

# APPLIED CLINICAL TRIALS

YOUR PEER-REVIEWED GUIDE TO GLOBAL CLINICAL TRIALS MANAGEMENT

[www.pharmaportal.com](http://www.pharmaportal.com)

February and March 2001



Patient compliance. Reliable data.  
Keys to a successful clinical trial.

patient compliance  
guaranteed  
[www.invivodata.com/guarantee](http://www.invivodata.com/guarantee)

invivodata uniquely delivers reliable patient experience data by integrating electronic patient diary technology with an in-depth understanding of behavioral science to drive compliance.

[www.invivodata.com/clinical](http://www.invivodata.com/clinical)

Visit us at the 13th Annual DIA EuroMeeting in Barcelona, Spain



© 2000 Invivodata. All rights reserved. | Third party products and brand names may be trademarks or registered trademarks of their respective owners.

©Reprinted from  
APPLIED CLINICAL TRIALS, February & March 2001

AN ADVANSTAR PUBLICATION

Printed in U.S.A.

# **APPLIED CLINICAL TRIALS**

---



**Subject Experience Diaries  
in Clinical Research**

**Part 1**

## **The Patient Experience Movement**

**Subject Experience Diaries  
in Clinical Research**

**Part 2**

## **Ecological Momentary Assessment**

**Saul Shiffman  
and  
Michael R. Hufford**

---

# The Patient Experience Movement

Saul Shiffman, Michael R. Hufford, and Jean Paty

Electronic experience diaries can provide high rates of diary compliance and high-quality data in clinical trials—and the quality-of-life measures that sponsors increasingly expect.



Understanding a patient's experience following administration of a drug is becoming a critical objective of pharmaceutical science. Although in clinical research we commonly use objective or researcher/physician-based measures (for example, blood pressure) to assess changes in bodily function, many drug effects can be determined only by asking the subjects. Examples abound—for instance, subject experience measures are the primary efficacy measures for pain medications and many psychiatric medications. Even when the focus is on objective events, subjects are uniquely positioned to report behavior changes or symptom reductions for many disorders, such as asthma or irritable bowel syndrome. Moreover, patient outcome measures, such as evaluations of quality of life, have become important means to differentiate drugs in the same class. Such

examples are part of a growing trend in the pharmaceutical industry that makes patient experience a decisive factor in determining a drug's success.

Obtaining reliable, valid subject experience data requires a thorough understanding of the processes through which subjects report about themselves and their behavior. For decades, scientists have studied the workings of memory, which is obviously relevant to accurate self-reporting. Another relevant area of scientific inquiry is psychometrics, the study and development of mental measurement. From these roots has arisen an approach called *ecological momentary assessment*, which entails the real-time measurement of patient experience in the real world, at the point of experience. The findings from these areas of research form the scientific foundation for capturing reliable, valid subject experience data in clinical trials.

This first of a two-part article explores the meaning and importance to clinical research of subject experience data. It also reviews the methods used to collect the data, focusing in particular on paper and electronic subject experience diaries. The second part discusses the scientific principles underlying ecological momentary assessment and how these principles can enhance protocol design and measurement sensitivity.

## What is patient experience?

To describe patient experience, we distinguish three kinds of patient experience measures: self-observation, subjective symptoms, and quality of life.

**Self-observations.** Clinical trials often rely on subjects to report observable signs and symptoms; some examples include coughing, incontinence episodes, or sleep disturbance. These targets are objective and could, in principle, be observed by others,

but subjects are in a unique position to report on those events as they occur. In such instances, the subject is acting as a research observer, uniquely positioned to observe symptoms over time and across a spectrum of real-world environments and circumstances.

**Subjective symptoms.** Subjects are also often asked to report on important symptoms that could not, even in principle, be observed by a third party—examples include pain, fatigue, or anxiety. In some cases, relief of these symptoms is a surrogate indicator that the treatment is addressing the underlying biological disease processes. In many other cases, relief of subjective distress is itself the major indication, or purpose, of treatment, and the subject's experience is the primary target and the primary efficacy measure.

**Quality of life.** Ultimately, the purpose of medical treatment is to improve patients' lives, as evaluated by the patients themselves. Patients, payers, and (in some jurisdictions) regulators are increasingly demanding that new treatments demonstrate improved quality of life—in other words, that they enhance a patient's sense of well-being and the ability to function in important life domains—physical, occupational, and social.

## Importance of subject experience

Researchers already gather subject experience measures in most phases of clinical trials (Phase 2 through Phase 4) as primary or secondary endpoints, or as outcome variables. A number of industry trends will further press the demand for these measures in clinical studies:

**Patient demand.** An aging, and increasingly demanding, patient population drives interest in patient experience. Thanks to advances in medical diagnostics and therapies, physicians have turned to the management and amelioration of

chronic disease, in which long-term symptom control is a priority. Patients expect not just to live longer and have less objective disease, but also to feel better.

**Growing competition** among drugs intended for the same indication also has focused attention on patient experience. When patients, physicians, and managed care formularies can choose from several drugs that have similar biological efficacy, the choice often will pivot on which formulation most improves patient well-being. Enhanced quality of life has become a critical element in determining a drug's delivered value and thus figures into the economic calculation of a drug's benefits. In making formulary decisions, these benefits are weighed against the drug's cost. DataEdge (Fort Washington, PA) has found that formulary managers will pay 10% more for a compound that delivers the same level of efficacy but with improved patient quality-of-life characteristics.<sup>1</sup>

Increased competition also places greater demands on competitive claims for similar drugs on the market. Many such claims will require patient experience measures to determine the differential effects of competing drugs. Direct-to-consumer advertising of drugs has accelerated this trend; consumers are more influenced by benefits they can experience than by the drug's biological profile.

### Capturing subject experience in clinical trials

Researchers typically assess subject experience in summary and in retrospect, by asking questions of subjects during clinic visits. Subjects are asked to recall and summarize their experience over some period of time (typically days or weeks) and make judgments about the occurrence of events or their frequency. Unfortunately, a large body of recent data on the workings of autobiographical memory indicates that such recall is inherently inaccurate and, worse, subject to many distorting biases.<sup>2</sup> Research on memory indicates that recall does not simply involve retrieval from some archival mental storeroom of records.<sup>3</sup> Rather, most recall actually involves active reconstruction of the target events, using fragmentary recall and a variety of heuristic methods to construct an account of past events. Such retrieval or reconstruction is subject

**TABLE 1 Biases affecting subject recall data**

Recall bias	Description	Effect on the data
Effort after meaning	Past events or symptoms are reconstructed so as to make them consistent with the subject's understanding of events, or with subsequent events or symptoms.	Subjects' guesses at whether they are on active medication may influence "recalled" symptoms.
Saliency	Events that are more salient (i.e., personally relevant) are more likely to be recalled. Emotionally important events are particularly dominant in memory.	Salient events may result in under- or overestimation of disease episodes or symptom severity.
Recency	More recent events are more accessible to memory, and thus exercise undue influence on recall.	Can lead to underestimation of therapeutic effects, if assessment occurs as symptoms begin to return.
Influence of current state	Subjects' current state influences recall to be consistent with that state.	Context for data collection (e.g., research site) may lead to systematic biases—akin to white-coat hypertension. <sup>4</sup> Because the research/recall context is uniform across groups, that narrows differences across treatment groups.
Telescoping	Subjects systematically recall events as having been more recent than they actually were.	Subjects may include pretreatment symptom experience in recall of symptoms while on therapy.

to multiple biases stemming from the layers of processing that are necessary to answer the typical research inquiry. Each step in the recall process has the potential to introduce significant bias into the data.

**Encoding.** In order for any content to be available for later recall, it must first be encoded into memory. Encoding is incomplete and imperfect and is influenced by a variety of processes. For example, the perceived salience at the time of experience influences whether content is encoded. This varies with the person's subjective state, distraction, and the like. An individual's state and situation similarly affect the way content is encoded (that is, what is stored), especially for evaluative information. In other words, the brain does not make a faithful, unbiased record of events, but selects and edits events based on complex factors that can bias later recall.

**Recall.** Recall of events is actually a reconstruction of events—a retrospective re-creation of what "must have" happened. A number of well-known inaccuracies and biases affect the retrieval of events from memory (Table 1).

**Aggregation and summary.** Subjects must not only retrieve the relevant information, but then also aggregate and summarize it over some period or sequence of events, to answer typical research questions:

"How many headaches did you have this week?" "How much gastric discomfort did you have this week?" Research on autobiographical memory makes it clear that the brain does not process these requests by counting or averaging but rather relies on a variety of heuristic strategies for estimating answers.<sup>2</sup> Those strategies introduce substantial biases, similar to those that affect recall. For example, recent research shows that recall of "average" pain using a pain episode is heavily influenced by the peak pain experienced (saliency) and the pain intensity at the end of the episode (recency). Thus, aggregate or summary recall of experience over an interval may produce additional biases.

### Autobiographical memory biases

In a research setting, autobiographical memory biases have two common effects:

- Collection of "dirty" (unreliable) data
- Collection of "biased" data.

Errors of recall make autobiographical data unreliable and inaccurate. Unreliability is introduced when the recall errors are random and exercise no systematic or directional influence over the reports or the study outcomes. Noisy data decreases the sensitivity of studies—the research may fail to detect effects (for example,

**TABLE 2 The limitations of paper diaries**

The problem	Explanation
Poor compliance	Many entries are omitted, and particular items are often missing.
Faked compliance (often impossible to detect)	Entries are not made in real time but are “backfilled,” often right before the study visit, to make the subject appear compliant.
Timing of diary entries introduces bias	Subjects choose when to make entries and often choose biased moments (e.g., when symptoms are at a peak or are quiescent).
Poor data quality	Many out-of-range and hard-to-capture responses, write-in comments, and other problems.
Limited branching and assessment sophistication	Assessment has to be simple.
Excessive subject burden	Difficult to complete diaries, especially if any branching is needed.
Processing paper diary data is slow and costly	High data entry costs; data management and query resolution very costly due to poor data quality and large volume.

drug-placebo differences) that are actually present. At a minimum, errors of recall drive demand for large numbers of subjects to achieve statistical power to overcome the noise in the data. Worse, they undermine attempts to demonstrate a drug effect even when it is present. Obviously, this failure is wasteful and undermines the very purpose of the research enterprise.

Recall errors can also introduce systematic errors, or bias. This bias is present across all groups, creating a uniform process that can blur differences between medication and placebo groups. Bias also arises from the use of a uniform clinical context for data collection. For example, if study personnel collect reports of past pain after a pleasant interaction with subjects, subjects will systematically underestimate past pain. If such reports are gathered after a painful exam, subjects will overestimate the discomfort. If most subjects' pain resolves by the end of the study, when reports are collected, past pain—and active-placebo differences in past pain—will systematically be underestimated in both treatment groups. Thus, biases introduced by recall can often lead to false negatives—that is, failure to detect treatment effects. The defects in recall we have outlined are neither the product of motivated distortion nor under subjects' control. These biases pervade all recall data, no matter how cooperative or industrious the subject may be. They simply reflect the way

human memory works and thus cannot easily be avoided.

### Subject diaries

The solution is clear: rather than rely on recall, clinical investigators must collect data about subject experience in real time in the real world—in other words, at the point of experience. To gather this information, many researchers choose a subject diary.

**Advantages of diaries.** Collecting data using subject diaries has major advantages over data gathered at the research site or during in-person interviews.<sup>5</sup> Problems of autobiographical memory and recall biases are avoided if subjects make their entries in real time. Whereas subject reports that rely on recollection and summary of past experience are biased by distortions of recall, data recorded in real time better represents actual experience. Diary data is more representative of experience in everyday life because subjects provide data in many real-world settings.

The completion of multiple assessments potentially increases the reliability of the data. Aggregation of repeated assessments reduces noise in data (unreliability), making the measurement more sensitive to treatment effects.

The richness of diary data may allow for a more detailed analysis of treatment effects, particularly for those processes that exert their influence over time. Thus, researchers can examine trends in subject experience over time (for example,

assessing speed of pain relief as well as overall efficacy), the relationship of medications to compliance, or their relationship to other variables (for example, stress, diet, workload).

Thus, compared with reports obtained periodically at the research site, diaries have several potential advantages in validity and efficiency.

**Paper diaries.** Since the 1940s, paper diaries have been widely used to collect data from subjects in their natural environment.<sup>6</sup> Paper diaries are intended to overcome inaccuracy and bias associated with retrospective reports. Subjects are typically given a booklet or set of cards on which to record their experience. They may be asked to complete assessments after a specified event (for example, an asthma attack), at a particular time (for example, “report your average level of pain every four hours between morning and bedtime”), or when signaled by a beeper or programmable wristwatch. Reports can include open-ended notes or may be restricted to structured responses to closed-ended queries.

**Limitations of paper diaries.** Although paper diaries potentially afford some advantages, they typically do not fulfill their potential and are fraught with problems that undermine the goal of providing a real-time record of subject experience (Table 2).

*Poor and/or faked compliance.* Compliance can be faked, and noncompliance is difficult or impossible to detect. Many subjects complete their diary entries after the fact, often just before a study visit when the diary will be collected.<sup>7</sup> Subject compliance is variable; some subjects may falsify many of their entries. A review of eight studies that have independently verified diary completion found that actual rates of completion were only 54%, while reported rates of paper diary completion were 88%.<sup>8–15</sup> Noncompliance with paper diaries undermines the entire purpose of collecting diary data and may exacerbate problems of recall. Indeed, as Goldsmith noted, 50% noncompliance can increase sample size requirements by a factor of four.<sup>16</sup> As a result, most diary studies are dramatically underpowered to test study hypotheses. Poor and/or faked compliance is the greatest limitation of paper diary methods.

*Timing-introduced bias.* The timing of

diary entries introduces bias. Even when subjects do complete diaries in the field, the timing of the recording can introduce bias into the data. For example, subjects who have been instructed to complete diaries at regular intervals may be reminded to make a diary entry when their symptoms flare up, which would result in an overestimation of symptom intensity and a blurring of group differences. Or subjects may remember to make diary entries when they are taking their medication, which is typically when drug effects are at their minimum because time-since-last-dose is greatest. Thus, allowing the subject to choose when to complete a diary entry often introduces unintended bias.

**Poor-quality data.** Paper diaries produce data of poor quality. Responses to structured diary items are often missing or out of range. Furthermore, written responses are often unreadable or uncodable. Subjects can usually see previous diary entries, which can introduce reactivity bias—subjects may either strive for consistency across responses or vary their responses to avoid the appearance of being “stuck.”

**Assessment limitations.** Paper diaries limit the length and the sophistication of assessments. Paper diaries are often forced into overly simplistic, single-stream statements that limit the data that can be collected. Sophisticated skip or contingency patterns often produce missing or uninterpretable data. This happens

## Electronic Diaries



The use of electronic subject experience diaries in clinical trials offers many potential benefits:

- Compliance can be improved.
- Compliance cannot be faked.
- Compliance can be assessed and tracked.
- Queries are dramatically reduced.
- Timing of subject entries can be controlled.
- Reactivity to monitoring is reduced.
- Subjects prefer using electronic diaries.
- Assessments can be as sophisticated as necessary.
- Diaries can be programmed to not overburden subjects.
- They facilitate trial completion.

when administration of an assessment depends on a prior item or on other variables, such as the subject’s treatment status or prior assessments. As a result, using skip or contingency patterns is unmanageable or simply impossible.

**Overburdened subjects.** Paper diaries can potentially overburden subjects with numerous assessments. When diary entries are set up to record events such as symptom flare-ups, the subject’s burden is proportional to symptom severity. That is, an increasing rate of symptoms can overwhelm subjects as they try to complete an assessment for each occurrence. Increasing

subject burden through an inflexible paper diary protocol can introduce systematic bias in the data, as well as offer a perverse incentive for subjects to avoid reporting symptoms.

**Slow and costly processing.** Processing paper diary data is slow and costly. Handwritten data is often hard to read, missing, or out of range. Moreover, subjects often enter extraneous information, which must then be transcribed, coded, and entered into a database. As a consequence of poor quality, diary data can generate an extraordinary number of queries, which take substantial time and money to resolve. These queries must be resolved before the data can be used, creating a significant burden in cost and time (it is estimated that resolving a single query can cost \$70 or more).<sup>17</sup> The net result is that paper diary data is often unusable, and the ensuing wastefulness has impeded the use of diary methods in clinical trials.

## The electronic solution

Technology, in the form of “digital paper” diaries, was initially seen as a panacea for the data quality and compliance problems associated with paper diaries. These digital paper diaries simply translate the paper diary into electronic form. These systems are not a complete solution because they fail to remediate many of the problems—especially low rates of subject compliance associated with paper diaries. Indeed, subjects can be just as noncompliant with digital paper diaries as with traditional paper diaries.<sup>18</sup>

Subject experience data can be collected using electronic experience diaries, which systematically apply principles from behavioral science implemented using sophisticated software solutions and handheld computers. When thoughtfully applied with an understanding of cognitive science, measurement theory, technology, and research design and analysis, electronic diaries have the potential to overcome many of the limitations of written, paper-based diaries (see Electronic Diaries box).

**Compliance is improved.** To ensure high compliance rates with electronic diaries, it is critical that several principles from behavioral science be incorporated into the diaries to achieve and maintain compliance (Table 3). These principles draw

**TABLE 3** Diary compliance principles<sup>a</sup>

Compliance principle	Advantage
Build compliance into the protocol	Facilitates compliance and discourages noncompliance. Also allows relevant compliance data to be tracked.
Subject training	Engenders compliance across sociodemographic status and familiarity with computers.
User interface issues	Subject experience diaries must be intuitive and easy to use, to help subjects succeed in the trial.
Drive the protocol	Real-time compliance reminders help subjects maintain good compliance. Can also drive other events (e.g., medication).
Guide subjects through assessments	Real-time data edit checks and programmed branching ensure that the right data is collected as required by the protocol.
Include “livability” functions	Successful solutions incorporate flexible features that allow subjects to incorporate the experience diary and protocol into their lives.
Create a sense of accountability	Subjects are more compliant with protocols if they know their compliance is being tracked and monitored. Subject feedback and reinforcement are essential.

<sup>a</sup>Principles from behavioral science that are required to ensure diary compliance.

on the empirical literature surrounding compliance in clinical trials broadly, behavioral science, user interface, and diary design.<sup>6,19–22</sup> Trials using electronic diaries with a compliance system can have far better compliance than their paper diary counterparts. Using an electronic diary solution developed over many years and dozens of trials, we have achieved 93%–99% timely compliance with all electronic data-collection prompts.<sup>23</sup>

**Compliance cannot be faked.** Sophisticated electronic diaries record missed entries, as well as the date and time of each entry. As a result, missed entries cannot be made up at the last minute, avoiding the faked compliance that corrupts paper diary data.

**Compliance can be assessed and tracked.**

*The expense may initially appear daunting, but the costs of electronic diaries are often offset by efficiencies and reduced charges for data cleaning.*

Instead of hoping for the best, subject compliance and data quality can be checked in real time, allowing for rapid intervention and saving both time and money. We have developed empirically derived statistical algorithms that can identify noncompliant subjects early, before they become costly resource drains.

**Dramatically reduces queries.** With proper system design and real-time edit checks, out-of-range responses are eliminated. Electronic diaries prevent missing data and the entry of invalid data. When properly applied, electronic diaries help to ensure correct, valid data entry in real time at the point of experience.

**Controls timing of subject entries.** The timing of entries can be scheduled by design, thereby eliminating biases due to subject selection of entry times. Entries can be scheduled at any time during the day, and may include random sampling throughout the day to achieve an unbiased, representative sample of the subject's experience.

**Reduces reactivity to monitoring.** Most electronic diaries present only the current assessment; subjects cannot review past assessments, which eliminates an important source of reactivity bias.

**Subjects prefer electronic diaries.** Considerable empirical evidence shows that subjects prefer electronic diaries to paper diaries.<sup>24–26</sup>

**Assessments can be as sophisticated as necessary.** Electronic diaries have the potential to handle complex branching and skip patterns, as well as a host of other contingencies—for example, responses based upon a preceding item, subject status, the stage of the protocol, or the passage of time—while keeping the subject's task simple.

**Subjects are not overburdened.** Unlike paper diaries, electronic diaries can administer complex sampling designs that equalize the assessment burden on subjects, even if the frequency of recorded events differs. For example, if the frequency of

symptoms increases, electronic diaries can modify the assessment strategy to ensure that compliance does not become difficult. The sampling can randomly assess target events to guarantee obtaining a representative sample of data. In this way, the systematic application of sophisticated algorithms can select occasions for assessment in an unbiased manner.

**Electronic diaries facilitate trial completion.** The scientific application of electronic diaries allows for rapid, efficient cleaning of data. Queries are reduced, and the data is processed more quickly and efficiently in parallel with data collection, delivering a locked database faster. Capturing data using electronic diaries eliminates the timely transcription, coding, and data entry associated with paper diary data.<sup>24,27–28</sup>

### **Challenges in using electronic diaries**

Although electronic diaries have the potential to overcome, or dramatically reduce, many of the drawbacks found in paper diaries, they present important challenges to trial managers when used in clinical research.

**Compliance.** Sophisticated electronic

diaries can produce very high rates of diary compliance, but not all electronic diaries are equal in this regard. High rates of subject compliance with electronic diaries cannot be assumed. The empirical evidence points to the importance of using behavioral science as a foundation for high rates of compliance in electronic diary trials.

**Cost.** Electronic diaries are more costly to use than paper diaries. Hardware must be purchased or leased; software development, configuration, testing, and validation must be completed; and monitors, study coordinators, and subjects must be trained. Although the expense may initially appear daunting, these “add-on” costs are often offset by the reduced charges for data cleaning and the efficiencies generated by well-designed electronic diaries. Also, some of the aforementioned costs—for example, staff and subject training—are incurred with paper diaries, but they tend to be hidden in omnibus site or monitoring costs.

**Technological demands on users.** New technologies can be complicated. If an electronic diary is not well designed, subjects and study staff will find it onerous to use. Thus, careful program design that incorporates human factors and user training is essential if electronic diaries are to provide value. With appropriate design and instruction, almost all subjects, even without any computer experience, can successfully navigate an electronic diary.

### **Fulfilling the promise of diary research**

Subject experience has become an important objective of clinical research. Yet studies and experience show that traditional ways of quantifying subject experience, which rely on questionnaires to collect subject recall and summary data, are not adequate to the task.

Data gathered through diaries can be a valuable source of information about the real-world, real-time effects of drugs, including their impact on symptoms and quality of life. Unfortunately, paper diaries are fraught with problems that prevent them from delivering the type and quality of data that provides value. By combining the principles of behavioral science with the technology of handheld computing, high rates of diary compliance and data quality can be achieved in clinical trials.

The second and final part of this article explores ecological momentary assessment, a methodological approach that integrates considerations of autobiographical memory, psychometric assessment, experience sampling, and technology to enhance the sensitivity of diary methods in clinical trials.

## References

- H.E. Glass, "Formulary Listings and the Quality of Life," *Scrip*, 1-3 (September 1995).
- R. Tourangeau, "Remembering What Happened: Memory Errors and Survey Reports," in *The Science of Self-Report*, A.A. Stone et al., Eds. (Lawrence Erlbaum Associates, Mahwah, NJ, 2000), pp. 29-47.
- S. Shiffman, "Real-Time Self-Report of Momentary States in the Natural Environment: Computerized Momentary Assessment," in *The Science of Self-Report*, A.A. Stone et al., Eds. (Lawrence Erlbaum Associates, Mahwah, NJ, 2000), pp. 277-296.
- T.G. Pickering, A. Coats, J.M. Mallion, G. Mancina, and P. Verdecchia, "Blood Pressure Monitoring. Task Force V: White-coat Hypertension," *Blood Pressure Monitoring*, 4, 333-341 (1999).
- J.B. Houpt, R. McMillan, C. Wein, and S.D. Paget-Dellio, "Effect of Glucosamine Hydrochloride in the Treatment of Pain of Osteoarthritis of the Knee," *Journal of Rheumatology*, 26, 2423-2430 (1999).
- L.M. Verbrugge, "Health Diaries," *Medical Care*, 18, 73-95 (1980).
- Stephen A. Raymond and Robert N. Ross, "Electronic Subject Diaries in Clinical Trials," *Applied Clinical Trials*, March 2000, 48-58.
- F. Chmelik and A. Doughty, "Objective Measurements of Compliance in Asthma Treatment," *Annals of Allergy, Asthma, and Immunology*, 73, 527-532 (1994).
- G. Jonasson, K.H. Carlsen, A. Sodal, C. Jonasson, and P. Mowinkel, "Patient Compliance in a Clinical Trial with Inhaled Budesonide in Children with Mild Asthma," *European Respiratory Journal*, 14, 150-154 (1999).
- R.S. Mazze, H. Shamoan, R. Pasmantier, D. Lucidio, J. Murphy, K. Hartmann, V. Kuykendall, and W. Lopatin, "Reliability of Blood Glucose Monitoring by Patients with Diabetes Mellitus," *The American Journal of Medicine*, 77, 211-217 (1984).
- H. Milgrom, B. Bender, L. Ackerson, P. Bowry, B. Smith, and C. Rand, "Noncompliance and Treatment Failure in Children with Asthma," *Journal of Allergy and Clinical Immunology*, 98, 1051-1057 (1996).
- M. Simmons, "Unpredictability of Deception in Compliance with Physician-Prescribed Bronchodilator Inhaler Use in a Clinical Trial," *Chest*, 118, 290-295 (2000).
- S.L. Spector, R. Kinsman, H. Mawhinney, S.C. Siegel, G.S. Rachelefsky, R.M. Katz, and A.S. Rohr, "Compliance in Patients with Asthma with an Experimental Aerolized Medication: Implications for Controlled Clinical Trials," *Journal of Allergy and Clinical Immunology*, 77, 65-70 (1986).
- R.J. Straka, J.T. Fish, S.R. Benson, and J.T. Suh, "Patient Self-Reporting of Compliance Does Not Correspond with Electronic Monitoring: An Evaluation Using Isosorbide Dinitrate as a Model Drug," *Pharmacotherapy*, 17, 126-132 (1997).
- P. Verschelden, A. Cartier, J. L'Archeveque, C. Trudeau, and J.L. Malo, "Compliance with and Accuracy of Daily Assessment of Peak Expiratory Flows (PEF) in Asthmatic Subjects over a Three Month Period," *European Respiratory Journal*, 9, 880-885 (1996).
- C.H. Goldsmith, "The Effect of Compliance Distributions on Therapeutic Trials," *Research Applications and Future Research*, 297-308 (1979).
- CenterWatch, Analyst Reports, "Will EDC Finally Catch On?" *CenterWatch Newsletter*, 6 (9) 1-5 (September 1999).
- M.E. Hyland, C.A. Kenyon, R. Allen, and P. Howarth, "Diary Keeping in Asthma: Comparison of Written and Electronic Methods," *British Medical Journal*, 306, 487-489 (1993).
- P. Boudes, "Drug Compliance in Therapeutic Trials: A Review," *Controlled Clinical Trials*, 19, 257-268 (1998).
- J. Raskin, *The Humane Interface: New Directions for Designing Interactive Systems* (Addison Wesley Longman, Inc., Reading, MA, 2000).
- L.R. Trick, "Patient Compliance—Don't Count on It!" *Journal of American Optometry Association*, 64, 264-270 (1993).
- J. Urquhart, "Patient Compliance with Crucial Drug Regimens: Implications for Prostate Cancer," *European Urology*, 29, 124-131 (1996).
- T.W. Kamarck, S. Shiffman, L. Smithline, J.L. Goodie, H. Thompson, P. Ituarte, J.Y. Jong, V. Pro, J. Paty, J. Kassel, M. Gnys, and W. Perz, "The Diary of Ambulatory Behavioral States: A New Approach to the Assessment of Psychosocial Influences on Ambulatory Cardiovascular Activity," in *Perspectives in Behavioral Medicine: Technology and Methodology in Behavioral Medicine*, D. Krantz, and A. Baum, Eds. (Lawrence Erlbaum Associates, Inc., Mahwah, NJ, 1998), pp. 163-193.
- H.E. Drummond, S. Ghosh, D. Ferguson, D. Brackenridge, and B. Tiplady, "Electronic Quality of Life Questionnaires: A Comparison of Pen-Based Electronic Questionnaires with Conventional Paper in Gastrointestinal Study," *Quality of Life Research*, 4, 21-26 (1995).
- C.B. Johannes, S.L. Crawford, J. Woods, R.B. Goldstein, D. Tran, S. Mehrotra, K.B. Johnson, and N. Santoro, "An Electronic Menstrual Cycle Calendar: Comparison of Data Quality with a Paper Version," *Menopause*, 7, 200-208 (2000).
- J.M. Rabin, J. McNett, and G.H. Badlani, "Compu-Void II: The Computerized Voiding Diary," *Journal of Medical Systems*, 20, 19-34 (1996).
- S. Shiffman, C. Elash, S. Paton, C.J. Gwaltney, J. Paty, D.B. Clark, K.S. Liu, and M.E. DiMarino, "Comparative Efficacy of 24-hour and 16-hour Transdermal Nicotine Patches for Relief of Morning Craving," *Addiction*, 95, 1185-1195 (2000).
- B. Tiplady, A.H. Jamieson, and G.K. Crompton, "Use of Pen-Based Electronic Diaries in an International Clinical Trial of Asthma," *Drug Information Journal*, 34, 129-136 (2000).

**Saul Shiffman, PhD**, is professor of psychology (clinical and health psychology), psychiatry, and pharmacy at the University of Pittsburgh and chief science officer at *invivodata, inc.* **Michael R. Hufford,\* PhD**, is manager of scientific affairs at *invivodata*, 2100 Wharton St., Suite 505, Pittsburgh, PA 15203, (412) 390-3008, fax (412) 390-3020, e-mail: mhufford@invivodata.com. **Jean Paty, PhD**, is executive vice president of solutions development at *invivodata* and research assistant professor of psychology at the University of Pittsburgh.

\*To whom correspondence should be addressed.

## Subject Experience Diaries in Clinical Research, Part 2

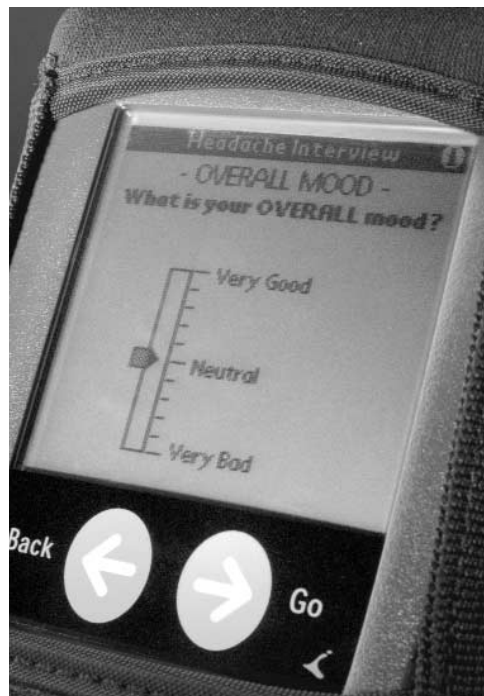
# Ecological Momentary Assessment

Saul Shiffman and Michael R. Hufford

**H**eightened competition for market share and growing concern about the role of human subjects in clinical research are leading pharmaceutical science to turn its attention to subject experience. Patients who become subjects in drug trials are being asked to observe their own symptoms, relief, and quality of life—and report their observations to investigators. Devising a way to capture their experience reliably and validly requires a deep understanding of the way people observe and report information about themselves and their behavior.

In part 1 of this article, “The Patient Experience Movement,” we discussed the importance of integrating cognitive science, measurement theory, technology, and statistical and research design in order to apply real-time measures to subject experience.<sup>1</sup> We also examined the utility of subject diaries and discussed biases and noncompliance issues affecting paper diaries. Part 1 showed how electronic experience diaries can overcome these problems. Here, we describe real-time methods used to capture subject experience: the science of ecological momentary assessment (EMA).

EMA is a methodological approach rooted in scientific principles evolving from cognitive psychology, statistical



**Representative samples of moments in each subject's experience can be as valuable to a study as a representative sample of subjects.**

sampling, psychometrics, and clinical study design. It enables researchers to quantify subject experience in the real world, in real time, using a collection of scientific methods to assess patient experience in real-world environments (the “ecological” aspect) in real time (the “momentary” aspect).

In a typical study using EMA, subjects' momentary experiences are assessed during the course of the day as they go about their normal activities. In other words, instead of relying on biased recall and summary to characterize the subjects' experience, EMA methods go directly to the source and collect data about real-world experience, as it is happening. After collecting a representative sample of subject experience, the data can be statistically aggregated to yield an accurate picture of a subject's status. EMA has three defining principles.<sup>2</sup>

Here, we describe each of the three characteristics of EMA—to maximize ecological validity, foster momentary assessment,

and ensure representation of experience across settings. To maximize ecological validity, researchers must study subjects in their natural environments. Ecological validity is concerned with the applicability of research findings to the real world. To foster momentary assessment and to avoid retrospective distortion of data, researchers must

collect real-time data about the subject's momentary (or nearly momentary) state. To ensure representative characterization of subject experience across the range of settings a person inhabits, EMA typically requires a sampling strategy that consists of many momentary assessments per day. Multiple time points also enable researchers to examine changes in subject experience over time.

### Characteristics of EMA

**Ecological validity.** Clinical research is meant to be relevant to real-world subject experience. Ecological validity refers to data that accurately reflects that experience. By collecting data in the environments that subjects typically inhabit, EMA can successfully capture real-world responses. Laboratory settings and clinicians' offices are both atypical environments. As a consequence, assessments made in those settings lack ecological validity. A good example of this effect is the well-documented phenomenon of "white-coat hypertension"—a condition in which blood pressures taken by "white-coated" medical personnel in clinical settings are falsely elevated, compared with real-world physiological parameters.<sup>3</sup> In other words, measures taken in artificial contexts often

*Real-world data from subjects' normal environments maximizes the likelihood that the data represents real-life experience and can be generalized to clinical application.*

misstate both the level and the range of the true response. By allowing subjects to provide real-world data in their normal environments, EMA maximizes the likelihood that the data represents actual real-life experience and can be generalized to clinical application.<sup>4-5</sup>

The ecological validity of subject experience data is too rarely considered in clinical trials. Researchers well understand the importance of gathering a representative sample of subjects in order to accurately gauge drug effects. To better understand drug effects, EMA adds another dimension to sampling—a representative sample of moments across a subject's experience.<sup>6</sup> Recall and reporting context influence subject reports. Relying on subject experience data collected in a single unusual environment (such as a research site) limits its generalizability, just as relying on a single demographic group in a clinical trial limits the generalizability of the results.

**Momentary assessment.** As we noted in part 1 of this article, the scientific literature demonstrates that data relying on retrospective recall is both unreliable and systematically biased. Thus, most recall data is suspect. Moreover, many studies of subject experience compound the problems associated with the retrospective reconstruction of events because data is collected in non-representative—indeed quite unusual—situations (a laboratory or other clinical research environment). Because the context of recall influences the content (see part 1), this has great potential to bias the data. By emphasizing real-time data collection in the real-world environment in which subjects live, EMA avoids the pernicious biases associated with retrospective self-report.<sup>4</sup>

Ecological momentary assessment brings scientific rigor to the capture and measurement of subjects' experiences by repeated and systematic sampling of their momentary experiences. This approach to understanding subject experience is grounded in the capture of multiple moments in a subject's life and is a distinct departure from methods that rely on recall or summary data, through questionnaires, interviews, or diaries. In those approaches, subjects are asked to recall and summarize their experiences over the past day or days, weeks, or even months. This process of recollection is prone to inaccuracy and bias.<sup>7</sup> Modern cognitive science research on autobiographical memory (how we recall our own experiences) demonstrates that recall is not only prone to inaccuracy but also to systematic biases that may distort subjects' reports and obscure treatment effects. Collecting data on momentary states avoids recall and thus avoids these biases. This is a central element of EMA and is supported by considerable evidence.<sup>4</sup>

Subjects' momentary states can be aggregated to achieve an accurate, valid picture of their global or average states. Multiple assessments of momentary states constitute samples of a subject's state. Just as thoughtful consideration of sampling frames and enrollment criteria is critical to achieving clinical endpoints, so too is the careful, unbiased sampling of key moments essential to drawing valid conclusions about subject experience. Biases can be introduced when such sampling considerations are not taken into account. For example, collecting data at bedtime, as is done in many diary studies, produces biased data because a subject's state at

bedtime is unrepresentative (arousal is at its ebb, and many symptoms show diurnal variations). That may bias the recall of earlier experiences. Research shows that reports of pain, for example, are influenced by pain experienced at the time of recall, by the peak level of pain throughout the day, and by the most recent levels of pain.<sup>8-9</sup> As a result, subjects' summaries of "average" pain seldom validly represent their average experience. By aggregating momentary data, investigators may create many indices of subject experience (for example, mean, mode, and peak) to understand how EMA data relates to subject satisfaction with treatment.

Sophisticated electronic diary solutions allow researchers to collect an unbiased, representative sample of momentary subject experience. For example, electronic diaries deployed within an EMA framework can prompt subjects at random intervals to report on their experience, thus avoiding both recall bias and the biases introduced when subjects choose when to complete an assessment. Such sampling schedules can be combined with subject-initiated reports to achieve a complete picture of the subject's experience. By allowing subjects to begin a report, sampling strategies will not miss important events. The momentary focus of EMA data allows the researcher to better understand drug effects on subject experience.

**Representative characterization.** Gathering real-time, real-world assessments requires careful attention to psychometrics in order to capture momentary states. Although studies using EMA typically rely on self-reports of subjective states or observed events, the methodology can also accommodate objec-

tive measures, such as performance and physical, physiological, and hormonal states. For both self-report and objective data, repeated administration of assessments raises both opportunities and challenges for subject performance and data aggregation and analysis. Careful protocol design, subject training, and data processing are crucial to successful EMA research.

It is important to stress that the credible study of subjects' momentary experience cannot be achieved merely by inserting a diary into a study or transferring a paper-and-pencil assessment to an electronic diary. Applying electronic technology to subject diaries is not an end in itself, although it can be the means to a solution. At their best, electronic diaries are an effective tool for applying the methodological discipline of EMA to the assessment of subject experience. What an EMA-driven electronic subject diary can achieve is sophisticated data assessment that incorporates branching, edit-checks, and real-time data processing.

### EMA study design and analysis

**Design issues.** To achieve the full richness of ecological momentary assessment, it is imperative to pay careful attention to study design and endpoints. Selecting the optimal frequency of assessments, for example, is critical. Too few assessments can undermine reliability and statistical power; too many can unduly burden subjects and data managers. Those determinations are specific to the study protocol, because they depend on variability in the predictor and outcome parameters as well as on the characteristics of the subject sample. Furthermore, a particular assessment strategy must account for changes in outcomes due to diurnal or environmental variation. Thus, a good understand-

*Careful planning of EMA methods and study designs permits a sophisticated data collection scheme, while keeping the analysis simple and powerful.*

ing of study objectives will help planners set up a suitable framework for sampling and data collection. It may include a mixture of time-based assessments (such as random time sampling), event-based assessments (for example, tracking events such as symptom flare-ups), and summary assessments (daily summaries, for example).

The potential for repeated real-time assessment of subject experience in real clinical contexts opens up opportunities for new insights into clinical phenomena. Rather than limiting assessment to average symptom severity, a sound EMA approach supported by sophisticated electronic diary technology can open a window onto the dynamics of symptom patterns over time. For example, drug effects on onset of symptom episodes can be distinguished from effects on severity or progress over time. Real-time data on the time medication is taken enables analysts to determine time-to-relief, duration of relief, and prevention of symptom episodes. The rich possibilities of EMA methods—and the potential to address refined clinical research questions—highlight the need for careful study

design and close linkage between protocol development and diary design.

**Data analysis.** The volume and density of data provided by ecological momentary assessment improve reliability and, as a consequence, statistical power. In many cases, it makes it possible to meet study objectives with smaller sample sizes, thus providing an opportunity to conduct highly focused statistical analyses. Careful planning of EMA methods and study designs permits data analysts to construct aggregated data sets that lend themselves to simple analysis. The data collection scheme can be sophisticated yet keep the analysis simple and powerful. Researchers can analyze disaggregated observations with more sophisticated statistical methods, such as random effects regression models and generalized estimating equations.<sup>9</sup>

With some designs, such as those that capture and study the time course-of-events (for example, time-to-relief, duration of episodes), it is crucial that analysts carefully select only those observations relevant to the process. They can do so by following approaches laid out in the protocol. Moreover, the analysis must account for the sampling designs included in the study. In other words, much of the strategic data processing lies in data summary and in the production of analytic data sets. In the end, close conceptual linkage between study objectives, study design, EMA design, and data analysis plans facilitates timely and successful study evaluation.

### Applications of EMA

Although its methods are relatively new to pharmaceutical science, EMA has been flourishing in basic and clinical research for more than a decade. Numerous studies sponsored by the National Institutes of Health have used the method, and the resulting studies have been widely published in the peer-reviewed literature.<sup>10</sup> Ecological momentary assessment has been used in research into treatments for arthritis, headaches, acute pain, hypertension, gastrointestinal disorders, urinary incontinence, asthma, depression, schizophrenia, anxiety disorders, alcoholism, smoking, cardiovascular disease, and other conditions (for a review, see Stone, Shiffman, and DeVries<sup>5</sup>). EMA data has been shown to be sensitive to drug effects and to differentiate similar drugs in Phase 4 studies.<sup>11</sup> Several studies have established that EMA data avoids the bias and inaccuracies present in retrospective data.<sup>12–13</sup> Moreover, such data has been related to many different types of momentary physiological data, including ambulatory blood pressure, peak expiratory flow, and hormonal parameters.<sup>14–18</sup>

In sum, EMA methods have been extensively used in research and have been validated against a variety of types of real-time physiological data. The methods can make a substantial contribution to pharmaceutical clinical trials.

Effective execution of EMA research requires strategists and planners with behavioral and cognitive science expertise and a facility for measurement theory, technology, and statistical and research design. Applied intelligently, such studies yield higher quality data and have a propensity to answer research questions that could not otherwise be addressed. EMA provides the scien-

tific foundation for using electronic experience diaries in clinical studies in a manner that ensures sensitive measurement of subject experience, subject compliance, data quality, timely trial completion, and overall cost-effectiveness. The ultimate benefit is successful clinical demonstration of drug effects.

## References

1. Saul Shiffman, Michael R. Hufford, and Jean Paty, "Subject Experience Diaries in Clinical Research, Part 1: The Patient Experience Movement," *Applied Clinical Trials*, February 2000, 46–56.
2. A.A. Stone and S. Shiffman, "Ecological Momentary Assessment in Behavioral Medicine," *Annals of Behavioral Medicine*, 16, 199–202 (1994).
3. T.G. Pickering, A. Coats, J.M. Mallion, G. Mancina, and P. Verdecchia, "Blood Pressure Monitoring, Task Force V: White-Coat Hypertension," *Blood Pressure Monitoring*, 4, 333–341 (1999).
4. S. Shiffman, "Real-Time Self-Report of Momentary States in the Natural Environment: Computerized Ecological Momentary Assessment," in *The Science of Self Report: Implications for Research and Practice*, A.A. Stone et al., Eds. (Lawrence Erlbaum Associates, Inc., Mahwah, NJ, 2000), pp. 277–296.
5. A.A. Stone, S. Shiffman, and M.W. DeVries, "Ecological Momentary Assessment," in *Well-Being: The Foundations of Hedonic Psychology*, D. Kahneman, E. Diener, and N. Schwarz, Eds. (Russell Sage Foundation, New York, 1999), pp. 26–39.
6. M.R. Hufford and S. Shiffman, "Capturing Real-Time, Real-World Quality of Life Data Using Ecological Momentary Assessment," *MAPI Research Institute Quality of Life Newsletter*, 25, 1–3 (2000).
7. M. Jensen, L. Turner, J. Turner, and J. Romano, "The Use of Multiple-Item Scales for Pain Intensity Measurement in Chronic Pain Patients," *Pain*, 67, 35–40 (1996).
8. D. Redelmeier and D. Kahneman, "Patients' Memories of Painful Medical Treatments: Real-Time and Retrospective Evaluations of Two Minimally Invasive Procedures," *Pain*, 66, 3–8 (1996).
9. J.E. Schwartz and A.A. Stone, "Strategies for Analyzing Momentary Assessment Data," *Health Psychology*, 17, 6–16 (1998).
10. S. Shiffman and A.A. Stone, "Introduction to the Special Section: Ecological Momentary Assessment in Health Psychology," *Health Psychology*, 17, 3–5 (1998).
11. S. Shiffman, C. Elash, S. Paton, C.J. Gwaltney, J. Paty, D.B. Clark, K.S. Liu, and M.E. DiMarino, "Comparative Efficacy of 24-Hour and 16-Hour Transdermal Nicotine Patches for Relief of Morning Craving," *Addiction*, 95, 1185–1195 (2000).
12. S. Shiffman, M. Hufford, M. Hickcox, J. Paty, M. Gnys, and J.D. Kassel, "Remember That? A Comparison of Real-Time Versus Retrospective Recall of Smoking Lapses," *Journal of Consulting and Clinical Psychology*, 65, 292–300 (1997).
13. A.A. Stone, J.M. Smyth, T. Pickering, and J. Schwartz, "Daily Mood Variability: Form of Diurnal Patterns and Determinants of Diurnal Patterns," *Journal of Applied Social Psychology*, 26, 1286–1305 (1996).
14. L. Jamner, D. Shapiro, I. Goldstein, and B. Rozanne Hug, "Ambulatory Blood Pressure and Heart Rate in Paramedics: Effects of Cynical Hostility and Defensiveness," *Psychosomatic Medicine*, 53, 393–406 (1991).
15. T.W. Kamarck, S. Shiffman, L. Smithline, J.L. Goodie, H. Thompson, P. Ituarte, J.Y. Jong, V. Pro, J. Paty, J. Kassel, M. Gnys, and W. Perz, "The Diary of Ambulatory Behavioral States: A New Approach to the Assessment of Psychosocial Influences on Ambulatory Cardiovascular Activity," in *Perspectives in Behavioral Medicine: Technology and Methodology in Behavioral Medicine*, D. Krantz and A. Baum, Eds. (Lawrence Erlbaum Associates, Inc., Mahwah, NJ, 1998), pp. 163–193.
16. M.C. Okenfels, L. Porter, J. Smyth, C. Kirschbaum, D.H. Hellhammer, and A.A. Stone, "The Effect of Chronic Stress Associated with Unemployment on Salivary Cortisol: Overall Cortisol Levels, Diurnal Rhythm, and Acute Stress Reactivity," *Psychosomatic Medicine*, 57, 460–467 (1995).
17. J. Smyth, M. Okenfels, L. Port, C. Kirschbaum, D.H. Hellhammer, and A.A. Stone, "Stressors and Mood Measured on a Momentary Basis Are Associated with Salivary Cortisol Secretion," *Psychoneuroendocrinology*, 23, 353–370 (1998).
18. M. VanEck, H. Berkhof, N. Nicolson, and J. Sulon, "The Effects of Perceived Stress, Traits, Mood States, and Stressful Events on Salivary Cortisol," *Psychosomatic Medicine*, 58, 447–458 (1996).

**Saul Shiffman, PhD**, is professor of psychology (clinical and health psychology), psychiatry, and pharmacy at the University of Pittsburgh and chief science officer at *invivodata, inc.* **Michael R. Hufford, PhD**,\* is manager of scientific affairs at *invivodata*, 2100 Wharton St., Suite 505, Pittsburgh, PA 15203, (412) 390-3008, fax (412) 390-3020, e-mail: mhufford@invivodata.com.

\*To whom correspondence should be addressed.

©Reprinted from  
APPLIED CLINICAL TRIALS  
Volume 10, Number 3  
pages 42-48  
March 2001  
AN ADVANSTAR PUBLICATION  
Printed in U.S.A.

Copyright Notice Copyright by Advanstar Communications Inc. Advanstar Communications Inc. retains all rights to this article. This article may only be viewed or printed (1) for personal use. User may not actively save any text or graphics/photos to local hard drives or duplicate this article in whole or in part, in any medium. Advanstar Communications Inc. home page is located at <http://www.advanstar.com>.