Use of the experience sampling method in the context of clinical trials

Simone J W Verhagen, Laila Hasmi, Marjan Drukker, J van Os, Philippe A E G Delespaul

Department of Psychiatry and Neuropsychology, Maastricht University, Maastricht, The Netherlands

Correspondence to Professor Philippe A EG Delespaul; Ph.delespaul@maastrichtuniversity.nl

ABSTRACT

Objective The experience sampling method (ESM) is a structured diary technique to appraise subjective experiences in daily life. It is applied in psychiatric patients, as well as in patients with somatic illness. Despite the potential of ESM assessment, the improved logistics and its increased administration in research, its use in clinical trials remains limited. This paper introduces ESM for clinical trials in psychiatry and beyond.

Methods ESM is an ecologically valid method that yields a comprehensive view of an individual’s daily life. It allows the assessment of various constructs (eg, quality of life, psychopathology) and psychological mechanisms (eg, stress-sensitivity, coping). These constructs are difficult to assess using cross-sectional questionnaires. ESM can be applied in treatment monitoring, as an ecological momentary intervention, in clinical trials, or in single case clinical trials. Technological advances (eg, smartphone applications) make its implementation easier.

Results Advantages of ESM are highlighted and disadvantages are discussed. Furthermore, the ecological nature of ESM data and its consequences are explored, including the potential pitfalls of ambiguously formulated research questions and the specificities of ESM in statistical analyses. The last section focuses on ESM in relation to clinical trials and discusses its future use in optimising clinical decision-making.

Conclusions ESM can be a valuable asset in clinical trial research and should be used more often to study the benefits of treatment in psychiatry and somatic health.

INTRODUCTION

In clinical trials, the effectiveness of a psychiatric intervention is often assessed retrospectively, by asking patients to complete self-report questionnaires. ESM can be applied in treatment monitoring, as an ecological momentary intervention, in clinical trials, or in single case clinical trials. Technological advances (eg, smartphone applications) make its implementation easier.

What is ESM?

ESM is an umbrella term for a family of momentary assessment techniques that use signals to trigger data collection in daily life. Alternative terms are ecological momentary assessment,6 ambulatory assessment,7 event sampling,8 beeper studies,9 the structured diary method,10 intensive longitudinal assessment11 and real-time data capture studies.3 ESM assesses sampled experiences and behaviours as well as moment-to-moment changes in mental states, embedded in normal daily life. It is an empirically validated structured diary technique. Typically, participants are asked to repeatedly complete short questionnaires (lasting no more than 2 min) in response to beep prompts. The questionnaires cover current mood, cognitions, perceptions, behaviours and descriptors of the momentary context (eg, location, company, activity).3,4,6 ESM is useful in psychiatry, as well as in patients with somatic illness. The method focuses on symptoms (ill health), as well as on adaptive functioning (well-being), aiming to map (normal) daily psychological functioning.15 Originally, paper diaries were used in combination with pagers or electronic wristwatches. As technology became more advanced, data collection logistics and reliability was improved by the use of personal digital assistants13 and smartphone applications14 (eg, the PsyMate, ExperienceSampler or apps by Invivo, lifeData). Typically, assessments are collected using an unpredictable (random) time sampling protocol. Assessments can also be triggered by an event (event sampling).8 Questionnaires are designed for quick and easy data collection and use open-ended questions, visual analogue scales, checklists or self-report Likert scales.4,13

A typical ESM data form

ESM assessments usually comprise a morning questionnaire, a beep questionnaire assessed repeatedly within a day and an evening questionnaire. The content of the items is subject to the assessment theme. Questions are short and can be rated quickly. An example of such an assessment form is included in table 1.

Why is ESM used in research?

ESM typically has a number of advantages, but its implementation varies.

▸ ESM has high ecological validity because assessments are made in the natural flow of real life.13,15

▸ ESM reduces memory strains and avoids aggregation because only the actual moment is assessed repeatedly over time. This increases accuracy and is comparatively easier.15

▸ The repeated assessments are collected in different situations (contextualised), which allows researchers to disentangle and understand the variability in mental states and psychological constructs.15
ESM yields a rich data set covering information on mental state, quality of life, mobility, social network and more, which reduces the need to use separate questionnaires measuring different constructs. Assessment error is reduced by repeated measures over time. This improves the validity, reliability, and transparency of individual pattern assessments, which is helpful in clinical practice. Sensitivity to detect change increases due to the collection of data at multiple time points. ESM makes it easier for patients during feedback sessions to acknowledge and translate the findings to daily life practice, because they have rated the items (eg, anxiety) repeatedly. This avoids mystification. Involvement and collaboration in care are improved and feelings of empowerment in the treatment process are heightened. A possible drawback of ESM is that the method is perceived as time-consuming and demanding. Ideally, assessments are kept as brief as possible, preferably <1 min, with a maximum of 2 min. A study protocol of 10 assessments for 6 consecutive days results in a cumulative time investment of <120 min. The set of questionnaires needed in cross-sectional studies easily exceeds this time investment. Selection bias is a second concern. Not all patients are willing to participate or comply with the protocol, and this subgroup possibly functions better than the non-responders. However, previous research has shown that the method is feasible in a wide variety of patients, including those suffering from severe mental illness. Most studies have reported good compliance. On the other hand, patients are inclined to miss assessments as a response to their current mood, which could also lead to selection bias. However, extensive analysis of missed responses demonstrate that beeps are predominantly missed at logical moments, such as the morning and evening (when participants are most likely sleeping), and lag analysis found no mood-related dropped responses. Third, ESM can induce reactivity. For instance, participants report increased awareness during ESM data collection. Monitoring is also used as a behavioural intervention, for instance, to reduce cigarette smoking. In this case, behavioural change is maximised by single target assessment of unwanted (automatic) behaviours. ESM, in contrast, is an open exploration of daily life using multiple targets. This open exploration reduces reactivity. Furthermore, repeated (exclusive) assessments of negative mood states can be confrontational and induce negative feelings. To avoid this, a careful and balanced construction of the questionnaire is needed; for example, changing the order of questions so that positive and negative momentary mood items are mixed. Finally, the complexity of ESM data analyses could form a drawback for both researchers and clinicians. Normal statistical methods, such as linear regression analysis and analysis of variance, cannot be used. Aggregation of data per participant removes the multilevel structure, allowing simple analyses. In addition, most statistical packages now have easily accessible multilevel regression tools. The present paper provides guidelines on how to analyse the data properly, allowing a reliable assessment of the richness of the ESM data (see below).
term at the participant level, in addition to the regular error term (at the assessment level) used in unilevel regression. Even when trends over time are not the primary interest, it is advisable to remove the potentially confounding effects of time in ESM data by adding a time variable (detrending). This is effective when values increase or decrease linearly over the time of the study. Furthermore, since adjacent assessments tend to be more similar to each other than to other assessments further apart in time (autocorrelation), ESM analyses also need more sophisticated adjustment methods. Care must be exercised when using such models, since observations are not evenly spaced due to randomisation and data collection usually proceeds over multiple days, resulting in a huge time lag for observations on consecutive days. Therefore, continuous-time structures are usually preferable, although their availability across different software packages differs. It is also highly recommended to allow random slopes in the model where applicable. This yields more realistic estimates of SEs, which otherwise might be underestimated. A random slope means that the regression coefficient is not fixed, but is allowed to vary between participants. This variation is realised by adding an extra error term for each predictor. In these models, a simple variance–covariance matrix of the random effects (eg, which assumes that all covariances are zero: the default in some statistical software packages) is not realistic. Therefore, an unstructured variance–covariance matrix, where all variances and covariances are allowed to be different, is preferable. However, this can lead to rather complex and time-consuming models. Sometimes, the statistical program may even fail to provide output. Alternatively, one can consider only using random slopes for the main independent variable(s). If the analysis includes multiple main independent variables, or if the analysis still does not yield valid results after reducing the number of random slopes, it may be necessary to remove essential random slopes. In that case, resampling methods (eg, permutation tests, bootstrapping) can be used to obtain valid tests and CIs.

**Clarifying ESM questions**

ESM allows the study of a broad range of research questions. The epidemiological finding that the mental health of people in urban environments is poor (a between-subject question) can be explored by assessing whether people in crowded environments feel worse (a within-subject question). ESM allows the exploration or ‘unpacking’ of relevant factors. The multilevel data structure of ESM can confuse researchers. Consider, for example, a study that explores the relation between depression and loneliness. What should be studied? Are depressive persons often alone, do lonely people become depressed, does depression make people lonely, are people in general more depressed when alone, and does this relation hold for depressed people? ESM data analyses require more specific research questions. The pitfall lies in the convergence of person-related questions and situation-related questions, both of which are represented in ESM data and can lead to divergent conclusions. In addition, results from ESM studies are easily misinterpreted. The conclusion that agoraphobia is a bogus diagnosis because ESM data show that people with this diagnosis feel worse when at home (and better when out) disregards the fact that people may select going out on days they feel best. In correlational relations, causal attributions should be avoided.

**DISCUSSION**

ESM was developed to sample moments from daily life. Since then, the method has yielded valuable results to better understand mental illness. ESM yields a comprehensive view of the participant’s daily life. This comprehensive view is valuable in clinical trials because the same instrument can measure various constructs, such as quality of life, psychopathology, social networks or productive activities. This broad measurement is done with high sensitivity to change. Even more, ESM allows assessment of mechanisms that are difficult to assess with cross-sectional questionnaires, such as stress-sensitivity, coping, reward-sensitivity and resilience. ESM has the potential to replace a larger body of instruments and sensitively assesses new and relevant parameters. It is, however, not yet the gold standard in clinical trials. This fact is mainly due to the heterogeneity in ESM applications. Nevertheless, it has the potential to become the de facto gold standard.

**ESM monitoring in clinical trials**

ESM data are ideal to inform and optimise personalised treatment in mental health. Feedback modules have been developed that translate ESM data into understandable graphs and figures that help promote patient engagement in treatment. Kramer et al assessed the benefits of an ESM feedback-intervention in a randomised control clinical trial. An ESM feedback-intervention was studied in addition to treatment as usual in patients diagnosed with depression. They demonstrated that ESM with weekly feedback moments reduced depression more than treatment as usual. ESM monitoring without feedback had a similar association, but this reduction did not last. The majority of ESM studies have short assessment periods, ranging from several days to a few weeks. In clinical practice, it is important to follow a patient for months or years. To reduce the burden, clinicians experiment with interrupted periodic ESM assessments. Psychiatrists and patients can mutually agree to customise ESM to monitor the effectiveness of treatment and update interventions accordingly. There is growing insight into what might be relevant to individual patients. For example, a patient who used ESM for a year had increased associations between symptoms in periods of exacerbation (personal communication, Dr. M. Bak). The patient’s subjective feeling of relapse coincided with changes in ESM towards higher severity. Monitoring patients this way provides both the clinician and the patient with additional information that can improve a personalised treatment process.

Ecological momentary intervention (EMI) has been cued as an umbrella term for mobile interventions in the daily living environment of participants. Unlike ESM, EMI aims not to observe, but to intervene. Only recently has mobile technology allowed one to bridge the gap between the therapist’s office and situations in daily life, where acquired skills should generalise and be autonomously applied. Several studies have demonstrated that EMI is feasible, acceptable and can be successful in mental health treatment.

**Single case clinical trials**

The translation of evidence from clinical trials assumes that patients are a homogeneous group. This is often not true. A large variability exists in treatment responses between subjects. Single case designs can be used to customise treatment to individual needs and accommodate to strengths. Since treatment effect (or outcome optimisation) can be difficult to detect, single case trials can add (experimental) control to clinical decision-making. This control requires that the timing of conditions is allocated randomly (independent from the clinical effect). Guidelines are available for the randomisation of single case trials and statistical software packages facilitate the analyses. An interesting innovation is a user-initiated program, which implements ESM to monitor the effects of gradual dose reductions in the user’s antidepressant medication. Over the course of a year, the user collected over a thousand ESM observations. Meanwhile, his clinician teamed up with a pharmacist to blindly implement a dose-reduction (tapering) schedule. From the data, early warning signs were detected for mood changes that predicted relapse. They could relate the mood change patterns to the medication reduction and were able to find the optimal, safe medication dosage. The use of ESM in classical clinical trials is less challenging because only aggregated group findings are pursued. Global constructs can be used to comprehensively describe the mental state of groups of individuals. In single-case studies, the data are highly individual and can benefit from the individual selection of items that maximally reflects the user’s variability.
Future of ESM in clinical trials

The advantages of single case trials to optimise treatment are increa-
sively recognised. In addition, there is increasing recognition of how
mobile phones can be used as medical instruments in the treatment of
mental illness. The benefits of combining mobile technologies with
single case designs are also recognised. Dissemination in regular
practice requires freely available user-friendly technology and software
that reduces the logistical burden and need for statistical skills. By
making this technology available, treatment efficiency can be improved.
Most participants are already used to carrying mobile phones, so the
threshold for using them to collect ESM data via an app is greatly
reduced. ESM progresses in parallel with evolutions in mobile technol-
ogy. Data collection logistics are reduced by innovations in tools and
programs, as well as cultural changes related to the use of mobile
devices. Furthermore, innovation in mobile tools facilitates the general
acceptance of momentary assessment. For broader clinical dispers-
ion, great investments are needed in software that allows on-the-fly
comprehensive feedback. More research should focus on the way data
become information. This knowledge is important for academically
trained professionals, as well as for clinicians who are more reluctant
to use and less able to understand statistical analyses. In addition, the
patients themselves should be able to comprehend the data. Ideally,
treatment planning, implementation and evaluation should be a colla-
borative process. ESM can prove its utility both in large-scale studies
and at the level of clinical practice. The latter is especially important,
as ESM can be used to optimise individual treatment and to help in
decision-making. In the future, assessment strategies that are more
sensitive in detecting change could be used in all prescriptions of pain
or psychotropic medication. Medication should be continued only when
efficacy is determined in a specific patient. Sensitive assessment
strategies can also be used in the domain of optimal dose finding.
Furthermore, recovery oriented care in psychiatry is not restricted to ill
health and symptoms, but extends to well-being. Daily life assessment
strategies can inform patients and clinicians how to improve resil-
ience by maximising strengths in patients.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

OPEN ACCESS

Open Access This is an Open Access article distributed in accordance with the
Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits
others to distribute, remix, adapt, build upon this work non-commercially, and license
their derivative works on different terms, provided the original work is properly cited
and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/
doi:10.1136/ebmental-2016-10241B

Received 14 June 2016; Revised 30 June 2016; Accepted 1 July 2016

REFERENCES

1. FDA. Guidance for industry: patient-reported outcome measures: use in medical
2. DeVries MW. The experience of psychopathology: investigating mental disorders in
in individuals with mental illness: reflections and guidance. Acta Psychiatr Scand
4. Delespaul PAEG. Assessing schizophrenia in daily life the experience sampling
method, UPM, Universitaire Pers Maastricht, University Library, Maastricht University
[Host], 1995.
5. Hektner JM, Schmidt JA, Csikszentmihalyi M. Experience sampling method:
8. Reis HT, Gable SL, Event-sampling and other methods for studying everyday
experience. In: Reis HT, Judd DM. Handbook of research methods in social
and personality psychology. New York, NY: Cambridge University Press,
2000:190–222.
9. Hurford R, Heavey CL. To beep or not to beep: obtaining accurate reports about
11. Bolger N, Laurenceau JP. Intensive longitudinal methods: an introduction to diary
14. Doherty ST, Lemieux CJ, Canally C. Tracking human activity and well-being in
on empowerment in a randomized controlled trial in patients with depression.
psychiatry: a multi-site collaborative study of acceptability, compliance, and
19. Silvia PJ, Kwapii TR, Edgington KM, et al. Missed beeps and missing data:
dispositional and situational predictors of nonresponse in experience sampling
research in schizophrenia and severe mental disorders. Schimph Bull
2012;38:386–95.
for the perplexed. In: delVries MW. The experience of psychopathology: investigating
mental disorders in their natural settings. New York, NY: Cambridge University Press,
1992:59–79.
23. Wang LP, MAXwell SE. On disaggregating between-person and within-person effects
24. Davison AC, Hinkley DN. Bootstrap methods and their application. New York, NY:
25. van Os J, Rutter BP, Poulton R. Gene-environment interactions in schizophrenia:
review of epidemiological findings and future directions. Schizophr Bull
2008;34:1066–82.
26. DeVries MW. Recontextualizing psychiatry: toward ecologically valid mental health
27. Mitchell V, Mathis MD. FDA briefing document, psychopharmacologic drugs
advisory committee meeting, topic: cognitive dysfunction in major depressive
experience sampling method in the treatment of depression: a randomized
29. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile
technology into psychosocial and health behaviour treatments. Br J Health Psychol
Interventions Through Single-Patient (N-of-1) Trials. Soc Psychol: Personality Compass
2014;8:408–21.
31. Gallely J, Raff BR. Optimizing behavioral health interventions with single-
32. Heyvaert M, Onghena P. Randomization tests for single-case experiments: state of
the art, state of the science, and state of the application. J Contextual Behav Sci
33. Bulté I, Onghena P. An R package for single-case randomization tests. Behav Res
Methods 2008;40:467–78.
34. Wickers M, Groot PC, Psychosystems E, et al. Critical slowing down as a personalized
35. Gravenhorst F, Masuree A, Bardram J, et al. Mobile phones as medical devices in
36. Vilardaga R, Bricker J, McDonell M. The promise of mobile technologies and single
case designs for the study of individuals in their natural environment. J Contextual