International Journal of Methods in Psychiatric Research
Int. J. Methods Psychiatr. Res. 18(1): 48–57 (2009)
Published online 4 February 2009 in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/mpr.276

Computerized ambulatory monitoring in psychiatry: a multi-site collaborative study of acceptability, compliance, and reactivity

ELIZABETH I. JOHNSON,[1] OLIVIER GRONDIN,[2] MARION BARRAULT,[2,6] MALIKA FAYTOUT,[2] SYLVIA HELBIG,[3] MATHILDE HUSKY,[4] ERIC L. GRANHOLM,[5] CATHERINE LOH,[5] LOUISE NADEAU,[6] HANS-ULRICH WITTCHEN[3] & JOEL SWENDSEN[1]

1 National Center for Scientific Research (CNRS 5231), Bordeaux, France
2 Clinical Psychology Research Laboratory, University of Bordeaux 2, France
3 Institute of Clinical Psychology and Psychotherapy, Technical University Dresden, Germany
4 Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA
5 Veterans Affairs San Diego Healthcare System, and the Department of Psychiatry, University of California, San Diego, CA, USA
6 Department of Psychology, University of Montreal, Canada

Key words
ambulatory monitoring, experience sampling method, ESM, ecological momentary assessment, EMA, feasibility, reactivity

Correspondence
Joel Swendsen, National Center for Scientific Research (CNRS 5231), 146 rue Léo Saignat, 33076 Bordeaux Cedex, France.
Email: Joel.Swendsen@u-bordeaux2.fr

Received 30 October 2007;
revised 14 May 2008;
accepted 29 May 2008

Abstract

Computerized ambulatory monitoring overcomes a number of methodological and conceptual challenges to studying mental disorders, however concerns persist regarding the feasibility of this approach with severe psychiatric samples and the potential of intensive monitoring to influence data quality. This multi-site investigation evaluates these issues in four independent samples. Patients with schizophrenia ($n = 56$), substance dependence ($n = 85$), anxiety disorders ($n = 45$), and a non-clinical sample ($n = 280$) were contacted to participate in investigations using computerized ambulatory monitoring. Micro-computers were used to administer electronic interviews several times per day for a one-week period. Ninety-five percent of contacted individuals agreed to participate in the study, and minimum compliance was achieved by 96% of these participants. Seventy-eight percent of all programmed assessments were completed overall, and only 1% of micro-computers were not returned to investigators. There was no evidence that missing data or response time increased over the duration of the study, suggesting that fatigue effects were negligible. The majority of variables investigated did not change in frequency as a function of study duration, however some evidence was found that socially sensitive behaviors changed in a manner consistent with reactivity.

Introduction

Ambulatory monitoring techniques such as the Experience Sampling Method (ESM) or Ecological Momentary Assessment (EMA) have been increasingly applied over the last decade to study of a range of psychiatric disorders, including mood disorders (Myin-Germeys *et al.*, 2003; Peeters *et al.*, 2006), psychosis (Delespaul *et al.*, 2002; deVries and Delespaul, 1989; Granholm *et al.*, 2007; Kimhy *et al.*, 2006; Myin-Germeys *et al.*, 2001, 2003), personality disorders (Ebner-Priember *et al.*, 2006; Farmer *et al.*, 2004; Loewenstein *et al.*, 1987; Stein, 1996), eating disorders (Hilbert and Tuschen-Caffier, 2007; Smyth *et al.*, 2007; Stein and Corte, 2003), social anxiety (Brown *et al.*, 2007; Kashdan and Steger, 2006; Lee *et al.*, 2006) and substance use (Cooney *et al.*, 2007; Freedman *et al.*, 2006; Hopper *et al.*, 2006; Krahn *et al.*, 2005; Lukasiewicz *et al.*, 2005; Swendsen *et al.*, 2000). A principal advantage of this approach is that it allows researchers to assess symptom expression in a manner that is often inaccessible to standard hospital or laboratory protocols, and to model complex within-person processes over time. Repeated within-day assessments also reduce retrospective recall bias (Bolger *et al.*, 2003; Hormuth, 1986; Hurlburt, 1997; Scollon *et al.*, 2003) and provide ecologically valid data for better understanding the process and management of psychiatric disorders.

Despite growing enthusiasm for ambulatory monitoring, issues continue to be raised regarding the feasibility of this approach. Compliance with repeated daily assessments is by far the most frequently raised concern, and research suggests that the self-reported response times recorded in paper-based protocols are often inaccurate (Broderick *et al.*, 2003; Stone *et al.*, 2003a). Due in part to these issues, a new generation of electronic monitoring methods using micro-computers or cellular telephones have become state-of-the-art in ambulatory monitoring investigations, with response rates that are generally comparable across clinical and healthy controls groups (Ebner-Priember *et al.*, 2006; Kimhy *et al.*, 2006). However, when considering all published investigations using computerized or electronic methods, few have provided information necessary for estimating full compliance rates, including the number of patients who refuse participation once informed of the ambulatory monitoring procedures, or the percentage of patients who agree to participate but are non-compliant with the study protocol. This information is particularly rare for severe psychiatric disorders, such as schizophrenia and drug dependence, where the use of ambulatory monitoring has been hindered due to doubts regarding the capacity of patients to fully participate or to concerns that the monitoring devices may be lost or sold.

An additional concern is that intensive monitoring by repeated electronic interviews may change the frequency or nature of variables under study. Repeated within-day assessments over periods spanning days to weeks may increase awareness of behaviors or psychological states for which patients may otherwise have been inattentive. Such insight may lead to clinical change depending on the perceived benefits or consequences of behaviors, or encourage response biases as a function of the social desirability of particular responses. For example, heavy drinkers have been found to report less alcohol consumption as ambulatory monitoring studies progress (Collins *et al.*, 1998; Magneberg, 1998), and subjective reports by participants in several investigations also suggest that additional variables may be sensitive to intensive repeated assessment (Aaron *et al.*, 2004; Freedman *et al.*, 2006; Litt *et al.*, 1998). Evidence for such 'reactive effects' (Nelson, 1977) raises the possibility that ambulatory monitoring may induce clinical change or that it may, under certain conditions, constitute a therapeutic intervention in itself. While recent investigations of persons with schizophrenia or substance dependence report satisfactory overall compliance with ambulatory methods (Freedman *et al.*, 2006; Granholm *et al.*, 2007), investigations of objective or electronically-verified reactive effects have been conducted almost exclusively in samples with alcohol dependence or chronic pain (Aaron *et al.*, 2004; Hufford *et al.*, 2002; Stein and Corte, 2003; Stone *et al.*, 2003b).

In light of the increasing application of ambulatory monitoring in psychiatry, a direct comparison of compliance and reactive effects associated with this methodology is needed across a range of mental disorders. The present investigation examines these issues through a multi-site collaborative study of three psychiatric samples (patients with schizophrenia, substance dependence, anxiety disorders), and a non-clinical comparison group. The specific objectives are to: (1) establish initial study acceptance rates as well as estimate participant compliance with the ambulatory monitoring procedures; (2) examine whether compliance rates vary as a function of time in the study; (3) assess the influence of time in the study on the frequency, intensity, and associations among specific daily life variables.

Method

Samples

Outpatients treated for schizophrenia or schizoaffective disorder ($n = 56$), substance dependence ($n = 85$), anxiety

disorders (n = 45) and a non-clinical sample (n = 280) were contacted to participate in four independent investigations using computerized ambulatory monitoring. All samples were composed of adults who provided informed written consent for participation. Exclusion criteria included current substance use disorder for the anxiety and schizophrenia samples, psychotic disorder for the substance dependence and anxiety samples, and bipolar disorder for the anxiety sample.

Patients with schizophrenia (n = 44) or schizoaffective disorder (n = 12) were enrolled through a larger psychosocial treatment outcome study at the University of California, San Diego, USA and VA San Diego Health Care System. Substance dependent participants were women recruited directly from public substance abuse treatment centers in Montreal, Canada. Approximately half (55.4%) of these individuals were dependent on at least two substances, with the most frequent substances at the origin of dependence being alcohol (56.6%), cocaine (51.8%), or heroin (25.3%). Patients with anxiety disorders were recruited through treatment clinics for panic disorder and agoraphobia in Dresden, Germany (n = 23) or for social phobia in Bordeaux, France (n = 22). The non-clinical sample was recruited from French universities using a two-phase selection strategy to increase generalizability to the overall population of young adults (for a description of sampling procedures, see Husky *et al.*, 2007). Of individuals in the non-clinical sample participating in ambulatory monitoring, a normal control group (n = 82) was further screened for an absence of any lifetime mental disorder.

Procedures

All sites administered a structured diagnostic interview soliciting the Diagnostic and Statistical Manual of Mental Disorder-Fourth Edition (DSM-IV) criteria as well as a seven-day period of ambulatory monitoring concerning diverse daily life behaviors and experiences. Prior to beginning the ambulatory monitoring phase, participants were trained in how to use the micro-computer and completed practice assessments with the investigator. Participants were then given either a Palm Zire 31 or Psion 'Revo' personal digital assistant (PDA) and instructed to carry it with them for seven days and to respond to an electronic interview following each alarm signal emitted by the micro-computer. The PDAs were programmed using either a modified version of the Purdue Momentary Assessment Tool (Weiss *et al.*, 2004) or the Bordeaux Experience Sampling Program 1.0 (Grondin *et al.*, 2002). Electronic assessments were then administered four times per day for the schizophrenia sample, and five times per day for samples with substance dependence, anxiety disorders, or no psychopathology. All assessments were programmed at fixed intervals (randomized across participants) and occurred once every three hours on average between 9:00 a.m. and 9:00 p.m. for the schizophrenia sample, and between 8:00 a.m. and 11:00 p.m. for all other samples. All entries were time-stamped, providing objective data on the timeliness of participants' responses. The start day for the study was counterbalanced across the different workdays of the week in each sample. All participants were contacted by telephone approximately halfway through the assessment period to monitor and encourage compliance. In the final phase of the study, the micro-computer was returned and its databases uploaded. Financial compensation ($35 to $100) was provided to the schizophrenia, substance dependence and normal samples.

Clinical measures

Diagnostic status

Diagnoses for the schizophrenia sample were established through the Structured Clinical Interview (SCID) (First *et al.*, 1995; First and Gibbon, 2003) for DSM-IV. The Mini International Neuropsychiatric Interview Version 4.4 for DSM-IV (MINI) (Wittchen, 1994) was used to establish the diagnostic status for all participants in the Canadian and French sites (Lecrubier *et al.*, 1997; Wittchen, 1994). The German participants were assessed using a computer-assisted version of the Munich Composite International Diagnostic Interview (M-CIDI/DIA-X) (Zanarini and Frankenburg, 2001). Validity and reliability of all the interviews have been established previously (Hoyer *et al.*, 2006; Lecrubier *et al.*, 1997; Wittchen, 1994; Wittchen *et al.*, 1998; Zanarini and Frankenburg, 2001).

Compliance and ambulatory measures

Minimum compliance and response rates

In order to minimize the extent to which data were influenced by retrospective recall biases, assessments were considered valid only if completed within 15 minutes of the signal. Valid data collected for each individual was first classified according to whether it corresponded to minimum compliance criteria, defined as the provision of more than one full day of valid assessments (five valid assessments for the schizophrenia sample and six valid assessments for all other samples). This criterion was established to identify individuals who provided

minimally sufficient data for examining both within-day and between-day variation, as well as to reduce response rate estimations that may be positively biased if based only on moderate to good compliers. Response rates among individuals meeting minimum compliance were then calculated based on the total number of valid observations generated by the individual divided by the total number of programmed observations.

Behaviors, environments, and social contexts

Participants at all sites were asked to describe their current behavior, physical environment, and social company (if any) at the moment of the signal. The response options provided to participants were those established through previous ambulatory monitoring investigations (Husky *et al.*, 2007, Swendsen, 1998; Verdoux *et al.*, 2003). In order to construct identical variables, specific categories of behavior, environment or social context (e.g. being with friends, being with family) were collapsed into the broadest category used at any one site (e.g. being with friends or family). This classification rendered five broad categories each for location, behavior, and social contact. For the purpose of the present investigation, two representative variables were selected from each category based on a sufficient frequency of occurrence across all samples. These variables included being 'Alone' or 'With friends or family' (for social categories), being 'At home' or 'Outside' (for environmental categories), and being 'Inactive' or 'Shopping and household chores' (for activity categories). In addition to these general categories, 'Alcohol use' and 'Self care' (which included personal hygiene care) were selected given their increased potential to be influenced by social desirability effects.

Depressed mood

Depressed mood states were assessed in all samples using a multi-point Likert scale that asked participants to evaluate their mood at the moment each signal occurred, ranging from 'not at all depressed' to 'extremely depressed'. As the numerical range used for this scale varied by site, this variable was standardized across samples.

Overview of analyses

Data were analyzed using Hierarchical Linear and Nonlinear Modeling 6.03 (HLM 6.03; Raudenbush *et al.*, 2005) to accommodate the multi-level structure of the data, with day of the study entered as the primary predictor variable. Fatigue effects were assessed by examining whether missing data or assessment duration varied by day of the study, and reactive effects were assessed by examining whether the frequency or intensity of behaviors, activities, and psychological states varied as a function of day in the study. The effects of day of the study on associations among the daily life variables were assessed for all social, environmental, or behavioral variables that were significantly correlated with depressed mood. To examine such change, dependent variables were regressed on depressed mood as well as on the interaction between day of study and depressed mood. Bonferroni corrections were applied to adjust alpha levels for the number of multiple comparisons within each variable category. All dichotomous variables were estimated using a Bernoulli distribution in HLM. In order to aide comparisons, unstandardized regression coefficients (γ) are presented for both dichotomous and continuous variables.

Results

Study acceptance and compliance

Study acceptance, completion, and compliance rates are presented in Table 1. In general, participation rates were high for all samples and across all research phases. After ambulatory monitoring procedures were described to participants, initial study acceptance rates were 95% for the total sample, and ranged from 89% to 100% across the individual samples. The minimum compliance rate was also high (96%), and ranged from 87% for the schizophrenia sample to 100% for the sample with no lifetime mental disorder. Overall, participants responded to 78% [standard deviation (SD) = 18%] of all assessments within 15 minutes of being signaled by the PDA. The nonclinical sample had the highest response rate (83% of all programmed assessments, SD = 16%), followed by participants with substance dependence [mean (M) = 80%, SD = 18%], anxiety disorders (M = 73%, SD = 18%) and schizophrenia (M = 69%, SD = 22%). On average, only 1% of PDAs (range 0% to 2%) were lost or stolen during the ambulatory data collection phase.

Effects associated with duration of ambulatory monitoring

Fatigue effects and reactivity were examined for all participants who met the minimum compliance criteria described earlier (see Table 2 for a description of these final participating samples). Contrary to the notion that ambulatory monitoring induces salient fatigue effects, day of the study was not significantly related to missing data in any of the samples (Table 3). To examine the possibility that fatigue effects may be more likely to be found

Table 1 Study acceptance rates among contacted individuals and compliance rates among participants in ambulatory monitoring

Variable	Schizophrenia $n = 56$	Substance dependence $n = 85$	Anxiety disorder $n = 45$	Non-clinical $n = 280$	Total $n = 466$
Study acceptance[a]	96%	98%	89%	93%	95%
Minimum compliance	87%	96%	98%	100%	96%
Material loss	2%	0%	0%	2%	1%
Average response rate[b]	Mean 69% (22%)	Mean 80% (18%)	Mean 73% (18%)	Mean 83% (16%)	Mean 78% (18%)
Assessment duration (in minutes)	3.4 (2.2)	2.9 (2.5)	4.2 (2.5)	2.9 (2.2)	3.2 (2.4)

[a] Study acceptance rates are based on all contacted individuals; all other statistics are based on participants in ambulatory monitoring. For the non-clinical group, all statistics other than study acceptance are based on the normal control group ($n = 82$).
[b] Standard deviations given in parentheses. Rate is calculated among patients having minimum compliance.

Table 2 Socio-demographic characteristics of the final samples demonstrating minimum compliance with ambulatory monitoring

Variable	Schizophrenia $n = 47$	Substance dependence $n = 80$	Anxiety disorder $n = 39$	No disorder $n = 82$	Total $n = 248$
Age (in years)[a]	Mean 44.1 (10.5)	Mean 35.1 (9.6)	Mean 29.1 (8.5)	Mean 19.4 (1.4)	Mean 30.6 (11.9)
Sex (percentage female)	37%	100%	59%	67%	71%
Married or cohabiting	7.4%	46.3%	33%	3.7%	23%
Employed	0%	40.0%	39%	0%	19%

[b] Standard deviations given in parentheses.

among individuals participating intensively, these analyses were then limited to the half of the sample demonstrating the highest compliance (corresponding to an average of 82.86% of all programmed assessments). Again, day of study was not significantly related to missing observations among these individuals [$\gamma = 0.02$, standard error (SE) = 0.04, $p = 0.61$], and assessment duration significantly decreased across all samples as the study progressed.

Table 3 also presents the results of analyses examining the effect of day of study on the frequency or intensity of daily life variables, as well as on the strength of associations among significantly correlated variables. Overall, results provide little consistent evidence of reactivity. However, the frequency with which individuals in the anxiety sample reported engaging in self-care behaviors increased across the observation period ($\gamma = 0.07$, SE = 0.03, $p < 0.05$), corresponding to an increase in frequency from 19% on day one to 26% on day seven. Reports of alcohol use also declined across the observation period for individuals with substance dependence ($\gamma = -0.11$, SE = 0.04, $p < 0.05$) as well as for the overall sample ($\gamma = -0.09$, SE = 0.03, $p < 0.05$). The magnitude of this decrease from day one to day seven was 60% for the substance dependent sample and 50% for the total sample. Concerning correlations among the daily life variables, depressed mood was positively associated with being alone ($\gamma = 0.11$, SE = 0.04, $p < 0.01$), being inactive ($\gamma = 0.19$, SE = 0.04, $p < 0.01$), being at home ($\gamma = 0.12$, SE = 0.05, $p < 0.05$), and negatively associated with being in the company of family or friends ($\gamma = -0.16$, SE = 0.04, $p < 0.01$), being outside ($\gamma = -0.07$, SE = 0.03, $p < 0.01$), or shopping/household chores ($\gamma = -0.10$, SE = 0.04, $p < 0.05$). However, as shown in Table 3, day of study did

Int. J. Methods Psychiatr. Res. 18(1): 48–57 (2009). DOI: 10.1002/mpr

Table 3 Frequency or intensity of behaviors, contexts and psychological states as a function of day in study

	Schizophrenia $n = 47$		Substance dependence $n = 80$		Anxiety disorder $n = 39$		No disorder $n = 82$		Total $n = 248$	
Variable	γ	SE	γ	SE	γ	SE	γ	SE	γ	SE
Compliance										
Missing data	−0.04	0.05	0.03	0.03	0.02	0.05	0.03	0.03	0.02	0.02
Assessment duration	−0.31	0.04**	−0.13	0.03**	−0.28	0.05**	−0.18	0.03**	−0.19	0.02**
Environment and social contexts										
Alone	−0.02	0.04	0.01	0.03	−0.02	0.05	−0.05	0.04	−0.02	0.02
With friends or family	−0.03	0.05	−0.03	0.03	−0.00	0.05	0.02	0.03	−0.01	0.02
At home	−0.02	0.05	−0.02	0.03	0.09	0.06	−0.02	0.06	−0.00	0.02
Outside	−0.06	0.07	0.02	0.03	−0.06	0.10	0.02	0.06	0.01	0.03
Activity and behavior										
Inactive	−0.02	0.04	0.03	0.02	−0.03	0.05	0.10	0.04	0.03	0.02
Shopping or household chores	0.04	0.09	−0.05	0.04	0.08	0.04	−0.03	0.03	−0.01	0.02
Personal care	0.06	0.07	0.00	0.02	0.07	0.03*	0.01	0.03	0.02	0.02
Alcohol use	0.09	0.11	−0.11	0.04*	−0.05	0.07	−0.13	0.06	−0.09	0.03*
Psychological states										
Depressed mood	−0.04	0.02	−0.01	0.02	0.02	0.02	0.03	0.02	0.00	0.01
Association among variables										
Alone – depressed mood	−0.03	0.09	0.03	0.05	−0.15	0.12	0.02	0.06	−0.01	0.04
With family – depressed mood	0.06	0.10	0.01	0.05	0.10	0.12	0.07	0.06	0.05	0.04
At home – depressed mood	−0.20	0.10	−0.06	0.05	0.09	0.11	0.12	0.10	−0.03	0.04
Outside – depressed mood	0.02	0.14	0.06	0.05	−0.36	0.14	−0.12	0.10	−0.01	0.04
Inactive – depressed mood	−0.06	0.07	−0.02	0.05	−0.08	0.09	0.09	0.08	−0.01	0.03
Shopping – depressed mood	−0.09	0.15	−0.04	0.07	0.08	0.09	−0.03	0.06	−0.02	0.04

Note: *p*-values reflect Bonferroni adjustments for multiple comparisons; * $p < 0.05$, ** $p < 0.01$.

not significantly alter the magnitude of any of these relationships. Finally, all results observed for the combined sample remained unchanged when adjusted for participant gender and age.

Discussion

Ambulatory monitoring has been used in psychiatry for more than two decades (Csikszentmihalyi and Larson, 1987; deVries and Delespaul, 1989; Dijkman and deVries, 1987), but recent advances in computerized techniques have overcome important biases associated with paper-based methods. Although these novel techniques have renewed enthusiasm for the capacity of ambulatory monitoring to provide insight into the daily expression of psychiatric disorders as well as risk factors for these conditions (Compton *et al.*, 2005; Moskowitz and Young, 2006), critical evaluations of this approach have generally been limited to normal or medical samples (Broderick *et al.*, 2003; Stone *et al.*, 2003b). Moreover, its feasibility in severe psychiatric populations is largely unknown. By combining samples that used nearly identical methods of ambulatory data collection, the objective of the present investigation was to provide comprehensive information concerning study acceptance, compliance and reactive effects across a range of mental disorders. Overall, the findings provide strong support for the use of computerized ambulatory monitoring, but also suggest that reactive effects may be observed for certain variables.

The most commonly reported statistic regarding feasibility of ambulatory monitoring is the average response rate to assessments administered over the course of a study. These estimates have often been based on those individuals with relatively good compliance that were retained in final analyses (Collins *et al.*, 1998; Delespaul *et al.*, 2002; Freedman *et al.*, 2006; Green *et al.*, 2006; Krahn *et al.*, 2005; Litt *et al.*, 1998; Magneberg, 1998; Myin-Germeys *et al.*, 2001, 2003; Peeters *et al.*, 2006; Swendsen, 1998), an approach that renders it difficult to

assess overall feasibility and compliance. In examining the acceptance and participation rates across all research phases, no evidence was found indicating that the intensive monitoring protocol was a deterrent to study enrollment or participation. Very few individuals in any sample failed to provide minimal data necessary for examining within and between-day variation, and participants in each sample responded on average to most of the programmed assessments. This finding is particularly notable for individuals with schizophrenia, who responded to more than two-thirds of the electronic interviews despite considerable cognitive and functional disability. It is also important to note that investigations of illicit drug dependence have rarely applied ambulatory methods due to fears that the micro-computers may be lost, traded or sold. Drug-dependent individuals in this study constitute the largest sample to date to participate in computerized monitoring, and although half were dependent on cocaine and one-quarter on heroin, study acceptance and compliance rates were high and all monitoring devices were returned. As such, the present findings provide strong support for the feasibility of computerized ambulatory monitoring across a wide range of psychiatric disorders.

Aside from general feasibility issues, a frequently raised concern is that intensive monitoring may induce fatigue effects and decrease compliance rates over time (Feldman Barrett and Barrett, 2001; Scollon *et al.*, 2003; Stone *et al.*, 1991). Contrary to this notion, day of the study was not associated with increases in missing data for any of the samples. Moreover, analyses of the most compliant participants revealed no tendency for missing data to increase as monitoring progressed. These findings are therefore consistent with previous studies using computerized methods with pain patients (Broderick *et al.*, 2003; Stone *et al.*, 2003a), but also confirm that computerized methods do not induce salient fatigue effects in psychiatric samples. Participants in all samples nonetheless became more rapid in providing their responses over the course of the study.

A final important criticism of ambulatory monitoring is the possibility that repeated assessments may alter the nature, frequency or intensity of variables under study (Bolger *et al.*, 2003; Hufford *et al.*, 2002; Hurlburt, 1997; Litt *et al.*, 1998; Scollon *et al.*, 2003). Previous investigations using computerized methods have found little evidence of such 'reactive effects' (Aaron *et al.*, 2004; Hufford *et al.*, 2002; Stone *et al.*, 2003b), but have been based primarily on the study of pain or stress in medical populations. The current findings demonstrated that most variables were not associated with day of study, and no evidence was found suggesting that the relationships among psychological states and behaviors varied over time. However, a notable exception concerned those variables that were selected specifically for their potential sensitivity to social desirability. In particular, the frequency of reported alcohol use declined across the observation period in the combined sample as well as in the substance-dependent sample. Whether this finding reflects a change in drinking behavior attributable to awareness or treatment, or a change in motivation to conceal drinking, it is consistent with previous research demonstrating that self-reported alcohol use in heavy drinkers typically declines as ambulatory monitoring studies progress (Collins *et al.*, 1998; Magneberg, 1998). Similarly, the increase in self care behaviors over time in persons with anxiety disorders may reflect awareness of the social desirability of appearance and hygiene, or other perceptions that may otherwise motivate an individual to change the manner in which these behaviors are reported. Future investigations may consider alternative means of assessing reactivity, such as randomizing participants into groups that receive different degrees of monitoring (e.g. Stone *et al.*, 2003b).

Several methodological constraints and characteristics should be considered in formulating conclusions based on the results of this study. Most importantly, the current investigation examined issues of feasibility and reactivity related to computerized ambulatory monitoring techniques. It did not address the validity of the resulting data, which is an additional and fundamental issue in assessing the benefits and limitations of this data collection technique with psychiatric samples. Concurrent and predictive validity may be assessed through several strategies that have been applied recently in severe psychiatric populations, including the examination of expected patterns among variables collected in daily life, or the estimation of correspondence between ambulatory monitoring data and data collected through laboratory-based measures (see Freedman *et al.*, 2006; Granholm *et al.*, 2008; Johnson *et al.*, 2009; Kimhy *et al.*, 2006). It is important to note, however, that low correspondence between different methods of data collection may not necessarily suggest poor validity of ambulatory monitoring measures but rather raise broader questions concerning what should be considered the 'gold standard' in psychiatric assessment.

Additional considerations are that the samples were selected from treatment settings for the specific goals of the collaborating research teams and that the ambulatory monitoring was administered over a one-week period of time; different findings may be observed in other samples or in studies that utilize more or less intensive monitoring

protocols. Moreover, the specific diagnostic groups were often confounded by geographic location. The generalization of findings should therefore be considered in light of the cultural and clinical similarities to other samples, in particular as the application of DSM-IV criteria may not in itself guarantee the identification of homogeneous clinical populations in different countries. Finally, the results may not be generalizable to other methods of data collection, including cellular telephones. Cellular devices may provide certain advantages over micro-computers such as allowing the immediate transfer of data or facilitating the intervention of researchers and clinicians when necessary. However, such devices may also place a heavier burden on auditory memory, have reduced screen dimensions for the reading of questions, and depend on external resources that vary in reliability (such as network coverage). Compared to paper-based approaches, however, both micro-computers and cellular telephones may be considered state-of-the-art and their respective benefits should be evaluated according to the specific study objectives. The recent technological advances in ambulatory monitoring should permit its widespread and cost effective use in the years to come, both as an independent investigation technique in psychiatry and as a complement to standard clinical research protocols.

Acknowledgements

This research was supported by a National Center for Scientific Research ATIP award (Dr Swendsen), the French Mission Against Drug Abuse (Dr Nadeau), the Office of Rehabilitation Research and Development, Department of Veterans Affairs, and the National Institute of Mental Health (R01MH071410) and the UCSD Advanced Center for Interventions and Services Research (ACISR: NIMH P30MH66248).

Declaration of interests statement

The authors have no competing interests.

References

Aaron L.A., Mancl L., Turner J.A., Sawchuk C.N., Klein K.M. (2004) Reasons for missing interviews in the daily electronic assessment of pain, mood, and stress. *Pain*, **109**, 389–398, DOI: 10.1016/j.pain.2004.02.014

Bolger N., Davis A., Rafaeli E. (2003) Diary methods: capturing life as it is lived. *Annual Review of Psychology*, **54**, 579–616, DOI: 10.1146/annurev.psych.54.101601.145030

Broderick J.E., Schwartz J.E., Shiffman S., Hufford M.R., Stone A.A. (2003) Signaling does not adequately improve diary compliance. *Annals of Behavioral Medicine*, **26**, 139–148, DOI: 10.1207/S15324796ABM2602_06

Brown L.H., Silvia P.J., Myin-Germeys I., Kwapil T.R. (2007) When the need to belong goes wrong: the expression of social anhednoia and social anxiety in daily life. *Psychological Science*, **18**, 778–782, DOI: 10.1111/j.1467-9280.2007.01978.x

Collins R.L., Morsheimer E.T., Shiffman S., Paty J.A., Gnys M., Papandonatos G.D. (1998) Ecological momentary assessment in a behavior drinking moderation training program. *Experimental and Clinical Psychopharmacology*, **6**, 306–315, DOI: 10.1037/1064-1297.6.3.306

Compton W.M., Thomas Y.F., Conway K.P., Colliver J.D. (2005) Developments in the epidemiology of drug use and drug use disorders. *American Journal of Psychiatry*, **162**, 1494–1502. DOI: 10.1176/appi.ajp.162.8.1494

Cooney N.L., Litt M.D., Cooney J.L., Pilkey D.T., Steinberg H.R., Oncken C.A. (2007) Alcohol and tobacco cessation in alcohol-dependent smokers: analysis of real-time reports. *Psychology of Addictive Behaviors*, **21**, 277–286, DOI: 10.1037/0893-164X.21.3.277

Csikszentmihalyi M., Larson R. (1987) Validity and reliability of the experience-sampling method. *Journal of Nervous and Mental Disease*, **175**, 526–536, DOI: 10.1097/00005053-198709000-00004

Delespaul P., deVries M., van Os J. (2002) Determinants of occurrence and recovery from hallucinations in daily life. *Social Psychiatry and Psychiatric Epidemiology*, **37**, 97–104, DOI: 10.1007/s001270200000

deVries M.W., Delespaul P.A. (1989) Time, context, and subjective experiences in schizophrenia. *Schizophrenia Bulletin*, **15**, 233–244.

Dijkman C.I.M., deVries M.W. (1987) The social ecology of anxiety: theoretical and quantitative perspectives. *Journal of Nervous and Mental Disease*, **175**, 550–557, DOI: 10.1097/00005053-198709000-00007

Ebner-Priember U.W., Kuo J., Welch S.W., Thielgen T., Witte S., Bohus M., Linehan M.M. (2006) A valence-dependent group-specific recall bias of retrospective self-reports: a study of Borderline Personality Disorder in everyday life. *Journal of Nervous and Mental Disease*, **194**, 774–779, DOI: 10.1097/01.nmd.0000239900.46595.72

Farmer R.F., Nash H.M., Dance D. (2004) Mood patterns and variations associated with personality disorder pathology. *Comprehensive Psychiatry*, **45**, 289–303, DOI: 10.1016/ j.comppsych.2004.03.009

Feldman Barrett L., Barrett D.J. (2001) An introduction to computerized experience sampling in psychology. *Social Science Computer Review*, **19**, 175–185.

First M.B., Gibbon M. (2003) The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II). In: *Comprehensive Handbook of Psychological Assessment* (eds Hilsenroth M.J., Segal D.L.), Volume 2, pp. 134–143, John Wiley & Sons.

First M.B., Spitzer R.L., Gibbon M., Williams J.B.W. (1995) *Structured Clinical Interview for DSM-IV Axis I Disorder*

Patient Edition (SCID I/P), Patient Edition, Biometrics Research Department, New York State Psychiatric Institute.

Freedman M.J., Lester K.M., McNamara C., Milby J.B., Schumacher J.E. (2006) Cell phones for ecological momentary assessment with cocaine-addicted homeless patients in treatment. *Journal of Substance Abuse Treatment*, **30**, 105–111, DOI: 10.1016/j.jsat.2005.10.005

Granholm E., Loh C., Swendsen J.D. (2008) Feasibility and validity of computerized ecological momentary assessment in schizophrenia. *Schizophrenia Bulletin*, **34**, 507–514. DOI: 10.1093/schbul/sbm113

Green A.S., Rafaeli E., Bolger N., Shrout P.E., Reis H.T. (2006) Paper or plastic? Data equivalence in paper and electronic diaries. *Psychological Methods*, **11**, 87–105, DOI: 10.1037/1082-989X.11.1.87

Grondin O.S., Husky M.M., Chakroun N. (2002) *Bordeaux Experience Sampling Program 1.0*. Bordeaux, unpublished manuscript.

Hilbert A., Tuschen-Caffier B. (2007) Maintenance of binge eating through negative mood: a naturalistic comparison of binge eating disorder and bulimia nervosa. *International Journal of Eating Disorders*, **40**, 521–530, DOI: 10.1002/eat.20401

Hopper J.W., Su Z., Looby A.R., Ryan E.T., Penetar D.M., Palmer C.M., Lukas S.E. (2006) Incidence and patterns of polydrug use and craving for ecstasy in regular ecstasy users: An ecological momentary assessment study. *Drug and Alcohol Dependence*, **85**, 221–235, DOI: 10.1016/j.drugalcdep.2006.04.012

Hormuth S.E. (1986) The sampling of experiences in situ. *Journal of Personality*, **54**, 262–293, DOI: 10.1111/j.1467-6494.1986.tb00395.x

Hoyer J., Scholz D., Ruhl U., Wittchen H.U. (2006) Patients' feedback after computerized clinical diagnostic interviews for mental disorders. *Psychotherapy Research*, **13**, 414–421, DOI: 10.1080/10503300500485540

Hufford M.R., Shields A.L., Shiffman S., Paty J., Balabanis M. (2002) Reactivity to ecological momentary assessment: an example using undergraduate problem drinkers. *Psychology of Addictive Behaviors*, **16**, 205–211, DOI: 10.1037/0893-164X.16.3.205

Hurlburt R.T. (1997) Randomly sampling thinking in the natural environment. *Journal of Consulting and Clinical Psychology*, **65**, 941–949, DOI: 10.1037/0022-006X.65.6.941

Husky M., Mazure C., Maciejewski P., Swendsen J. (2007) A daily life comparison of sociotropy-autonomy and hopelessness theories of depression. *Cognitive Therapy and Research*, **31**, 659–676, DOI: 10.1007/s10608-006-9025-x

Johnson E.I., Barrault M., Nadeau L., Swendsen J. (2009) Feasibility and validity of computerized ambulatory monitoring in drug-dependent women. *Drug and Alcohol Dependence,* **99**, 322–326, DOI: 10.1016/j.drugalcdep.2008.06.010

Kashdan T.B., Steger M.F. (2006) Expanding the topography of social anxiety: an experience-sampling assessment of positive emotions, positive events, and emotion suppression. *Psychological Science*, **17**, 120–128, DOI: 10.1111/j.1467-9280.2006.01674.x

Kimhy D., Delespaul P., Corcoran C., Ahn H., Yale S., Malaspina D. (2006) Computerized experience sampling method (ESMc): assessing feasibility and validity among individuals with schizophrenia. *Journal of Psychiatric Research*, **40**, 231–230, DOI: 10.1016/j.jpsychires.2005.09.007

Krahn D.D., Bohn M.J., Henk H.J., Grossman J.L., Gosnell B. (2005) Patterns of urges during early abstinence in alcohol-dependent subjects. *American Journal on Addictions*, **14**, 248–255, DOI: 10.1080/10550490590949424

Lee M.R., Okazaki S., Yoo H.C. (2006) Frequency and intensity of social anxiety in Asian Americans and European Americans. *Cultural Diversity and Ethnic Minority Psychology*, **12**, 291–305, DOI: 10.1037/1099-9809.12.2.291

Lecrubier Y., Sheehan D.V., Weiller E., Amorim P., Bonora I., Sheehan K.H., Janavs J., Dunbar G.C. (1997) The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *European Psychiatry*, **12**, 224–231, DOI: 10.1016/S0924-9338(97)83296-8

Litt M.D., Cooney N.L., Morse P. (1998) Ecological momentary assessment (EMA) with treated alcoholics: methodological problems and potential solutions. *Health Psychology*, **17**, 48–52, DOI: 10.1037/0278-6133.17.1.48

Loewenstein R.J., Hamilton J., Alagna S., Reid N., deVries M. (1987) Experience sampling in the study of multiple personality disorder. *American Journal of Psychiatry*, **144**, 19–24.

Lukasiewicz M., Benyamina A., Reynaud M., Falissard B. (2005) An *in vivo* study of the relationship between craving and reaction time during alcohol detoxification using the Ecological Momentary Assessment. *Alcoholism: Clinical and Experimental Research*, **29**, 2135–2143, DOI: 10.1097/01.alc.0000191760.42980.50

Magneberg R. (1998) The effect of electronic surveillance on alcohol misuse and everyday behavior. *Addictive Behaviors*, **23**, 281–301, DOI: 10.1016/S0306-4603(97)00067-1

Moskowitz D.S., Young S.N. (2006) Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology. *Journal of Psychiatry and Neuroscience*, **31**, 13–20.

Myin-Germeys I., van Os J., Schwartz J.E., Stone A.A., Delespaul P.A. (2001) Emotional reactivity to daily life stress in psychosis. *Archives of General Psychiatry*, **58**, 1137–1144, DOI: 10.1001/archpsyc.58.12.1137

Myin-Germeys I., Peeters F., Havermans R., Nicolson N.A., deVries M.W., Delespaul P., van Os J. (2003) Emotional

reactivity to daily life stress in psychosis and affective disorder: an experience sampling study. *Acta Psychiatrica Scandinavica*, **107**, 124–131, DOI: 10.1034/j.1600-0447.2003.02025.x

Nelson R.O. (1977) Assessment and therapeutic functions of self-monitoring. In: *Progress in Behavioral Modification* (eds Hersen M., Eisler R.M., Miller P.), pp. 3–41, Academic Press.

Peeters F., Berkhof J., Delespaul P., Rottenberg J., Nicolson N.A. (2006) Diurnal mood variation in major depressive disorder. *Emotion*, **6**, 383–391, DOI: 10.1037/1528-3542.6.3.383

Raudenbush S., Bryk A., Congdon R. (2005) *HLM for Windows, Version 6.03*, Scientific Software International.

Scollon C.N., Kim-Prieto C., Diener E. (2003) Experience sampling: promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies*, **4**, 5–34, DOI: 10.1023/A:1023605205115

Smyth J.M., Wonderlich S.A., Heron K.E., Sliwinski M.J., Crosby R.D., Mitchell J.E., Engel S.G. (2007) Daily and momentary mood and stress are associated with binge eating and vomiting in bulimia nervosa patients in the natural environment. *Journal of Consulting and Clinical Psychology*, **75**, 629–638, DOI: 10.1037/0022-006X.75.4.629

Stein K.F. (1996) Affect instability in adults with a borderline personality disorder. *Archives of Psychiatric Nursing*, **10**, 32–40, DOI: 10.1016/S0883-9417(96)80084-7

Stein K.F., Corte C.M. (2003) Ecological momentary assessment of eating-disordered behaviors. *International Journal of Eating Disorders*, **34**, 349–360. DOI: 10.1002/eat.10194

Stone A.A., Kessler R.C., Haythornthwaite J.A. (1991) Measuring daily events and experiences: decisions for the researcher. *Journal of Personality*, **59**, 575–607, DOI: 10.1111/j.1467-6494.1991.tb00260.x

Stone A.A., Shiffman S., Schwartz J.E., Broderick J.E., Hufford M.R. (2003a) Patient compliance with paper and electronic diaries. *Controlled Clinical Trials*, **24**, 182–199, DOI: 10.1016/S0197-2456(02)00320-3

Stone A.A., Broderick J.E., Schwartz J.E., Shiffman S., Litcher-Kelly L., Clavanese P. (2003b) Intensive momentary report of pain with an electronic diary: reactivity, compliance and patient satisfaction. *Pain*, **104**, 343–351, DOI: 10.1016/S0304-3959(03)00040-X

Swendsen J.D. (1998) The helplessness-hopelessness theory and daily mood experience: an idiographic and cross-situational perspective. *Journal of Personality and Social Psychology*, **74**, 1398–1408, DOI: 10.1037/0022-3514.74.5.1398

Swendsen J.D., Tennen H., Carney M.A., Affleck G., Willard A., Hromi A. (2000) Mood and alcohol consumption: an experience sampling test of the self-medication hypothesis. *Journal of Abnormal Psychology*, **109**, 198–204, DOI: 10.1037/0021-843X.109.2.198

Verdoux H., Husky M., Tournier M., Sorbara F., Swendsen J.D. (2003) Social environments and daily life occurrence of psychotic symptoms: an experience sampling test in a non-clinical population. *Social Psychiatry and Psychiatric Epidemiology*, **38**, 654–661, DOI: 10.1007/s00127-003-0702-8

Weiss H.M., Beal D.J., Lucy S.L., MacDermid S.M. (2004) Constructing EMA Studies with PMAT: The Purdue Momentary Assessment Tool User's Manual, Military Family Research Institute, Purdue University.

Wittchen H.U. (1994) Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): a critical review. *International Journal of Psychiatric Research*, **28**, 57–84, DOI: 10.1016/0022-3956(94)90036-1

Wittchen H.U., Lachner G., Wunderlich U., Pfister H. (1998) Test–retest reliability of the computerized DSM-IV version of the Munich-Composite International Diagnostic Interview (M-CIDI). *Social Psychiatry and Psychiatric Epidemiology*, **33**, 568–578, DOI: 10.1007/s001270050095

Zanarini M.C., Frankenburg F.R. (2001) Attainment and maintenance of reliability of axis I and axis II disorders over the course of a longitudinal study. *Comprehensive Psychiatry*, **42**, 369–374, DOI: 10.1053/comp.2001.24556