving changes may for instance feature in depression, cognitive decline, or burnout.

So far, EWMA has been applied to two empirical data sets: one on recurrent depression and another on bipolar disorder. In the first data set, EWMA was applied to ESM data of 41 remitted

[3] EWMA control limits can be calculated using (1) steady-state limits as described here, or (2) the time-dependent formula, which includes an extra term $[1-(1-\lambda)^{2t}]$. The time-dependent limits are relevant when control limits are computed during—and not after—phase I, and adjust for the initial uncertainty about the phase I estimates (i.e., average and standard deviation).

individuals who gradually discontinued their antidepressant medication, 22 of whom experienced a recurrence of depression (Marieke J. Schreuder et al., 2024; Snippe et al., 2023). These studies demonstrated that monitoring changes in the intensity and variability of mental states using the EWMA procedure allows for predicting recurrent depression with 59-91% sensitivity and 47-90% specificity, depending on the mental state that was monitored. In the most optimal scenario, 91% of recurrences were correctly identified, while false alarms were detected in 42% of those who remained in remission. In the second data set, EWMA was used for predicting 26 depressive and 20 (hypo)manic episodes in 28 patients with bipolar disorder (Ludwig et al., 2025). This study monitored both passive sensing data and a bipolar mood item, and found that the mood item was more predictive of shifts in depression and mania compared to the passive sensing data. However, it was also concluded that the frequency of false alarms undermined clinical utility. Together, these first empirical findings are promising, but also point to further research questions that warrant attention, which we revisit below (see Section 3, *Advances of SPC in clinical psychology and research agenda*).

2.2 How does SPC compare to other change-detection methods?

In order to put the performance of SPC procedures into perspective, it is important to also consider other methods for forecasting mental health problems based on intensive longitudinal data. In this section, we will therefore review two alternative forecasting approaches which received considerable attention from researchers in the field of clinical psychology over the past decade, including complex systems-derived methods (Helmich et al., 2024; Scheffer et al., 2024) and machine learning methods (Leaning et al., 2024).

Firstly, complex-systems based methods rely on the idea that–if mental health behaves according to principles from complex dynamic systems–sudden transitions in mental health should be anticipated by changing affect dynamics (Scheffer et al., 2024). Unlike in SPC, these changes are not based on a comparison with phase I data, but rather continuously tracked based on moving window methods (Dakos et al., 2012). This means that affect dynamics are computed within a segment of the

time series (e.g., t = 1 to t = 11), recomputed on a next segment (e.g., t = 2 to t = 12), and so on. The result is a new time series of affect dynamics, where rising trends are considered indicative of warning signs for transitions in mental health. So far, however, the latter warning signs seem not sufficiently accurate for anticipating mental disorders (Helmich et al., 2024). Further, complex-systems based warning signs have only been detected in retrospect–*after* delineating sudden changes in mental health–which complicates clinical applications. SPC procedures address this limitation because they are designed to detect warning signs in real time, that is, *before* a sudden shift in mental health has emerged. Given that such real-time detection is imperative from the perspective of prevention and early intervention, SPC procedures may better accommodate clinical needs.

A second approach to leveraging intensive longitudinal data for forecasting the onset or worsening of mental disorders is to use machine learning methods. While SPC is a univariate technique that requires an a priori idea of the type of construct that should be monitored in order to foresee impending mental health problems, these methods allow for extracting those predictors that best predict future mental disorders in a data-driven manner (e.g., based on prediction errors). This means that, from a set of variables, machine learning methods can stratify those variables that best predict mental health problems. Without proper cross-validation, however, it remains uncertain how well those variables predict worsening mental health in a new sample (Leaning et al., 2024). This links to an essential limitation of many machine learning studies, namely that these studies often predict mental disorders at the between-person level instead of the within-person level. Specifically, machine learning is often used to discriminate *between* healthy controls and individuals with mental disorders, which sharply contrasts with the goal of detecting worsening mental health *within* individuals (Langener & Jacobson, 2025). As an exception, several recent studies used machine learning models to derive person-specific predictions from ESM data, about smoking (Fisher & Soyster, 2019), loneliness, procrastination, and studying (Beck & Jackson, 2022), or binge eating and alcohol use (Leenaerts et al., 2024). These studies found that such behaviors and experiences can be predicted reasonably well. However, these predictions were derived by fitting machine learning models for each person

separately, which is only possible *after* individuals reported the outcome (e.g., smoking, feeling lonely, etc.) and therefore not suited for predicting more seldom but significant impairments such as mental disorders. Again, the comparative advantage of SPC is that it is designed to yield real-time predictions, without requiring prior occurrences of the outcome in question.

3. Advances of SPC in clinical psychology and research agenda

While the first empirical applications of SPC procedures yielded promising results, several open questions need to be answered prior to implementing SPC procedures in (clinical) practice. In this section, we review recent advances in applying SPC within clinical psychology and outline a research agenda that lists what we consider to be the most important challenges ahead. Addressing these questions is not just relevant within the realm of SPC research, but rather may advance the broader research field covering the detection of (Helmich et al., 2024) and intervention on (Schneider et al., 2024) momentary changes in mental health.

3.1 Assumptions of SPC procedures

SPC procedures build on the assumption that the monitored data are independently sampled from a normal distribution, and that a sufficient amount of phase I data are available to reliably compute the control limits (for alternative approaches to establishing a baseline, see Section *3.4 Phase I data*). These assumptions are often violated in intensive longitudinal clinical data; for example, ESM measures of negative affect of healthy individuals are typically right skewed (Haslbeck et al., 2023), and temporal dependence between observations is common (Houben et al., 2015; Kuppens et al., 2010). Violations of these assumptions can lead to increased false alarms or reduced sensitivity to detect changes (see Schat et al., 2023, for a detailed discussion of assumption violations in EWMA and Shewhart procedures). To address violations of normality and independence, computing and monitoring day averages has been proposed (Schat et al., 2023). Taken together, researchers should carefully consider the alignment of intensive longitudinal data characteristics with SPC assumptions.

Future research could examine the performance of SPC procedures that explicitly accommodate non-normal or discrete data (Borror et al., 1998; Weiß & Atzmüller, 2010).

3.2 Long-term intensive longitudinal data

A crucial challenge for studying the performance of SPC is the limited availability of long-term intensive longitudinal data in which meaningful changes in mental health are well established. An important next step for advancing the field—and for rigorously evaluating SPC—is therefore to collect intensive longitudinal data with SPC applications or broader monitoring goals in mind. It is also essential that the incidence of relevant changes (e.g., depression relapse, binge-eating episodes) is sufficiently high to allow robust evaluation of SPC performance (e.g., sensitivity, specificity; Ebner-Priemer & Santangelo, 2020). In addition, accurately recording the timing of these changes using methods independent of SPC (e.g., based on clinical interviews, exceeding cut-offs on validated questionnaires, or change point analyses) is crucial, as it provides a reference for validating whether SPC correctly identifies moments of increased risk or deterioration. Hence, in an ideal scenario, future studies should last long enough to capture clinically relevant changes in mental states, while also having sufficient temporal resolution to pinpoint these changes in time.

To make long-term, intensive longitudinal designs more feasible, strategies are needed that reduce participant burden while maintaining data quality. A first step in this direction is to reconsider the optimal sampling frequency when applying SPC. In case higher sampling frequencies are necessary, it is possible to adopt adaptive sampling approaches—such as variable size EWMA—which lower the frequency of assessments when the process appears stable and increase it when there may be a potential change (De Pauw et al., 2025). It is also possible to lower the burden of long-term ESM data collection through adopting adaptive testing (Schneider et al., 2024), planned missingness designs (Rein et al., 2023), and increasing the degree of personalization (Olthof et al., 2023).

3.3 What to measure

So far, SPC research has mainly focused on monitoring the mean and variability of affective states and repetitive negative thinking. Yet, it may well be that monitoring person-specific indices of risk, more complex measures of affect dynamics, or alternative data modalities yields more accurate warning signs. First, regarding the monitoring of personalized indicators of risk, earlier work has shown substantial inter-individual variation in terms of what precipitates mental health problems (Hewitt & Birchwood, 2012; Pede et al., 2017). This implies that for real-time monitoring purposes, personalized risk indicators–such as the early signs of deterioration that are listed in relapse prevention plans when patients discontinue treatment–may be preferred over more generic mental states. Second, regarding the monitoring of complex measures of affect dynamics, it could be that increased reactivity to stressors or a slowing recovery of affect yield reliable warning signs (Olthof et al., 2025). In agreement, it has consistently been shown that mental disorders coincide with altered reactions to and recoveries from negative or stressful experiences (Kalisch et al., 2021; Ong & Leger, 2022). Third, alternative data modalities may offer a less burdensome way to collect large amounts of data over long periods, although their utility for identifying warning signs for impending mental disorders still awaits empirical substantiation. One prominent example is passive sensing data collected via smartphone or wearables. Although some studies report encouraging associations between passive sensing indicators and mental health/disorders (Bufano et al., 2023; Choi et al., 2024; Sun et al., 2023; Terhorst et al., 2025), others have raised methodological concerns, noting that many indicators lack thorough validation and thus calling their usefulness into question (Amin et al., 2025; De Angel et al., 2022; Elmer, Fernandez, et al., 2025; Kunkels et al., 2023; Niemeijer & Kuppens, 2024; Ortiz & Mulsant, 2024; van der Mee et al., 2025). Moreover, passive sensing data and their corresponding self-report measures may not necessarily measure the same constructs (Siepe et al., 2025), and initial evidence suggests that applying SPC to passive sensing data can be challenging (Ludwig et al., 2025). Hence, further research is needed to explore the utility of passively collected data for detecting warning signs for worsening mental health (Matcham et al., 2024), starting with a rigorous evaluation of the validity and reliability of these data

(Giurgiu et al., 2025). Another potential modality is speech-derived data, though—as with passive sensing—findings are mixed, with some studies reporting promising associations between speech-derived indicators and mental disorders (Koops et al., 2021) and others questioning their utility (Parola et al., 2020). Ultimately, as with many design-related choices in ESM research, the question of what to measure is both a theoretical and a practical one: It depends on our theoretical insight in the outcome and the period that precedes it, as well as practicalities (e.g., considering the burden and feasibility of the data collection).

3.4 Phase I data

By definition, SPC considers person-specific control limits which are derived from baseline data (i.e., phase I data collected when a person is in good mental health). However, this dependence on baseline data limits the feasibility of SPC. Some studies have explored whether phase I can be (partially) replaced with historical data from the general population. For example, historical data can be conditioned on relevant covariates (e.g., age, sex, etc.; Schreuder, Kuppens, et al., 2024) or augmented with a limited amount of person-specific data (Schat et al., 2024). The latter approach is similar to dynamically updating control limits based on newly gathered ESM data (cf. the idea of proximal baseline states; Smyth et al., 2023). While using historical data may provide an appealing solution to the need for large amounts of person-specific phase I data, the aforementioned approaches have shown limited effectiveness. It could be that other methods for leveraging historical data–for instance, using K-nearest neighbors methods to identify individuals who are similar to the monitored individual–prove more promising (Lutz et al., 2005). Still, however, using historical data inevitably reduces person-specificity, which may impact the identification of moments of risk. The extent of this impact likely depends on what is being monitored: for some variables (e.g., amount of alcohol consumptions per day, or instruments with established cut-offs), relatively generic or one-size-fits-all control limits may suffice (Saulnier et al., 2024), while for others (e.g., feeling sad), a more person-specific control limit may be needed. An additional complexity is that monitored variables may show response shifts–which in turn

may lead to a shift in a person's baseline and control limits—suggesting that phase I data should be revisited or even periodically updated.

3.5 Relevant timeframe

The pursuit of warning signs for upcoming mental health problems brings up a range of questions that tap into the clinical relevance of SPC. At what timeframe can indicators of imminent risk for mental health problems be expected (e.g., weeks, days, or hours in advance), are momentary (preventive) interventions effective, and if so, at what timeframe, and how do individuals experience the monitoring of their mental health and the provision of momentary interventions? With regard to the timeframe of warning signs, it should be noted that it is still unknown *when*, after the detection of warning signs, mental health problems may arise. Yet, such insight is crucial for intervention purposes: in case mental health problems are projected to emerge within several days, prevention is more urgent compared to when mental health problems are predicted to occur only after weeks or months. To gain further insight in the timeframe of warning signs, it may be relevant to consider (1) the outcome of interest and theoretical or clinical insights in (inter- and intra-individual differences in) the period preceding this outcome as well as (2) the timescale at which potential interventions are expected to operate. With regard to the former, it is for instance known that in case of depression, lack of energy and anhedonia commonly emerge in the month(s) prior to full-blown episodes (Lacoviello et al., 2010; Lobban et al., 2011), whereas psychosis is often anticipated by suspicion and social withdrawal that may already emerge years in advance (Häfner et al., 1999). In case prodromal states only emerge a short time before the mental disorder–or do not occur at all, which may be the case for patients with bipolar disorder (Jackson et al., 2003)–it is impossible to detect warning signs anticipating the disorder. Hence, although individual differences in the presence and duration of prodromal states complicate generic guidelines on the optimal timeframe of warning signs, insight in the period preceding the outcome of interest could guide the evaluation of warning signs. With regard to the timescale of interventions (point 2), interventions that motivate individuals to move (e.g., in the context of

monitoring sedentary behavior) can be expected to sort effects immediately, while interventions that are meant to build skills (e.g., re-appraisal) may require more training and hence can only be expected to sort effects over longer periods of time. An additional important point to consider when evaluating the timeframe of warning signs concerns the perspective of those who are receiving the warnings. Indeed, risk monitoring comes with ethical considerations that should be acknowledged in order to prevent, for instance, self-fulfilling prophecies (e.g., feelings of helplessness or despair when warning signs appear shortly before recurrent mental health problems) or apathy (e.g., when warnings appear too far in advance, or are perceived as false alarms; Neiders et al., 2025).

3.6 JITAIs

One of the most promising use cases for detecting warning signs for worsening mental health is to inform preventive interventions. The underlying assumption is that the prevention of mental disorders is most effective when provided in a personalized, timely manner. Indeed, preventive interventions–which are targeted at those individuals who are at increased risk for mental disorders–have been considered key in transitioning towards personalized psychiatry (Fusar-Poli et al., 2021). JITAIs provide an example of how such personalized and timely interventions could be implemented in individuals' daily lives. Briefly, JITAIs concern interventions that are delivered just-in-time–namely, when individuals are both in need and receptive of an intervention–and adaptively–namely, adjusted according to individuals' needs and context (Nahum-Shani et al., 2018). With regard to the just-in-time premise, many earlier studies have focused on identifying moments of high stress or symptom severity, reasoning that such moments may indicate a need for interventions (Olthof et al., 2025; Reininghaus et al., 2024; Saulnier et al., 2024; Schreuder et al., 2025). However, these moments may interfere with receptiveness, illustrating that moments of opportunity (need) and receptivity may not align well. Indeed, a recent study showed that JITAIs initiated by participants are received better than JITAIs that are triggered upon proposed moments of need (Elmer, Wolf, et al., 2025). This raises the question how the optimal moments for triggering JITAIs should be identified, which may be addressed through

systematically varying design factors (e.g., the constructs and thresholds that may inform the triggering of interventions, the contextual information that should be taken into account, etc.; van Genugten et al., 2025).

4. Conclusion

SPC offers a promising framework for real-time monitoring of person-specific indicators of risk, helping to detect worsening mental health *before* full-blown disorders may emerge. This perspective highlighted the potential of SPC in clinical psychology, summarized key empirical findings, and outlined important challenges for future research aimed at enhancing the feasibility and practical utility of SPC procedures. Addressing these challenges will not just benefit research into SPC, but rather may advance the broader field of research into using intensive longitudinal data for understanding changes in mental health.

References

Amin, R., Schreynemackers, S., Oppenheimer, H., Petrovic, M., Hegerl, U., & Reich, H. (2025). Use of mobile sensing data for longitudinal monitoring and prediction of depression severity: Systematic review. *Journal of Medical Internet Research*, *27*, e57418. https://doi.org/10.2196/57418

Beck, E. D., & Jackson, J. J. (2022). Personalized prediction of behaviors and experiences: An idiographic person-situation test. *Psychological Science*, *33*(10), 1767–1782. https://doi.org/10.1177/09567976221093307

Borror, C. M., Champ, C. W., & Rigdon, S. E. (1998). Poisson EWMA Control Charts. *Journal of Quality Technology*. https://doi.org/10.1080/00224065.1998.11979871

Bufano, P., Laurino, M., Said, S., Tognetti, A., & Menicucci, D. (2023). Digital phenotyping for monitoring mental disorders: Systematic review. *Journal of Medical Internet Research*, *25*, e46778. https://doi.org/10.2196/46778

Carson, P. K., & Yeh, A. B. (2008). Exponentially weighted moving average (EWMA) control charts for monitoring an analytical process. *Industrial and Engineering Chemistry Research*, *47*(2), 405–411. https://doi.org/10.1021/ie070589b

Choi, A., Ooi, A., & Lottridge, D. (2024). Digital phenotyping for stress, anxiety, and mild depression: Systematic literature review. *JMIR MHealth and UHealth*, *12*, e40689. https://doi.org/10.2196/40689

Dakos, V., Carpenter, S. R., Brock, W. A., Ellison, A. M., Guttal, V., Ives, A. R., Kéfi, S., Livina, V., Seekell, D. A., van Nes, E. H., & Scheffer, M. (2012). Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS ONE*, *7*(7), e41010. https://doi.org/10.1371/journal.pone.0041010

De Angel, V., Lewis, S., White, K., Oetzmann, C., Leightley, D., Oprea, E., Lavelle, G., Matcham, F., Pace, A., Mohr, D. C., Dobson, R., & Hotopf, M. (2022). Digital health tools for the passive monitoring of depression: a systematic review of methods. *Npj Digital Medicine*, *5*(1), 1–14. https://doi.org/10.1038/s41746-021-00548-8

De Pauw, F., Schat, E., Tuerlinckx, F., & Ceulemans, E. (2025). *Reducing participant burden of ESM-based monitoring of mood disorder onset through variable sample size EWMA* (Unpublished manuscript).

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

Ebner-Priemer, U., & Santangelo, P. (2020). Digital phenotyping: Hype or hope? *The Lancet Psychiatry*, *7*(4), 297–299. https://doi.org/10.1016/S2215-0366(19)30380-3

Elmer, T., Fernandez, A., Stadel, M., Kas, M. J. H., & Langener, A. M. (2025). Bidirectional associations between smartphone usage and momentary well-being in young adults: Tackling

methodological challenges by combining experience sampling methods with passive smartphone data. *Emotion*, *25*(5), 1065–1078. https://doi.org/10.1037/emo0001485.supp

Elmer, T., Wolf, M., Snippe, E., & Scholz, U. (2025). A social support just-in-time adaptive intervention for individuals with depressive symptoms: Feasibility study with a microrandomized trial design. *JMIR Mental Health*, *12*(e74103), 1–19. https://doi.org/10.2196/74103

Fisher, A. J., & Soyster, P. D. (2019). *Generating accurate personalized predictions of future behavior: A smoking exemplar*. https://doi.org/10.31234/osf.io/e24v6

Fusar-Poli, P., Correll, C. U., Arango, C., Berk, M., Patel, V., & Ioannidis, J. P. A. (2021). Preventive psychiatry: A blueprint for improving the mental health of young people. *World Psychiatry*, *20*(2), 200–221. https://doi.org/10.1002/wps.20869

Giurgiu, M., von Haaren-Mack, B., Fiedler, J., Woll, S., Burchartz, A., Kolb, S., Ketelhut, S., Kubica, C., Nigg, C., Timm, I., Thron, M., Schmidt, S., Wunsch, K., Müller, G., Nigg, C. R., Woll, A., Reichert, M., Ebner-Priemer, U., & Bussmann, J. B. (2025). The wearable landscape: Issues pertaining to the validation of the measurement of 24-h physical activity, sedentary, and sleep behavior assessment. *Journal of Sport and Health Science*, *14*, 1–9. https://doi.org/10.1016/j.jshs.2024.101006

Häfner, H., Löffler, W., Maurer, K., Hambrecht, M., & An Der Heiden, W. (1999). Depression, negative symptoms, social stagnation and social decline in the early course of schizophrenia. *Acta Psychiatrica Scandinavica*, *100*(2), 105–118. https://doi.org/10.1111/j.1600-0447.1999.tb10831.x

Haslbeck, J., Ryan, O., & Dablander, F. (2023). Multimodality and skewness in emotion time series. *Emotion*, *23*(8), 2117–2141. https://doi.org/10.1037/emo0001218

Helmich, M. A., Schreuder, M. J., Bringmann, L. F., Riese, H., Snippe, E., & Smit, A. C. (2024). Slow down and be critical before using early warning signals in psychopathology. *Nature Reviews Psychology*, *3*(11), 767–780. https://doi.org/10.1038/s44159-024-00369-y

Hewitt, L., & Birchwood, M. (2012). Preventing relapse of psychotic illness: Role of self-monitoring of prodromal symptoms. *Disease Management & Health Outcomes*, *10*(7), 395–407.

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

Jackson, A., Cavanagh, J., & Scott, J. (2003). A systematic review of manic and depressive prodromes. *Journal of Affective Disorders*, *74*(3), 209–217. https://doi.org/10.1016/S0165-0327(02)00266-5

Jumah, J. A. Bin, Burt, P., & Buttram, M. B. (2012). An exploration of quality control in banking and finance. *International Journal of Business and Social Science*, *3*(6), 273–277.

Kalisch, R., Köber, G., Binder, H., Ahrens, K. F., Basten, U., Chmitorz, A., Choi, K. W., Fiebach, C. J., Goldbach, N., Neumann, R. J., Kampa, M., Kollmann, B., Lieb, K., Plichta, M. M., Reif, A., Schick, A., Sebastian, A., Walter, H., Wessa, M., … Engen, H. (2021). The frequent stressor and mental health monitoring-paradigm: A proposal for the operationalization and measurement of

resilience and the identification of resilience processes in longitudinal observational studies. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.710493

Koops, S., Brederoo, S. G., de Boer, J. N., Nadema, F. G., Voppel, A. E., & Sommer, I. E. (2021). Speech as a biomarker for depression. *CNS & Neurological Disorders - Drug Targets*, *22*(2), 152–160. https://doi.org/10.2174/1871527320666211213125847

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

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

Lacoviello, B. M., Alloy, L. B., Abramson, L. Y., & Choi, J. Y. (2010). The early course of depression: A longitudinal investigation of prodromal symptoms and their relation to the symptomatic course of depressive episodes. *Journal of Abnormal Psychology*, *119*(3), 459–467. https://doi.org/10.1037/a0020114

Langener, A. M., & Jacobson, N. C. (2025). *Just in time or just a guess? Addressing challenges in validating prediction models based on longitudinal data*. https://doi.org/10.31219/osf.io/38v42_v1

Leaning, I. E., Ikani, N., Savage, H. S., Leow, A., Beckmann, C., Ruhé, H. G., & Marquand, A. F. (2024). From smartphone data to clinically relevant predictions: A systematic review of digital phenotyping methods in depression. In *Neuroscience and Biobehavioral Reviews* (Vol. 158). Elsevier Ltd. https://doi.org/10.1016/j.neubiorev.2024.105541

Leenaerts, N., Soyster, P., Ceccarini, J., Sunaert, S., Fisher, A., & Vrieze, E. (2024). Person-specific and pooled prediction models for binge eating, alcohol use and binge drinking in bulimia nervosa and alcohol use disorder. *Psychological Medicine*, *54*, 2758–2773. https://doi.org/10.1017/S0033291724000862

Lobban, F., Solis-Trapala, I., Symes, W., & Morriss, R. (2011). Early warning signs checklists for relapse in bipolar depression and mania: Utility, reliability and validity. *Journal of Affective Disorders*, *133*(3), 413–422. https://doi.org/10.1016/j.jad.2011.04.026

Ludwig, V. M., Bittendorf, C. A., Reinhard, I., Guth, M., Mühlbauer, E., Hartnagel, L.-M., Severus, W. E., Bauer, M., Ritter, P., & Ebner-Priemer, U. W. (2025). Predicting depressive and manic episodes in patients with bipolar disorder using statistical process control methods on passive sensing data. *Journal of Psychopathology and Clinical Science*, *134*(8), 971–981. https://doi.org/10.1037/abn0001002

Lutz, W., Barkham, M., Stiles, W. B., Leach, C., Lucock, M., Evans, C., Noble, R., & Iveson, S. (2005). Predicting change for individual psychotherapy clients on the basis of their nearest neighbors. *Journal of Consulting and Clinical Psychology*, *73*(5), 904–913. https://doi.org/10.1037/0022-006X.73.5.904

Matcham, F., Carr, E., Meyer, N., White, K. M., Oetzmann, C., Leightley, D., Lamers, F., Siddi, S., Cummins, N., Annas, P., de Girolamo, G., Haro, J. M., Lavelle, G., Li, Q., Lombardini, F., Mohr, D. C., Narayan, V. A., Penninx, B. W. H. J., Coromina, M., … Hotopf, M. (2024). The relationship between wearable-derived sleep features and relapse in Major Depressive Disorder. *Journal of Affective Disorders*, *363*, 90–98. https://doi.org/10.1016/j.jad.2024.07.136

Montgomery, D. C. (2019). *Introduction to statistical quality control* (8th ed.). John Wiley & Sons, Inc.

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

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine*, *52*(6), 446–462. https://doi.org/10.1007/s12160-016-9830-8

Neiders, I., Mežinska, S., & van Haren, N. E. M. (2025). Ethical and social issues in prediction of risk of severe mental illness: A scoping review and thematic analysis. *BMC Psychiatry*, *25*, 501. https://doi.org/10.1186/s12888-025-06949-3

Niemeijer, K., & Kuppens, P. (2024). Emotion detection with mobile sensing. In M. R. Mehl, M. Eid, C. Wrzus, G. M. Harari, & U. Ebner-Priemer (Eds.), *Mobile Sensing in Psychology: Methods and Applications* (pp. 561–580). Guilford Press.

Olthof, M., Hasselman, F., Aas, B., Lamoth, D., Scholz, S., Daniels-Wredenhagen, N., Goldbeck, F., Weinans, E., Strunk, G., Schiepek, G., Bosman, A. M. T., & Lichtwarck-Aschoff, A. (2023). The best of both worlds? General principles of psychopathology in personalized assessment. *Journal of Psychopathology and Clinical Science*, *132*(7), 808–819. https://doi.org/10.1037/abn0000858

Olthof, M., Mestdagh, M., Hasselman, F., Maciejewski, D., Bunge, A., Dejonckheere, E., Schat, E., Ceulemans, E., & Lichtwarck-Aschoff, A. (2025). *Micro interventions based on real-time monitoring of affective instability: A proof-of-concept study*. https://doi.org/10.31234/osf.io/e7djf_v1

Ong, A. D., & Leger, K. A. (2022). Advancing the study of resilience to daily stressors. *Perspectives on Psychological Science*, *17*(6), 1591–1603. https://doi.org/10.1177/17456916211071092

Ortiz, A., & Mulsant, B. H. (2024). Beyond step count: Are we ready to use digital phenotyping to make actionable individual predictions in psychiatry? *Journal of Medical Internet Research*, *26*(e59826), 1–6. https://doi.org/10.2196/59826

Parola, A., Simonsen, A., Bliksted, V., & Fusaroli, R. (2020). Voice patterns in schizophrenia: A systematic review and Bayesian meta-analysis. *Schizophrenia Research*, *216*, 24–40. https://doi.org/10.1016/j.schres.2019.11.031

Pede, V. B., Jaiswal, S. V, & Sawant, V. A. (2017). Study of prodromal and residual symptoms of depression. *Ind Psychiatry J.*, *26*(2), 121–127. https://doi.org/10.4103/ipj.ipj_19_18

Perez-Benitez, B. E., Tercero-Gomez, V. G., & Khakifirooz, M. (2023). A review on statistical process control in healthcare: Data-driven monitoring schemes. *IEEE Access*, *11*, 56248–56272. https://doi.org/10.1109/ACCESS.2023.3282569

Rein, M. T., Vogelsmeier, L. V. D. E., & Bolsinova, M. (2023). *Assessing life satisfaction in everyday life: An example of scale development for ecological momentary assessment*. https://doi.org/10.31234/osf.io/47bry

Reininghaus, U., Daemen, M., Postma, M. R., Schick, A., Hoes-Van Der Meulen, I., Volbragt, N., Nieman, D., Delespaul, P., De Haan, L., Van Der Pluijm, M., Breedvelt, J. J. F., Van Der Gaag, M., Lindauer, R., Boehnke, J. R., Viechtbauer, W., Van Den Berg, D., Bockting, C., & Van Amelsvoort, T. (2024). Transdiagnostic ecological momentary intervention for improving self-esteem in youth exposed to childhood adversity: The SELFIE randomized clinical trial. *JAMA Psychiatry*, *81*(3), 227–239. https://doi.org/10.1001/jamapsychiatry.2023.4590

Roberts, S. W. (1959). Control chart tests based on geometric moving averages. *Technometrics*, *1*(3), 239–250. https://doi.org/10.2307/1266443

Saulnier, K. G., Koscinski, B., Flynt, S., Accorso, C., & Allan, N. P. (2024). Brief observable anxiety sensitivity treatment: Intervention development and a pilot randomized-controlled acceptability and feasibility trial to evaluate a brief intervention for anxiety sensitivity social concerns. *Cognitive Behaviour Therapy*, *53*(2), 190–206. https://doi.org/10.1080/16506073.2023.2288551

Schat, E., Tuerlinckx, F., Schreuder, M. J., De Ketelaere, B., & Ceulemans, E. (2024). Forecasting the onset of depression with limited baseline data only: A comparison of a person-specific and a multilevel modeling based exponentially weighted moving average Approach. *Psychological Assessment*, *36*(6–7), 379–394. https://doi.org/10.1037/pas0001314

Schat, E., Tuerlinckx, F., Smit, A. C., de Ketelaere, B., & Ceulemans, E. (2023). Detecting mean changes in experience sampling data in real time: A comparison of univariate and multivariate statistical process control methods. *Psychological Methods*, *28*(6), 1335–1357. https://doi.org/10.1037/met0000447

Scheffer, M., Bockting, C. L., Borsboom, D., Cools, R., Delecroix, C., Hartmann, J. A., Kendler, K. S., Van De Leemput, I., Van Der Maas, H. L. J., Van Nes, E., Mattson, M., McGorry, P. D., & Nelson, B. (2024). A dynamical systems view of psychiatric disorders - practical implications: A review. *JAMA Psychiatry*, *81*(6), 624–630. https://doi.org/10.1001/jamapsychiatry.2024.0228

Schneider, S., Junghaenel, D. U., Smyth, J. M., Fred Wen, C. K., & Stone, A. A. (2024). Just-in-time adaptive ecological momentary assessment (JITA-EMA). *Behavior Research Methods*, *56*(2), 765–783. https://doi.org/10.3758/s13428-023-02083-8

Schreuder, M J, Kuppens, P., Schat, E., de Jonge, P., Hartman, C. A., & Ceulemans, E. (2024). Warning signals for mental health problems in at-risk young adults may be informed by momentary emotions reported by the general population: A novel application of the principles of statistical process control. *Mental Health Science*, *2*(3), e27. https://doi.org/10.1002/mhs2.72

Schreuder, Marieke J., Schat, E., Smit, A. C., Snippe, E., & Ceulemans, E. (2024). Monitoring emotional intensity and variability in real-time to forecast depression recurrence in remitted adults.

Journal of Consulting and Clinical Psychology, *92*(8), 505–516. https://doi.org/10.1037/ccp0000871

Schreuder, M. J., Schorrlepp, L., Maciejewski, D., Schipper, A., van Roekel, E., & Schreuder, M. (2025). *Just-in-time adaptive interventions to improve positive emotions in daily life: A randomized controlled experience sampling study*. https://doi.org/10.31234/osf.io/qyt2v_v2

Shewhart, W. A. (1931). *Economic control of quality of manufactured product*. Macmillan And Co Ltd.

Siepe, B. S., Tutunji, R., Rieble, C. L., Proppert, R. K. K., & Fried, E. I. (2025). Associations between ecological momentary assessment and passive sensor data in a large student sample. *Journal of Psychopathology and Clinical Science*. https://doi.org/10.1037/abn0001013

Smit, A. C., Schat, E., & Ceulemans, E. (2023). The exponentially weighted moving average procedure for detecting changes in intensive longitudinal data in psychological research in real-time: A tutorial showcasing potential applications. *Assessment*, *30*(5), 1354–1368. https://doi.org/10.1177/10731911221086985

Smyth, J. M., Zawadzki, M. J., Marcusson-Clavertz, D., Scott, S. B., Johnson, J. A., Kim, J., Toledo, M. J., Stawski, R. S., Sliwinski, M. J., & Almeida, D. M. (2023). Computing components of everyday stress responses: Exploring conceptual challenges and new opportunities. *Perspectives on Psychological Science*, *18*(1), 110–124. https://doi.org/10.1177/17456916221082108

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

Sun, S., Folarin, A. A., Zhang, Y., Cummins, N., Garcia-Dias, R., Stewart, C., Ranjan, Y., Rashid, Z., Conde, P., Laiou, P., Sankesara, H., Matcham, F., Leightley, D., White, K. M., Oetzmann, C., Ivan, A., Lamers, F., Siddi, S., Simblett, S., … Dobson, R. J. B. (2023). Challenges in using mHealth data from smartphones and wearable devices to predict depression symptom severity: Retrospective analysis. *Journal of Medical Internet Research*, *25*, e45233. https://doi.org/10.2196/45233

Terhorst, Y., Messner, E. M., Asare, K. O., Montag, C., Kannen, C., & Baumeister, H. (2025). Investigating smartphone-based sensing features for depression severity prediction: Observation study. *Journal of Medical Internet Research*, *27*, e55308. https://doi.org/10.2196/55308

van der Mee, D. J., Koyuncu, Z., & Lemmers-Jansen, I. L. J. (2025). Are you stressed or just excited? What the Garmin Stress Score can say about your mood. *Journal of Affective Disorders Reports*, *21*. https://doi.org/10.1016/j.jadr.2025.100974

van Genugten, C. R., Thong, M. S. Y., van Ballegooijen, W., Kleiboer, A. M., Spruijt-Metz, D., Smit, A. C., Sprangers, M. A. G., Terhorst, Y., & Riper, H. (2025). Beyond the current state of just-in-time adaptive interventions in mental health: A qualitative systematic review. *Frontiers in Digital Health*, *7*, 1460167. https://doi.org/10.3389/fdgth.2025.1460167

Vigo, D., Thornicroft, G., & Atun, R. (2016). Estimating the true global burden of mental illness. *The Lancet Psychiatry*, *3*(2), 171–178. https://doi.org/10.1016/S2215-0366(15)00505-2

Weiß, C. H., & Atzmüller, M. (2010). EWMA control charts for monitoring binary processes with applications to medical diagnosis data. *Quality and Reliability Engineering International*, *26*(8), 795–805. https://doi.org/10.1002/qre.1098

Zhang, Y., Wu, K., Yu, C., Zhang, S., & Cai, X. (2022). Application of statistical process control for structural health monitoring of a high-speed railway track system. *Applied Sciences*, *12*(12), 6046. https://doi.org/10.3390/app121260