Cite as: Kornacka M, Barnes S, Skorupski J, Skorupski MS, Jaskulska A, Para M, Szastok M: Psychotherapy in motion: opportunities and challenges of using experience sampling methods to assess the effectiveness of psychotherapy. Psychiatr Psychol Klin 2026; 26 (2): 155–166

Monika Kornacka[1], Steven Barnes[1], Jan Skorupski[1], Michał Stefan Skorupski[1], Anna Jaskulska[2], Małgorzata Para[1], Marta Szastok[1]

Received: 07.04.2025
Accepted: 17.12.2025
Published: 04.08.2026

Psychotherapy in motion: opportunities and challenges of using experience sampling methods to assess the effectiveness of psychotherapy

Psychoterapia w rytmie dnia: możliwości i wyzwania związane z użytkowaniem metod próbkowania doświadczeń w ocenie skuteczności i efektywności psychoterapii

[1] Emotion Cognition Lab, Institute of Psychology, SWPS University, Katowice, Poland
[2] KOBO Association, Warsaw, Poland
Correspondence: Monika Kornacka, Emotion Cognition Lab, Instytut Psychologii, Uniwersytet SWPS, Techników 9, 40-326 Katowice, Poland, e-mail: mkornacka@swps.edu.pl

[1] Emotion Cognition Lab, Instytut Psychologii, Uniwersytet SWPS, Katowice, Polska
[2] Stowarzyszenie KOBO, Warszawa, Polska
Adres do korespondencji: Monika Kornacka, Emotion Cognition Lab, Instytut Psychologii, Uniwersytet SWPS, ul. Techników 9, 40-326 Katowice, e-mail: mkornacka@swps.edu.pl

https://doi.org/10.15557/PiPK.2026.0018

ORCID iDs

1. Monika Kornacka https://orcid.org/0000-0003-2737-9236
2. Steven Barnes https://orcid.org/0000-0002-5114-2178
3. Jan Skorupski https://orcid.org/0009-0004-4997-2932
4. Michał Stefan Skorupski https://orcid.org/0000-0001-8661-8669
5. Anna Jaskulska https://orcid.org/0000-0002-2539-3934
6. Małgorzata Para https://orcid.org/0009-0008-7178-0063
7. Marta Szastok https://orcid.org/0000-0001-7674-7651

Abstract

Introduction and objective: Psychotherapy efficacy and effectiveness are traditionally assessed using pre-post symptom reports, which often fail to capture the dynamic nature of therapeutic change and may overlook individual patterns of interaction between symptoms and psychological processes. Experience sampling methods (ESM), encompassing daily diaries, ecological momentary assessment (EMA), and ambulatory assessment via wearable devices, address this limitation by collecting real-time, *in-situ* data on symptoms and psychological processes through mobile technologies. This approach enables the development of personalised models of patient functioning and dynamically tracks changes during psychotherapy, offering a more nuanced understanding of treatment mechanisms. Despite increasing recommendations for their use in clinical practice and exponential growth in access to technology, ESM remains relatively underutilised in evaluating psychotherapy efficacy and effectiveness. **Materials and methods:** This paper provides a narrative review exploring the opportunities and challenges of using ESM to assess psychotherapy effectiveness and efficacy. **Results:** ESM is shown to complement, rather than replace, traditional methods by capturing dynamic changes and contextual factors, recognising that retrospective self-reported questionnaires and ESM data do not always measure the same aspects of a given psychological process. The review outlines key areas, including an overview of ESM methodologies, the rationale for their use in psychotherapy evaluation, user experience and methodological challenges (e.g. compliance, UX, Open Science principles, and participatory approaches), data structure and analytical approaches (such as multilevel and network models), and ethical considerations related to ESM. **Conclusions:** Practical recommendations for integrating ESM into psychotherapy research are presented, drawing on general ESM guidelines (e.g. Löchner et al., 2025) and the findings of this review.

Keywords: experience sampling methods (ESM), ecological momentary assessment (EMA), psychotherapy efficacy, psychotherapy effectiveness, mechanisms of change in psychotherapy

Streszczenie

Wprowadzenie i cel: Tradycyjnie skuteczność i efektywność psychoterapii ocenia się na podstawie samooceny objawów przed i po interwencji, jednak często nie oddają one ani dynamicznego charakteru mechanizmów zmiany terapeutycznej, ani unikalnych wzorców interakcji między objawami a procesami psychologicznymi u danego pacjenta. Metody próbkowania doświadczeń (*experience sampling methods*, ESM), obejmujące pomiary dzienniczkowe, wielokrotne dzienne pomiary (*ecological momentary assessment*, EMA) oraz monitorowanie ambulatoryjne za pomocą noszonych elektronicznych urządzeń pomiarowych (*wearables*), odpowiadają na te wyzwania poprzez gromadzenie danych dotyczących objawów i procesów

Authors: Kornacka M, Barnes S, Skorupski J, Skorupski MS, Jaskulska A, Para M et al. Proofreading by ENSKA Agnieszka Kosarzycka.

psychologicznych w czasie rzeczywistym i w naturalnym otoczeniu, za pośrednictwem urządzeń mobilnych. Takie podejście umożliwia tworzenie spersonalizowanych wzorców funkcjonowania pacjentów i dynamiczne śledzenie zmian w trakcie psychoterapii, oferując bogatsze zrozumienie mechanizmów leczenia. Pomimo aktualnych rekomendacji dotyczących ich stosowania w praktyce klinicznej i istotnego wzrostu dostępności technologicznej metody ESM pozostają niedostatecznie wykorzystywane w ocenie skuteczności i efektywności psychoterapii. **Materiał i metody:** Niniejszy przegląd narracyjny przedstawia możliwości i wyzwania związane ze stosowaniem ESM w ocenie skuteczności i efektywności psychoterapii. **Wyniki:** Autorzy pracy podkreślają, że ESM uzupełnia, a nie zastępuje tradycyjne metody pomiaru, oceniając dynamiczne zmiany i czynniki kontekstualne; w tym kontekście istotne jest, że retrospektywne kwestionariusze samoopisowe i dane z ESM nie zawsze mierzą dokładnie te same aspekty danego procesu psychologicznego. Praca zawiera przegląd metodologii stosowanych w ESM, uzasadnia ich zastosowanie w ocenie psychoterapii, omawia wyzwania związane z metodologią i doświadczeniem użytkownika, strukturę danych i podejścia analityczne (takie jak modele wielopoziomowe i sieciowe) oraz kwestie etyczne związane z ESM. **Wnioski:** Pracę kończą praktyczne rekomendacje dotyczące włączania ESM do badań nad psychoterapią, oparte na ogólnych wytycznych dotyczących ESM (np. Löchner *et al.*, 2025) oraz na niniejszym przeglądzie.

Słowa kluczowe: metody próbkowania doświadczeń (ESM), wielokrotne dzienne pomiary (EMA), skuteczność psychoterapii, efektywność psychoterapii, mechanizm zmiany w psychoterapii

INTRODUCTION

When the effectiveness of psychotherapy is assessed using traditional approaches, the evaluation typically relies on the same self-reported symptom measures collected before and after the therapeutic intervention. However, two patients with the same disorder might exhibit entirely different clinical patterns (Rief et al., 2024). Furthermore, "before and after" measures alone overlook the dynamic processes unfolding during therapy. Current research in clinical psychology offers a solution: experience sampling methods (ESM). These involve brief assessments of symptoms and psychological processes delivered via mobile devices multiple times per day in patients' everyday lives. ESM captures real-time data, enabling the development of a personalised model of each patient's disorder and showing how it dynamically changes over the course of psychotherapy (Myin-Germeys et al., 2018). The present paper, provides a comprehensive review of the current use of ESM in evaluating psychotherapy effectiveness, together with its main benefits and challenges.

EXPERIENCE SAMPLING METHODS – THE IDEA

Experience sampling refers to a group of methods in which patients complete very short questionnaires at least once a day to assess their symptoms and psychological processes in ecological, real-world settings (Myin-Germeys and Kuppens, 2022; Myin-Germeys et al., 2009; Nezlek, 2012). It includes daily diaries, ecological momentary assessment (EMA), and daily sampling, and has been used in clinical psychology research for several decades (Myin-Germeys et al., 2018; Scollon et al., 2003). ESM can be also classified within the broader group of methods known as ambulatory assessment, which includes different forms of mobile sensing in addition to self-report questions (Myin-Germeys and Kuppens, 2022). The use of ESM has grown exponentially in recent years due to the increased accessibility of technological tools, the widespread adoption of smartphones and smartwatches by both researchers and patients, and the availability of survey platforms or solutions dedicated to ESM in research, such as m-Path (Mestdagh et al., 2023) or movisens (movisens GmbH, Karlsruhe, Germany). Experience sampling is also increasingly recommended in clinical practice (Hayes et al., 2020; Hofmann and Curtiss, 2018; Hofmann and Hayes, 2019; Myin-Germeys et al., 2024). Some authors suggest that ESM opens another "black box" of psychology – namely, patients' everyday life functioning (Löchner et al., 2025; Myin-Germeys et al., 2009). Despite these developments and recommendations, ESM remains relatively underutilised in evaluating psychotherapy efficacy and effectiveness (Hofmann, 2025; Mink et al., 2025; Verhagen et al., 2016), although its use has increased in recent years (e.g. Peterson et al., 2020; Robinaugh et al., 2020b). Likewise, therapists often lack practical tools to integrate it effectively into their clinical work (Kornacka et al., 2023; Weermeijer et al., 2023). It is important to emphasise that experience sampling is not intended to replace traditional pre-post and follow-up measures commonly used in clinical trials but rather to complement them by assessing dynamic psychological changes, contextual factors, and potentially aspects related to the therapeutic alliance (Myin-Germeys et al., 2018). This is particularly relevant because recent research has shown that the same psychological processes assessed in participants' everyday lives through ESM and in retrospective questionnaires do not always correlate, indicating that the two approaches do not capture the same process (or the same aspects of this given process) through ESM and traditional questionnaires (Koval et al., 2023; Solhan et al., 2009). One provides a momentary snapshot of the mechanisms at play in a patient's daily functioning, the other reflects the patient's overall perception of their functioning.

Ambulatory assessment (including ESM) may take various forms – from daily diaries, usually completed once in the evening, through real-time assessments administered multiple times per day, to multimodal ambulatory assessment including for example self-reports with psychophysiological measures. The choice of the precise method should be informed by the theoretical model and characteristics of the assessed process or symptom (Löchner et al., 2025; Myin-Germeys and Kuppens, 2022). While subjective sleep quality is often assessed only once a day after waking, challenges in regulating or expressing anger may fluctuate numerous times throughout the day. Consequently, they require more frequent, real-time evaluation. Relying on retrospective assessment in such cases may introduce bias, as participants may already have reappraised their difficulties (Solhan et al., 2009). Process or symptom characteristics also lead to another methodological decision in ESM – the choice between event-contingent, signal-contingent, or hybrid assessment – a mixture of both (Löchner et al., 2025). In signal-contingent assessment, participants receive prompts several times a day. Usually, their daily activity window is divided into slots, and they randomly receive a signal within each slot. This method is suited for assessing processes or symptoms that unfold alongside everyday activities and may not be noticed without explicit prompting (for example, mind-wandering or affect). In the signal-contingent method, researchers relying on process features and theoretical background have a choice between asking patients how they are feeling or what they are doing "at the moment of receiving the signal", or "since the last signal". In event-contingent assessment, participants decide when to complete questionnaires. This method is appropriate for evaluating processes/symptoms that occur only under certain circumstances (for example, when feeling pain or entering an elevator). It is also applied when measured variables are brief and/or do not occur often during the day/week, so a randomly sent signal would be very unlikely to fall at the exact moment (e.g. panic symptoms). Event-contingent designs may also include micro-bursts, i.e. additional questions that appear with a delay after the first reporting (e.g. checking how patients cope with pain two hours after reporting an acute episode).

A significant challenge in ESM self-reports is finding a balance between participants' burden in providing the assessment, often multiple times a day, and the validity of the measures obtained (Eisele et al., 2025, 2023). It is recommended to use questionnaires validated specifically for ESM (e.g. Ciarrochi et al., 2022; Jover Martínez et al., 2024, 2025; Rosenkranz et al., 2020). Although single-item assessments of a given process or symptom often show good validity, they may still require additional validation (Dejonckheere et al., 2022; Schuurman and Hamaker, 2019).

THE RATIONALE FOR USING EXPERIENCE SAMPLING METHODS FOR EVALUATING PSYCHOTHERAPY EFFECTIVENESS

Using ESM to assess psychotherapy outcomes can minimise assessment error and retrospective biases that typically arise when patients are asked to recall how they felt over the past week(s) while completing traditional questionnaires. Collecting real-time data in participants' natural environments is particularly important for variables that could be easily affected by various environmental and interpersonal factors, such as mood or emotion regulation (Solhan et al., 2009; Trull and Ebner-Priemer, 2009). Moreover, the severity of psychological disorders and underlying processes* (e.g. rumination) tends to fluctuate significantly throughout the day (e.g. Chen et al., 2022; Wirz-Justice, 2008). Consequently, single-point assessments are unlikely to capture the actual intensity of these processes and may fail to detect temporal trends and changes over time (Scholten et al., 2025b). In this context, questionnaire measures are best understood as capturing relatively stable, trait-like tendencies in the use of processes such as rumination across time and situations (Aldao et al., 2010), whereas ESM enables examination of how individuals actually select and implement these processes in daily life. For example, the combined use of questionnaire and ESM measures to assess a single psychological process in the context of intervention research is illustrated by Barnes et al. (2025). An excellent example justifying the use of this dual approach (traditional questionnaires vs. ESM), or the exclusive choice of just one based on the study's purpose, is the assessment of metacognitive processes (e.g. beliefs about emotions). The choice depends on whether the aim is to capture a generalised tendency, such as stable negative convictions about negative emotions (best measured with retrospective questionnaires; e.g. Becerra et al., 2020), or to measure how those metacognitions change depending on context and verify their real-time dynamics, such as instability or entropy (best captured through ESM; e.g. Bailen et al., 2019; Kneeland et al., 2020). Moreover, some researchers have recently extended this line of inquiry by showing that self-report questionnaires, weekly assessments, and ESM may each capture different aspects of symptoms and processes in the context of psychological disorders (Tamm et al., 2024). Dejonckheere et al. (2019) demonstrated that many indices of emotional dynamics substantially overlap and provide limited incremental validity compared to average affect. This suggests that researchers should carefully justify the selection of dynamic indices and prioritise those with clear theoretical relevance, rather than assuming added value by default.

Tamm et al. (2024) compared two approaches to evaluating psychotherapy effectiveness: EMA and weekly questionnaire-based assessments. They concluded that EMA was more sensitive in detecting changes, provided more accurate estimates, and improved the likelihood of detecting significant effects. A similar conclusion was reached

* In line with the literature on process-based cognitive behavioral therapy, the term "process" refers to transdiagnostic mechanisms such as rumination, experiential avoidance, or perfectionism. These mechanisms can be assessed both using questionnaires, which capture relatively stable tendencies, and with intensive longitudinal methods, which reveal their dynamic and context-dependent fluctuations (Harvey et al., 2004; Hofmann and Hayes, 2019).

by Moore et al. (2016), who examined the effectiveness of mindfulness-based stress reduction (MBSR) and found that EMA was more sensitive than traditional questionnaires in detecting changes in patient-reported outcomes. However, findings from clinical populations suggest that the added value of emotional dynamics may vary across contexts. Sperry et al. (2020) showed that fluctuations in negative affect predicted the course of depressive and bipolar spectrum disorders, whereas Houben and Kuppens (2020) reported limited predictive power of emotion dynamics beyond mean affect. Similarly, in individuals living with HIV, Pięta-Lendzion et al. (2024) found that daily emotional inertia and innovation were stable over time but unrelated to long-term changes in posttraumatic growth or depreciation. Longitudinal studies involving multiple assessments collected via ESM allow researchers to capture mental disorders as complex, dynamic networks of symptoms and psychological processes (e.g. Borsboom and Cramer, 2013; Bringmann, 2021; Scholten et al., 2025b). This is particularly relevant for one of the key elements in establishing evidence-based interventions and their efficacy: understanding the mechanism of change (Bastiaansen et al., 2020; Rief et al., 2024). In this framework, network models consist of nodes representing observed symptoms and psychological processes, and edges indicating associations between them (Epskamp et al., 2018). Mental disorders are thus conceptualised as pathological states emerging from systems of interrelated variables (e.g. Contreras et al., 2019; Robinaugh et al., 2020a; Verhagen et al., 2016). Applying intensive daily measurements and a network perspective to psychotherapy outcome research enables the assessment of intervention effects not only on individual symptoms but also on the relationships between them (Schumacher et al., 2024). Such changes are viewed more broadly, not merely as increases or decreases in the intensity of single variables (e.g. depression scale scores), but also as modifications in the dynamic interconnections between measurable psychological processes and symptoms. When the research question focuses on how treatment affects the structure of symptom interrelations, symptom networks can be compared between groups exposed to different therapeutic conditions (Schumacher et al., 2024).

Furthermore, repeated self-assessments via ESM can support patients' understanding of their own patterns during feedback sessions (Scholten et al., 2025a). Familiarity with the data helps to demystify the treatment process, making it easier to relate therapeutic insights to everyday experiences. As a result, ESM may enhance patient engagement, strengthen the sense of collaboration, and promote empowerment within the therapeutic process (Simons et al., 2015).

KEY EXAMPLES OF STUDY DESIGNS FOR USING ESM IN THERAPY EFFICACY AND EFFECTIVENESS ASSESSMENT

Randomised controlled trials (RCTs) have been criticised for several reasons (Rief et al., 2024; Tolin et al., 2015). Although this is not inherent to the RCT design itself (unlike randomisation and the use of control groups), in research on therapeutic interventions, RCT primarily rely on self-reported pre-post and follow-up measures, which require retrospective evaluation of symptoms. Such evaluations are often far removed from ecological settings and are unable to capture the dynamic psychological processes which change during an intervention. Additionally, existing guidelines are predominantly built on group-based, nomothetic research. It is crucial to note that this criticism is not directed at the RCT design itself, but rather the methodological choices regarding how outcomes and/or psychological processes are measured. While RCTs remain the gold standard (and currently, there is no viable alternative), scientific and clinical communities constantly work on improving the methodology for assessing treatment effectiveness. This is particularly important now, given the multitude of intervention types available (e.g. face-to-face therapy, internet-based interventions). There is no doubt that all interventions offered to patients should demonstrate evidence-based efficacy (Rief et al., 2024).

One solution to these limitations is the integration of ESM within RCTs. ESM can be applied at several stages – for establishing a baseline and comparing it with post-intervention ESM assessments to assess efficacy through a dynamic approach, or for fully capitalising on ESM's possibilities by using it during therapy not only to capture dynamic processes of change but also to provide participants with its benefits (e.g. empowerment; Simons et al., 2015) and to evaluate therapeutic relationships (Thomas et al., 2014). However, even in RCTs using ESM it is challenging to disentangle the specific impact of intervention components. Currently, RCTs often use identical network nodes across all participants, overlooking ESM's potential for personalised measurement (Schumacher et al., 2024).

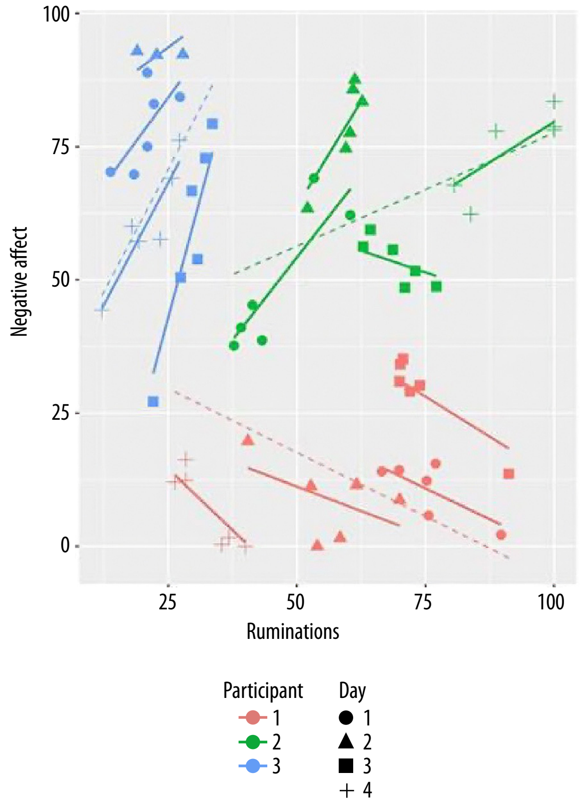
An alternative approach, particularly well-suited for ESM, is the Sequential Multiple Assignment Randomised Trial (SMART; Kidwell and Almirall, 2023), designed to assess adaptive interventions (Almirall et al., 2014). SMART is a multi-stage, factorial, randomised design in which some or all participants are randomised at two or more decision points. Treatment options – and whether a participant is re-randomised – depend on their response to previous treatment components. SMART designs are used to construct personalised sequences of treatments based on a patient's evolving needs and responses. This requires frequent monitoring of patient psychological processes and outcomes, making ESM an excellent solution fit for SMART designs. However, SMARTs do not provide overall evidence for therapy effectiveness, yet they can be very informative of its efficacy by exploring mechanisms of change and contribute significantly to understanding and improving psychotherapy (Kidwell and Almirall, 2023). Nevertheless, in this approach, a crucial decision is whether, from a methodological point of view, it is possible to use personalised assessment for each patient (Green, 2016).

DOI: 10.15557/PiPK.2026.0018

THE STATISTICAL CHALLENGES OF USING ESM IN ASSESSING THERAPY EFFECTIVENESS (DATA STRUCTURE AND ANALYTICAL APPROACH)

The full value of ESM data can be realised only when appropriate analytical approaches are used, such as multilevel models (Garson, 2013; Myin-Germeys and Kuppens, 2026; Nezlek, 2011; Woltman et al., 2012) or network models (Hevey, 2018). Unlike cross-sectional data, where observations are assumed to be independent, time series data in ESM are collected sequentially from multiple participants across numerous time points, often spanning multiple days with repeated daily measurements. As a result, ESM data inherently combine cross-sectional and longitudinal structures. This creates a nested hierarchy – i.e. single observations (for example, responses to questions assessing mood at a given signal) are nested within participants. It is not unusual that the data structure includes three levels due the frequency of sampling based on variable characteristics. For example, momentary assessment of mood five times per day may be nested within daily observations (e.g. sleep quality assessment), which are in turn nested within participants. If trait-level characteristics are taken into account in the analysis, they are incorporated into the model at the participant level. This data structure must be considered analytically. For example, a correlation observed at the momentary level may disappear at the individual or daily level. One clear example of this is the assessment of the emotional consequences of avoidance – it may decrease momentary negative affect, but has deleterious longer-term consequences and may be linked, for example, to increased depressive symptoms at the daily level (e.g. Skorupski et al., 2026). The potential for a single psychological mechanism to yield different outcomes at momentary, daily, or long-term levels highlights the crucial need to carefully consider both sampling frequency and the hierarchy of analytical levels. Multilevel modelling (MM) (Myin-Germeys and Kuppens, 2026; Woltman et al., 2012) – also known as hierarchical linear or linear mixed models (Garson, 2013) – accommodates such hierarchies. These models allow intercepts or slopes to vary across data groupings (e.g. by participant or day), while pooling information across units. As shown in Fig. 1, patterns observed on a single day are influenced by the overall structure across other days, stabilising estimates when data are sparse. From the perspective of psychotherapy effectiveness and efficacy, the most common approach involves comparing pre-post treatment evaluation and/or group differences by adding moderators (e.g. time/group) at the between-person level in multilevel models (e.g. Asselmann et al., 2023; Eddington et al., 2017).

In cases when the complex interplay between measured variables and symptoms needs to be assessed, network analysis (NA) can be used to estimate interrelations



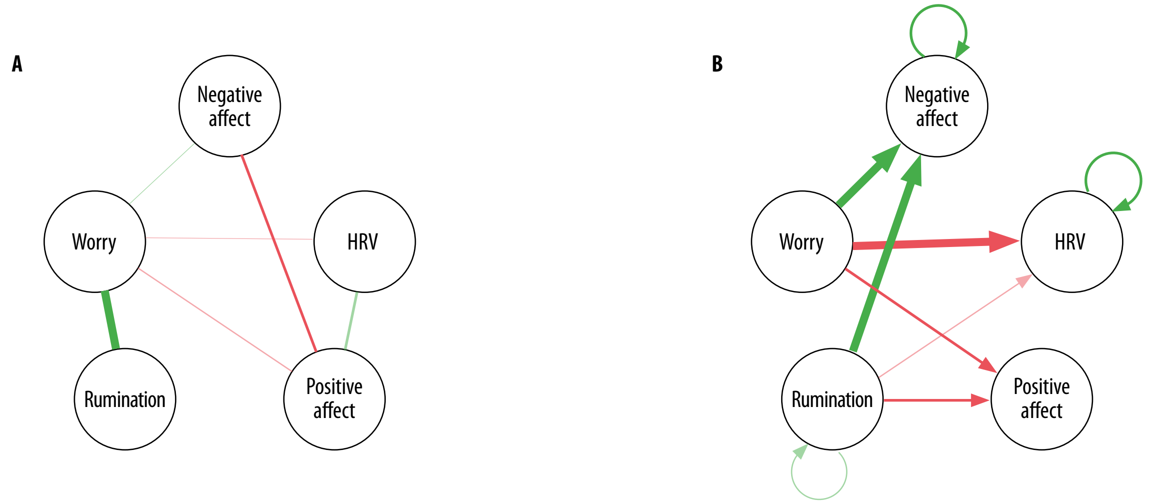
Note. Solid lines represent model predictions for individual days; dashed lines show estimated participant-level slopes across days. The model pools data across days to improve estimates, especially when data are limited (e.g. participant 3, day 2).

Fig. 1. Graphical visualisation of a multilevel model on simulated data illustrating the relation between rumination and mood in patients' daily life

within a system and assess variable centrality (Hevey, 2018; McNally, 2021). NA constructs networks in which nodes represent variables (i.e. symptoms or psychological processes) and edges indicate conditional relationships (e.g. partial correlations, regression weights) (see Fig. 2). For time-series data, networks can incorporate lagged relations, producing directed graphs indicating how one process influences another at the next time point. Centrality metrics quantify the relative importance of each variable, and bootstrapping procedures evaluate network stability and test selected edges or centralities for statistical significance (Castro et al., 2024; Epskamp et al., 2018).

Temporal dynamics can be analysed explicitly by regressing variables at time t on values at $t–i$ (Shojaie and Fox, 2022). This is achieved by adding lagged variables as predictors and can be integrated into MM or NA. Although causal inference must be cautious, such methods allow hypothesis testing for directed effects. In NA, this produces two networks: a contemporaneous (undirected) network showing associations independent of temporal influence, and a temporal (directed) network reflecting lagged dependencies. In psychotherapy research, the network approach is still

© PSYCHIATR PSYCHOL KLIN 2026, 26 (2), 155–166 DOI: 10.15557/PiPK.2026.0018



Note. The figure shows the contemporaneous network (**A**) and the temporal network (**B**) for a simulated dataset. Red and green edges represent negative and positive relations, respectively. The width of each line is proportional to edge weight. **HRV** – heart rate variability.

Fig. 2. Networks calculated for simulated data

being actively explored. The two most prevalent strategies are to compare network dynamics pre-post or during interventions across groups (e.g. Snippe et al., 2017) and to compare the centrality of nodes (symptoms/processes; e.g. Meine et al., 2024). Centrality indicators are also explored as potential markers of psychotherapy dynamics and dropout (e.g. Lutz et al., 2018) from the perspective of psychotherapy effectiveness.

Despite their advantages, ESM-based data analysis methods also present specific challenges. One arises even before data collection: determining sample size through a priori power analysis is difficult (Fritz et al., 2024). For MM, this requires assumptions about effect sizes across multiple levels and participant compliance, which varies widely (for practical recommendations on how to conduct power analysis for MM, see Lafit, 2022; Snijders, 2005). Beyond power analysis, missing data can be particularly problematic for lagged analyses, where a missing value at t_0 disrupts both $t-1$ and $t+1$ associations. Even with high total measurement counts, unevenly distributed missing prompts can eliminate large portions of usable lagged data. For instance, 40% missingness across 100 prompts may remove over 80% of potential lagged observations if consecutive measurements are interrupted. Unlike in non-lagged models, this issue can only be partially mitigated by establishing a minimum compliance rate for each participant.

While powerful, NA has limitations when applied to psychological data. A directed network with n nodes involves $n(n-1)$ edges – meaning 90 parameters for 10 variables. Large samples are required for stable estimation, and statistical testing is typically restricted to preselected edges or centrality metrics (Epskamp et al., 2018). Larger models often become impractical, forcing researchers to analyse smaller subsets of variables. This process is further complicated by the lack of reliable power analysis. Although methods exist (Constantin et al., 2026), network power depends heavily on structural features (e.g. density, clustering), limiting the precision of sample-size estimates. Furthermore, the interpretation of centrality can be problematic, as the assumptions underlying these measures do not always align with psychological theory (Bringmann et al., 2019). A recent review underlines that current research on therapy effectiveness does not take full advantage of ESM data, particularly when analysed using network approaches, for example by restricting the analysis to cross-sectional (undirected) networks, even though the longitudinal structure of the data enables the exploration of directed networks taking into account temporal dynamic in the interplay between processes and symptoms (Schumacher et al., 2024).

Finally, the network approach enables also to use an idiographic perspective by allowing the creation of a personalised network for a given participant and further assessment of changes in network structure reflecting the dynamics of the psychotherapeutic process (Hofmann and Hayes, 2019). However, researchers must decide whether to apply relatively straightforward comparisons of the same networks (i.e. measuring the same processes and symptoms) across groups or use personalised networks, making between-group comparisons more challenging. The latter approach involves estimating networks on small datasets from one participant. Although researchers are constantly working on idiographic networks, it remains challenging and is still extremely used only rarely (Mansueto et al., 2023; McNally, 2021). Nevertheless, it seems that once these challenges are addressed, the network approach may be the future of assessing psychotherapy effectiveness, particularly within a process-based framework and for patients with high comorbidity (Hofmann, 2025).

DOI: 10.15557/PiPK.2026.0018

THE COMPLIANCE, USER EXPERIENCE AND ETHICAL CHALLENGES OF USING ESM IN ASSESSING THERAPY EFFECTIVENESS

ESM-based assessments appear simple and convenient for patients, who are already accustomed to using their smartphones and mobile applications in everyday life (Verhagen et al., 2016). However, certain challenges associated with this research approach persist. The most relevant include maintaining compliance, the requirement to monitor and support use; minimising the burden imposed on participants by repeated daily assessments; and providing effective technical and onboarding support to minimise frustration, reduce the risk of potential data loss (Doherty et al., 2020; Trull and Ebner-Priemer, 2009), and limit the potential for attrition (Santangelo et al., 2013).

Decisions regarding the sampling scheme and study duration are critical, as they directly affect the quality and interpretability of ESM data. As Fritz et al. (2024) emphasise, the context in which participants provide their responses (e.g. holidays, stressful events) can strongly influence assessments and must be considered during study design. To address these challenges, researchers are encouraged to align sampling schemes with the construct and its expected fluctuations and may employ short, intensive periods of measurement and/or control for contextual factors to capture processes more accurately.

The most effective strategy to create a user-friendly, accessible, and actionable ESM solution is a participatory design approach, which actively involves future users in the development process. Evidence from participatory design studies (e.g. Kopeć et al., 2021, 2018, 2019; Kornacka et al., 2023) shows that this approach allows tailoring the solution to user needs while significantly reducing the time and cost of prototyping and testing.

Design complications also emerge in the use of mobile technologies for ESM in relation to maintaining user-engagement. Assessments and associated reminders must be effectively implemented in a way that provides satisfactory flexibility for users (to adjust the schedule to their needs and daily life) (Edney et al., 2023), while adhering to a frequency that maintains methodological rigor. Although reminders may serve as an effective tool to minimise the potential for data loss, they must be delivered without becoming intrusive or frustrating. Additionally, push notifications must be delivered at times and in context that allow for the discreet use of mental health tools, where preferred by users. While some studies suggest that gamification elements may improve motivation and help reduce attrition in digital mental health interventions (Litvin et al., 2020), potentially improving the rate and quality of responses obtained (van Berkel et al., 2017), their implications for treatment outcomes (e.g. whether gamification may distract from intervention content) are as yet unclear (Kleiman et al., 2025).

Although early ESM research raised many doubts regarding the use of ESM within specific clinical populations, the current state of the art suggests that it might be a method suitable for nearly all clinical populations (Naslund et al., 2016; Vachon et al., 2019; Verhagen et al., 2016) including patients with psychotic disorders (Bogudzińska et al., 2024) and in different contexts (e.g. psychopharmacological treatment; Bos et al., 2015). Nevertheless, researchers underline the non-random nature of dropout in ESM studies and the need for systematic monitoring of dropout predictors (Rintala et al., 2020, 2019).

Digitally delivered ESM protocols must also overcome potential barriers to acceptability, such as privacy concerns related to the collection and storage of sensitive data (Trull and Ebner-Priemer, 2013). While this issue is more broadly observed in digital health tools, it is particularly salient in ESM systems, which may collect extensive amounts of personal information, particularly in cases of multimodal assessments where self-report data may be combined with physiological measures. In some cases, combinations of passively acquired data may also be challenging to fully anonymise (e.g. GPS location). In one study, 79% of participants' home locations were identifiable (Brownstein et al., 2006), highlighting the potential for significant privacy violations and justifiable privacy concerns among users.

Beyond the challenges of implementation and engagement, further design considerations relate to the broader effects of regular self-assessment on habituation, reactivity, and honest disclosure. Ongoing engagement in self-assessment may be subject to social desirability effects and awareness of being measured (König et al., 2022), which presents not only a methodological challenge, but also a potential design limitation (Doherty et al., 2020). Although some studies have also underlined positive effects of patients using ESM on a daily basis (e.g. Simons et al., 2015; van Os et al., 2017), this should be carefully taken into account when including ESM into the evaluation of psychotherapy effectiveness, as other studies suggest potentially deleterious effects (e.g. habituation and decreased emotional awareness over time, Eisele et al., 2023; Widdershoven et al., 2019). Finally, relatively little is known how ESM use may affect the therapeutic relationship, but on the other hand, this method can also be an effective tool for monitoring therapeutic dynamics (Thomas et al., 2014).

Moreover, it is crucial to consider the impact of the ESM procedure itself on patients' emotion regulation, behaviour, and cognition. This reactivity has been explored primarily in the context of addiction (e.g. Rowan et al., 2007). While one might expect frequent self-monitoring to increase focus on the behaviour and thus amplify it, some studies using EMA in addiction research have surprisingly found that intensive daily evaluation might actually reduce addictive behaviours (McCarthy et al., 2015). Husky et al. (2014) found that endorsement of suicidal thoughts was not reactive to study duration or repeated measurements, even in high-risk groups. At the same time, repeated self-assessment may foster beneficial changes, such as increased emotional awareness or improved differentiation of affective states

 DOI: 10.15557/PiPK.2026.0018

(e.g. Eisele et al., 2023; Hoemann et al., 2021). Nevertheless, more dedicated research on ESM reactivity is needed within the context of psychotherapy and other psychological interventions. This need is further underscored by the fact that the potential impact of ESM on patients' behaviour is not merely a methodological concern, but has also been harnessed therapeutically, as demonstrated by the use of self-monitoring in ecological momentary interventions (EMI) (Balaskas et al., 2021; Kramer et al., 2014).

Recently, researchers have also explored how the measurement of user experience (UX) in complex digital interventions may benefit from ESM frameworks (e.g. Lukka et al., 2024). Such approaches aim to reduce the burden of retrospective recollection in UX examinations, while also addressing some of the challenges associated with obtaining meaningful insights into user experience by considering it within the immediate context of technology use (Doherty and Doherty, 2018). ESM may also assist with more rapid identification of digital health literacy deficits, which can negatively impact engagement with digital health tools (Borghouts et al., 2021). When assessing therapeutic intervention, an additional challenge is to take the practitioner's perspective into account (Weermeijer et al., 2024).

Apart from safety, the impact of ESM on patients, and the need to adapt the burden of ESM to participants' resources, there are three other key ethical considerations. First, given the considerable burden that ESM assessments can place on participants, it is imperative to fully leverage the potential of the large datasets collected through these methods to better understand and improve psychotherapy (Bastiaansen et al., 2020; Schumacher et al., 2024). Second, adherence to the principles of Open Science is essential (Löchner et al., 2025; Tuerlinckx et al., 2026). Most resources, ranging from those concerning study design and methodology to data and codes for analysing them, are freely available (a good example is open access to the REAL EMA book (Myin-Germeys and Kuppens, 2022), to databases of ESM items such as the ESM Item Repository, where the whole community of ESM is invited to contribute (Kirtley et al., 2020), or working together on developing the quality of ESM evaluation (Eisele et al., 2025). Third, the research process should involve co-designing and actively engaging experts by experience across all study phases, including methodology development, data analysis, and result dissemination (Eisele et al., 2025; Green et al., 2020; Smyth et al., 2021). Using user-centred design might overcome some challenges linked to acceptability and compliance in studies using ESM (Srinivas et al., 2018).

THE PROMISE OF PASSIVE ASSESSMENT AND WEARABLES

One of the solutions to address participants' burden and complete the self-reported measures in ESM is through passive data collection and wearables (Knight and Bidargaddi, 2018; Mehl et al., 2023). Mobile devices (e.g. smartphones and smartwatches) continuously collect information about momentary activity (e.g. accelerometer, sleep parameters, heart rate – HR), and these data may also serve as relevant indicators for psychological research and markers of therapeutic change (Mehl et al., 2023).

However, using passive data (sometimes referred to as parallel or multimodal data) also poses several challenges, particularly ethical (Behnke et al., 2024) and methodological (De Calheiros Velozo et al., 2024; Schoenmakers et al., 2025). First, researchers must clearly explain what types of data will be collected and for what purpose. Second, usually for psychophysiological measures, commercial devices not designed for research often do not provide access to the raw data needed for analysis, while research-grade devices are more expensive. Another important challenge is sensitive biological data protection. Finally, for many passive indicators, consensus is still lacking on how to use them within experience sampling – for example, what time window should be applied when merging these data with self-reports, and how to distinguish activity driven by psychological processes or symptoms from other forms of activity (Verkuil et al., 2016). Furthermore, a consensus remains elusive regarding which specific psychological processes are reflected by various psychophysiological measures. Consequently, current insights into the use of wearables for evaluating psychotherapy efficacy and effectiveness should not be viewed as a fully established method (Adams et al., 2017). Nevertheless, the number of accessible wearables and their use in clinical research continue to grow (for a review, see Hilty et al., 2021; Miyakoshi and Ito, 2024).

CONCLUSIONS AND RECOMMENDATIONS

A recent study based on three meta-analyses clearly indicates the urgent need to improve mental healthcare treatment (Hofmann et al., 2025). This goal can be achieved through several avenues, including: (1) a better understanding of the mechanism of change in psychotherapy; (2) improving methods for verifying psychotherapy effectiveness; and (3) adapting interventions and their evaluation to patients' individual needs. All three areas can be enhanced through the use of ESM in evaluating psychotherapy effectiveness.

ESM provides a powerful means to understand change mechanisms and improve effectiveness tracking by enabling dynamic monitoring of psychological processes, symptoms, and their interplay. It also reduces retrospective biases, increases ecological validity, and incorporates contextual factors into evaluation. Additionally, employing a network-based approach can facilitate individually tailored assessments, which is vital for patients with comorbidity and for applying process-based interventions. Crucially, however, these benefits are realised only when several factors are meticulously considered and appropriately addressed. Notably, while using ESM, researchers often fail to fully exploit its potential and in some cases do not use parts of the collected dataset at all (Bastiaansen et al., 2020).

DOI: 10.15557/PiPK.2026.0018 © PSYCHIATR PSYCHOL KLIN 2026, 26 (2), 155–166

In 2025, twenty six experts in the field summarised key recommendations for using experience sampling methods in mental health research in general (Löchner et al., 2025). Some of these recommendations are crucial for assessing psychotherapy efficacy and effectiveness. Below is a non-exhaustive list based on Löchner et al. (2025) and the above literature review:

- **Adapt study design:** tailor the study design, sampling scheme, and duration to the specific research question, ensuring it best matches the dynamics of the evaluated processes or symptoms (e.g. Myin-Germeys and Kuppens, 2022).
- **Evaluate psychometric properties:** assess the psychometric properties (e.g. within-person consistency) of ESM items (e.g. Eisele et al., 2025; Neubauer et al., 2020) also taking into account participants' burden vs. validity of assessment.
- **Consider biomarkers and passive data:** explore the use of biomarkers, objective measures, and passive data collection (e.g. Mehl et al., 2023).
- **Optimise compliance and monitor dropout:** implement strategies to optimise participant compliance, taking into account the burden associated with ESM assessment and the potential impact of ESM on participants' behaviour and emotion regulation. Account for the possibility of non-random dropout and monitor its predictors.
- **Prioritise ethics and UX:** place a high priority on ethical considerations and user experience throughout study design and implementation.
- **Enhance involvement:** consider expert-by-experience involvement and integrate practitioners' perspectives into the research process.
- **Plan data management and embrace Open Science:** develop a clear data management plan, for example, by adhering to FAIR (findable, accessible, interoperable, and reusable) principles (Wilkinson et al., 2016). Follow Open Science practices by freely sharing study designs, methodology, data, and analytical code, fostering community collaboration and transparency in ESM research (e.g. Tuerlinckx et al., 2026).
- **Utilise adequate data analysis:** employ appropriate data analysis for answering research questions (e.g. evaluating effectiveness vs. mechanism of change), taking into account data structure (e.g. Borsboom and Cramer, 2013; McNally, 2016; Nezlek, 2011).
- **Maximise data potential:** ensure that the large datasets generated through ESM are fully utilised, rather than leaving parts of the data unused (e.g. Bastiaansen et al., 2020).

Conflict of interest

The authors do not report any financial or personal connections with other persons or organisations which might negatively affect the content of this publication and/or claim authorship rights to this publication.

Author contribution

Original concept of study: MK. Collection, recording and/or compilation of data; analysis and interpretation of data; final approval of manuscript: MK, SB, JS, MSS, AJ, MP, MS. Writing of manuscript: MK, SB, JS, MSS, AJ, MS. Critical review of manuscript: MP.

References

Adams ZW, McClure EA, Gray KM et al.: Mobile devices for the remote acquisition of physiological and behavioral biomarkers in psychiatric clinical research. J Psychiatr Res 2017; 85: 1–14.

Aldao A, Nolen-Hoeksema S, Schweizer S: Emotion-regulation strategies across psychopathology: a meta-analytic review. Clin Psychol Rev 2010; 30: 217–237.

Almirall D, Nahum-Shani I, Sherwood NE et al.: Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research. Transl Behav Med 2014; 4: 260–274.

Asselmann E, Zenker M, Rückert F et al.: Ecological momentary assessment and applied relaxation: results of a randomized indicated preventive trial in individuals at increased risk for mental disorders. PLoS One 2023; 18: e0286750.

Bailen NH, Wu H, Thompson RJ: Meta-emotions in daily life: associations with emotional awareness and depression. Emotion 2019; 19: 776–787.

Balaskas A, Schueller SM, Cox AL et al.: Ecological momentary interventions for mental health: a scoping review. PLoS One 2021; 16: e0248152.

Barnes S, Szastok M, Para M et al.: A mobile ecological momentary intervention for reducing experiential avoidance in the context of rumination: protocol for a randomized controlled trial. JMIR Res Protoc 2025; 14: e66067.

Bastiaansen JA, Kunkels YK, Blaauw FJ et al.: Time to get personal? The impact of researchers choices on the selection of treatment targets using the experience sampling methodology. J Psychosom Res 2020; 137: 110211.

Becerra R, Preece DA, Gross JJ: Assessing beliefs about emotions: development and validation of the Emotion Beliefs Questionnaire. PLoS One 2020; 15: e0231395.

Behnke M, Saganowski S, Kunc D et al: Ethical considerations and checklist for affective research with wearables. IEEE Trans Affect Comput 2024; 15: 50–62.

van Berkel N, Goncalves J, Hosio S et al.: Gamification of mobile experience sampling improves data quality and quantity. Proc ACM Interact Mob Wearable Ubiquitous Technol 2017; 1: 107.

Bogudzińska B, Jaworski A, Zajdel A et al.: The experience sampling methodology in psychosis risk states: a systematic review. J Psychiatr Res 2024; 175: 34–41.

Borghouts J, Eikey E, Mark G et al.: Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021; 23: e24387.

Borsboom D, Cramer AOJ: Network analysis: an integrative approach to the structure of psychopathology. Annu Rev Clin Psychol 2013; 9: 91–121.

Bos FM, Schoevers RA, aan het Rot M: Experience sampling and ecological momentary assessment studies in psychopharmacology: a systematic review. Eur Neuropsychopharmacol 2015; 25: 1853–1864.

Bringmann LF: Person-specific networks in psychopathology: past, present, and future. Curr Opin Psychol 2021; 41: 59–64.

Bringmann LF, Elmer T, Epskamp S et al.: What do centrality measures measure in psychological networks? J Abnorm Psychol 2019; 128: 892–903.

Brownstein JS, Cassa CA, Mandl KD: No place to hide – reverse identification of patients from published maps. N Engl J Med 2006; 355: 1741–1742.

Castro D, Gysi D, Ferreira F et al.: Centrality measures in psychological networks: a simulation study on identifying effective treatment targets. PLoS One 2024; 19: e0297058.
Chen Z, Zhao S, Tian S et al.: Diurnal mood variation symptoms in major depressive disorder associated with evening chronotype: evidence from a neuroimaging study. J Affect Disord 2022; 298: 151–159.
Ciarrochi J, Sahdra B, Hofmann SG et al.: Developing an item pool to assess processes of change in psychological interventions: the Process-Based Assessment Tool (PBAT). J Context Behav Sci 2022; 23: 200–213.
Constantin MA, Schuurman NK, Vermunt JK: A general Monte Carlo method for sample size analysis in the context of network models. Psychol Methods 2026; 31: 385–405.
Contreras A, Nieto I, Valiente C et al.: The study of psychopathology from the network analysis perspective: a systematic review. Psychother Psychosom 2019; 88: 71–83.
De Calheiros Velozo J, Vaessen T, Claes S et al.: Investigating adverse daily life effects following a psychosocial laboratory stress task, and the moderating role of psychopathology. Stress 2024; 27: 2380403.
Dejonckheere E, Demeyer F, Geusens B et al.: Assessing the reliability of single-item momentary affective measurements in experience sampling. Psychol Assess 2022; 34: 1138–1154.
Dejonckheere E, Mestdagh M, Houben M et al.: Complex affect dynamics add limited information to the prediction of psychological well-being. Nat Hum Behav 2019; 3: 478–491.
Doherty K, Doherty G: The construal of experience in HCI: Understanding self-reports. Int J Hum Comput Stud 2018; 110: 63–74.
Doherty K, Balaskas A, Doherty G: The design of ecological momentary assessment technologies. Interact Comput 2020; 32: 257–278.
Eddington KM, Burgin CJ, Silvia PJ et al.: The effects of psychotherapy for major depressive disorder on daily mood and functioning: a longitudinal experience sampling study. Cognit Ther Res 2017; 41: 266–277.
Edney S, Goh CM, Chua XH et al.: Evaluating the effects of rewards and schedule length on response rates to ecological momentary assessment surveys: randomized controlled trials. J Med Internet Res 2023; 25: e45764.
Eisele G, Hiekkaranta A, Kunkels YK et al.: ESM-Q: a consensus-based quality assessment tool for experience sampling method items. Behav Res Methods 2025; 57: 124.
Eisele G, Vachon H, Lafit G et al.: A mixed-method investigation into measurement reactivity to the experience sampling method: the role of sampling protocol and individual characteristics. Psychol Assess 2023; 35: 68–81.
Epskamp S, Borsboom D, Fried EI: Estimating psychological networks and their accuracy: a tutorial paper. Behav Res Methods 2018; 50: 195–212.
Fritz J, Piccirillo ML, Cohen ZD et al.: So you want to do ESM? 10 essential topics for implementing the experience-sampling method. Adv Methods Pract Psychol Sci 2024; 7: 25152459241267912.
Garson GD: Hierarchical Linear Modeling: Guide and Applications. SAGE Publications, Thousand Oaks 2013.
Green D: Making the case for using personalised outcome measures to track progress in psychotherapy. Eur J Psychother Couns 2016; 18: 39–57.
Green T, Bonner A, Teleni L et al.: Use and reporting of experience-based codesign studies in the healthcare setting: a systematic review. BMJ Qual Saf 2020; 29: 64–76.
Harvey A, Watkins E, Mansell W et al.: Cognitive Behavioural Processes across Psychological Disorders: A Transdiagnostic Approach to Research and Treatment. Oxford University Press, Oxford 2004.
Hayes SC, Hofmann SG, Ciarrochi J: A process-based approach to psychological diagnosis and treatment: the conceptual and treatment utility of an extended evolutionary meta model. Clin Psychol Rev 2020; 82: 101908.
Hevey D: Network analysis: a brief overview and tutorial. Health Psychol Behav Med 2018; 6: 301–328.
Hilty DM, Armstrong CM, Luxton DD et al.: A scoping review of sensors, wearables, and remote monitoring for behavioral health: uses, outcomes, clinical competencies, and research directions. J Technol Behav Sci 2021; 6: 278–313.
Hoemann K, Barrett LF, Quigley KS: Emotional granularity increases with intensive ambulatory assessment: methodological and individual factors influence how much. Front Psychol 2021; 12: 704125.
Hofmann SG: A network control theory of dynamic systems approach to personalize therapy. Behav Ther 2025; 56: 199–212.
Hofmann SG, Curtiss J: A complex network approach to clinical science. Eur J Clin Invest 2018; 48: e12986.
Hofmann SG, Hayes SC: The future of intervention science: process-based therapy. Clin Psychol Sci 2019; 7: 37–50.
Hofmann SG, Kasch C, Reis A: Effect sizes of randomized-controlled studies of cognitive behavioral therapy for anxiety disorders over the past 30 years. Clin Psychol Rev 2025; 117: 102553.
Houben M, Kuppens P: Emotion dynamics and the association with depressive features and borderline personality disorder traits: unique, specific, and prospective relationships. Clin Psychol Sci 2020; 8: 226–239.
Husky M, Olié E, Guillaume S et al.: Feasibility and validity of ecological momentary assessment in the investigation of suicide risk. Psychiatry Res 2014; 220: 564–570.
Jover Martínez A, Lemmens LHJM, Fried EI et al.: Developing a transdiagnostic ecological momentary assessment protocol for psychopathology. Int J Methods Psychiatr Res 2024; 33: e2028.
Jover Martínez A, Lemmens LHJM, Fried EI et al.: Validation of a transdiagnostic psychopathology ecological momentary assessment protocol in a university student sample. Psychol Assess 2025; 37: 46–61.
Kidwell KM, Almirall D: Sequential, multiple assignment, randomized trial designs. JAMA 2023; 329: 336–337.
Kirtley O, Hiekkaranta AP, Kunkels YK et al.: The Experience Sampling Method (ESM) Item Repository. OSF Preprints 2020. Available from: https://doi.org/10.17605/OSF.IO/KG376 [cited: 30 June 2025].
Kleiman EM, Glenn CR, Auad EC et al.: A simple way to gamify ecological momentary assessment studies and improve survey adherence with adolescents: The Emoji Game. Psychol Assess 2025; 37: 288–296.
Kneeland ET, Goodman FR, Dovidio JF: Emotion beliefs, emotion regulation, and emotional experiences in daily life. Behav Ther 2020; 51: 728–738.
Knight A, Bidargaddi N: Commonly available activity tracker apps and wearables as a mental health outcome indicator: a prospective observational cohort study among young adults with psychological distress. J Affect Disord 2018; 236: 31–36.
König LM, Allmeta A, Christlein N et al.: A systematic review and meta-analysis of studies of reactivity to digital in-the-moment measurement of health behaviour. Health Psychol Rev 2022; 16: 551–575.
Kopeć W, Biele C, Kornacka M et al.: Participatory design landscape for the human-machine collaboration, interaction and automation at the frontiers of HCI (PDL 2021). In: Ardito C, Lanzilotti R, Malizia A et al. (eds.): Human-Computer Interaction – INTERACT 2021. INTERACT 2021. Lecture Notes in Computer Science. Springer International Publishing, Cham 2021: 564–569.
Kopeć W, Nielek R, Wierzbicki A: Guidelines towards better participation of older adults in software development processes using a new SPIRAL method and participatory approach. In: Proceedings of the 11th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE '18). Association for Computing Machinery, New York 2018: 49–56.
Kopeć W, Wichrowski M, Kalinowski K et al.: VR with older adults: participatory design of a virtual ATM training simulation. IFAC-PapersOnLine 2019; 52: 277–281.
Kornacka M, Jaskulska A, Skorupska K et al.: A solution for creating dynamic networks of symptoms in cognitive-behavioral functional analysis and therapy – participatory design protocol. Ment Health Soc Incl 2023; 27: 167–176.
Koval P, Kalokerinos EK, Greenaway KH et al.: Emotion regulation in everyday life: Mapping global self-reports to daily processes. Emotion 2023; 23: 357–374.

Kramer I, Simons CJ, Hartmann JA et al.: A therapeutic application of the experience sampling method in the treatment of depression: a randomized controlled trial. World Psychiatry 2014; 13: 68–77.

Lafit G: Sample size selection in ESM studies. In: Myin-Germeys I, Kuppens P (eds.): The Open Handbook of Experience Sampling Methodology: A Step-by-Step Guide to Designing, Conducting, and Analyzing ESM Studies. 2nd ed., The center for Research on Experience sampling and Ambulatory methods Leuven (REAL), Leuven 2022: 217–237.

Litvin S, Saunders R, Maier MA et al.: Gamification as an approach to improve resilience and reduce attrition in mobile mental health interventions: a randomized controlled trial. PLoS One 2020; 15: e0237220.

Löchner J, Santangelo PS, Ansell E et al.: Ambulatory assessment in mental health research: expert consensus on current practices and future directions. OSF Preprints 2025. Available from: https://doi.org/10.31234/osf.io/mhb5g_v1.

Lukka L, Karhulahti VM, Bergman VR et al.: Measuring digital intervention user experience with a novel ecological momentary assessment (EMA) method, CORTO. Internet Interv 2024; 35: 100706.

Lutz W, Schwartz B, Hofmann SG et al.: Using network analysis for the prediction of treatment dropout in patients with mood and anxiety disorders: a methodological proof-of-concept study. Sci Rep 2018; 8: 7819.

Mansueto AC, Wiers RW, van Weert JCM et al.: Investigating the feasibility of idiographic network models. Psychol Methods 2023; 28: 1052–1068.

McCarthy DE, Minami H, Yeh VM et al.: An experimental investigation of reactivity to ecological momentary assessment frequency among adults trying to quit smoking. Addiction 2015; 110: 1549–1560.

McNally RJ: Can network analysis transform psychopathology? Behav Res Ther 2016; 86: 95–104.

McNally RJ: Network analysis of psychopathology: controversies and challenges. Annu Rev Clin Psychol 2021; 17: 31–53.

Mehl MR, Eid M, Wrzus C et al. (eds.): Mobile Sensing in Psychology: Methods and Applications. Guilford Press, New York 2023.

Meine LE, Müller-Bardorff M, Recher D et al.: Network analyses of ecological momentary emotion and avoidance assessments before and after cognitive behavioral therapy for anxiety disorders. J Anxiety Disord 2024; 106: 102914.

Mestdagh M, Verdonck S, Piot M et al.: m-Path: an easy-to-use and highly tailorable platform for ecological momentary assessment and intervention in behavioral research and clinical practice. Front Digit Health 2023; 5: 1182175.

Mink F, Lutz W, Hehlmann MI: Ecological momentary assessment in psychotherapy research: a systematic review. Clin Psychol Rev 2025; 117: 102565.

Miyakoshi T, Ito YM: Assessing the current utilization status of wearable devices in clinical research. Clin Trials 2024; 21: 470–482.

Moore RC, Depp CA, Wetherell JL et al.: Ecological momentary assessment versus standard assessment instruments for measuring mindfulness, depressed mood, and anxiety among older adults. J Psychiatr Res 2016; 75: 116–123.

Myin-Germeys I, Kuppens P (eds.): The Open Handbook of Experience Sampling Methodology: A Step-by-Step Guide to Designing, Conducting and Analysing ESM Studies. 3rd ed., Leuven University Press, Leuven 2026.

Myin-Germeys I, Kuppens P (eds.): The Open Handbook of Experience Sampling Methodology: A Step-by-Step Guide to Designing, Conducting, and Analyzing ESM Studies. 2nd ed., The center for Research on Experience sampling and Ambulatory methods Leuven (REAL), Leuven 2022.

Myin-Germeys I, Kasanova Z, Vaessen T et al.: Experience sampling methodology in mental health research: new insights and technical developments. World Psychiatry 2018; 17: 123–132.

Myin-Germeys I, Oorschot M, Collip D et al.: Experience sampling research in psychopathology: opening the black box of daily life. Psychol Med 2009; 39: 1533–1547.

Myin-Germeys I, Schick A, Ganslandt T et al.: The experience sampling methodology as a digital clinical tool for more person-centered mental health care: an implementation research agenda. Psychol Med 2024; 54: 2785–2793.

Naslund JA, Aschbrenner KA, Bartels SJ: Wearable devices and smartphones for activity tracking among people with serious mental illness. Ment Health Phys Act 2016; 10: 10–17.

Neubauer AB, Voelkle MC, Voss A et al.: Estimating reliability of within-person couplings in a multilevel framework. J Pers Assess 2020; 102: 10–21.

Nezlek JB: Diary Methods for Personality and Social Psychology. SAGE, Los Angeles 2012.

Nezlek JB: Multilevel Modeling for Social and Personality Psychology. SAGE, Los Angeles 2011.

van Os J, Verhagen S, Marsman A et al.: The experience sampling method as an mHealth tool to support self-monitoring, self-insight, and personalized health care in clinical practice. Depress Anxiety 2017; 34: 481–493.

Peterson CB, Engel SG, Crosby RD et al.: Comparing integrative cognitive-affective therapy and guided self-help cognitive-behavioral therapy to treat binge-eating disorder using standard and naturalistic momentary outcome measures: a randomized controlled trial. Int J Eat Disord 2020; 53: 1418–1427.

Pięta-Lendzion M, Rzeszutek M, Tsukayama E et al.: Daily emotional dynamics and changes in posttraumatic growth and posttraumatic depreciation among people living with HIV. J Psychosom Res 2024; 176: 111552.

Rief W, Asmundson GJG, Bryant RA et al.; PsyChange consortium: The future of psychological treatments: the Marburg Declaration. Clin Psychol Rev 2024; 110: 102417.

Rintala A, Wampers M, Myin-Germeys I et al.: Momentary predictors of compliance in studies using the experience sampling method. Psychiatry Res 2020; 286: 112896.

Rintala A, Wampers M, Myin-Germeys I et al.: Response compliance and predictors thereof in studies using the experience sampling method. Psychol Assess 2019; 31: 226–235.

Robinaugh DJ, Brown ML, Losiewicz OM et al.: Towards a precision psychiatry approach to anxiety disorders with ecological momentary assessment: the example of panic disorder. Gen Psychiatry 2020a; 33: e100161.

Robinaugh DJ, Hoekstra RHA, Toner ER et al.: The network approach to psychopathology: a review of the literature 2008–2018 and an agenda for future research. Psychol Med 2020b; 50: 353–366.

Rosenkranz T, Takano K, Watkins ER et al.: Assessing repetitive negative thinking in daily life: development of an ecological momentary assessment paradigm. PLoS One 2020; 15: e0231783.

Rowan PJ, Cofta-Woerpel L, Mazas CA et al.: Evaluating reactivity to ecological momentary assessment during smoking cessation. Exp Clin Psychopharmacol 2007; 15: 382–389.

Santangelo PS, Ebner-Priemer UW, Trull TJ: Experience sampling methods in clinical psychology. In: Comer JS, Kendall PC (eds.): The Oxford Handbook of Research Strategies for Clinical Psychology. Oxford University Press, Oxford 2013: 188–210.

Schoenmakers M, Saygin M, Sikora M et al.: Stress in action wearables database: a database of noninvasive wearable monitors with systematic technical, reliability, validity, and usability information. Behav Res Methods 2025; 57: 171.

Scholten S, Klintwall L, Glombiewski JA et al.: Updating patient perceptions with intensive longitudinal data for enhanced case conceptualizations: an approach with Bayesian informative priors. J Psychopathol Clin Sci 2025a; 134: 585–599.

Scholten S, Rubel JA, Glombiewski JA et al.: What time-varying network models based on functional analysis tell us about the course of a patient's problem. Psychother Res 2025b; 35: 637–655.

Schumacher L, Klein JP, Hautzinger M et al.: Predicting the outcome of psychotherapy for chronic depression by person-specific symptom networks. World Psychiatry 2024; 23: 411–420.

Schuurman NK, Hamaker EL: Measurement error and person-specific reliability in multilevel autoregressive modeling. Psychol Methods 2019; 24: 70–91.
Scollon CN, Kim-Prieto C, Diener E: Experience sampling: promises and pitfalls, strengths and weaknesses. J Happiness Stud 2003; 4: 5–34.
Shojaie A, Fox EB: Granger causality: a review and recent advances. Annu Rev Stat Appl 2022; 9: 289–319.
Simons CJP, Hartmann JA, Kramer I et al.: Effects of momentary self-monitoring on empowerment in a randomized controlled trial in patients with depression. Eur Psychiatry 2015; 30: 900–906.
Skorupski MS, Krejtz I, Barnes S et al.: Not all task-unrelated thoughts (TUT) are created equal – TUT characteristics as predictors of affective states and heart-rate variability. Sci Rep 2026; 16: 13292.
Smyth JM, Jones DR, Wen CKF et al.: Influence of ecological momentary assessment study design features on reported willingness to participate and perceptions of potential research studies: an experimental study. BMJ Open 2021; 11: e049154.
Snijders TAB: Power and sample size in multilevel linear models. In: Everitt BS, Howell DC (eds.): Encyclopedia of Statistics in Behavioral Science. Wiley, Chichester 2005; 3: 1570–1573.
Snippe E, Viechtbauer W, Geschwind N et al.: The impact of treatments for depression on the dynamic network structure of mental states: two randomized controlled trials. Sci Rep 2017; 7: 46523.
Solhan MB, Trull TJ, Jahng S et al.: Clinical assessment of affective instability: comparing EMA indices, questionnaire reports, and retrospective recall. Psychol Assess 2009; 21: 425–436.
Sperry SH, Walsh MA, Kwapil TR: Emotion dynamics concurrently and prospectively predict mood psychopathology. J Affect Disord 2020; 261: 67–75.
Srinivas P, Bodke K, Ofner S et al.: Context-sensitive ecological momentary assessment: application of user-centered design for improving user satisfaction and engagement during self-report. JMIR Mhealth Uhealth 2018; 7: e10894.
Tamm J, Takano K, Just L et al.: Ecological momentary assessment versus weekly questionnaire assessment of change in depression. Depress Anxiety 2024; 2024: 9191823.
Thomas KM, Hopwood CJ, Woody E et al.: Momentary assessment of interpersonal process in psychotherapy. J Couns Psychol 2014; 61: 1–14.
Tolin DF, McKay D, Forman EM et al.: Empirically supported treatment: recommendations for a new model. Clin Psychol Sci Pract 2015; 22: 317–338.
Trull TJ, Ebner-Priemer U: Ambulatory assessment. Annu Rev Clin Psychol 2013; 9: 151–176.
Trull TJ, Ebner-Priemer UW: Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: introduction to the special section. Psychol Assess 2009; 21: 457–462.
Tuerlinckx F, Kuppens P, Ariens S et al.: New developments in experience sampling methodology. Br J Math Stat Psychol 2026; 79: 46–65.
Vachon H, Viechtbauer W, Rintala A et al.: Compliance and retention with the experience sampling method over the continuum of severe mental disorders: meta-analysis and recommendations. J Med Internet Res 2019; 21: e14475.
Verhagen SJW, Hasmi L, Drukker M et al.: Use of the experience sampling method in the context of clinical trials. Evid Based Ment Health 2016; 19: 86–89.
Verkuil B, Brosschot JF, Tollenaar MS et al.: Prolonged non-metabolic heart rate variability reduction as a physiological marker of psychological stress in daily life. Ann Behav Med 2016; 50: 704–714.
Weermeijer J, Kiekens G, Wampers M et al.: Practitioner perspectives on the use of the experience sampling software in counseling and clinical psychology. Behav Inf Technol 2024; 43: 540–550.
Weermeijer JDM, Wampers M, de Thurah L et al.: Usability of the experience sampling method in specialized mental health care: pilot evaluation study. JMIR Form Res 2023; 7: e48821.
Widdershoven RLA, Wichers M, Kuppens P et al.: Effect of self-monitoring through experience sampling on emotion differentiation in depression. J Affect Disord 2019; 244: 71–77.
Wilkinson MD, Dumontier M, Aalbersberg IJJ et al.: The FAIR guiding principles for scientific data management and stewardship. Sci Data 2016; 3: 160018.
Wirz-Justice A: Diurnal variation of depressive symptoms. Dialogues Clin Neurosci 2008; 10: 337–343.
Woltman H, Feldstain A, MacKay JC et al.: An introduction to hierarchical linear modeling. Tutor Quant Methods Psychol 2012; 8: 52–69.