Effects of Antidepressant Treatment on the Quality of Daily Life: An Experience Sampling Study

Daniela Q. C. M. Barge-Schaapveld, Ph.D., and Nancy A. Nicolson, Ph.D.

Background: Although some depression trials have included quality of life (QoL) as an outcome measure, assessments were retrospective and relatively infrequent. Because QoL varies in relation to everyday experience, intensive time-sampling approaches may be useful.

Method: The experience sampling method (ESM) was used to assess effects of antidepressant treatment on the quality of life, as measured from moment to moment in daily life (mQoL), and related aspects of daily experience. Primary care patients with a DSM-III-R/DSM-IV diagnosis of major depressive disorder were randomly assigned to imipramine (N = 32) or placebo (N = 31) treatment for 6 weeks, with possible prolongation to 18 weeks. A healthy control group (N = 22) provided normative data.

Results: Treatment-related increases in frequency and severity of physical complaints, including those not reported to the general practitioner as side effects, were associated with lowered mQoL; this negative association was especially strong in treatment dropouts. Despite greater clinical improvement at week 6, imipramine patients did not report greater increases than placebo patients in mean mQoL ratings. However, imipramine treatment stabilized mQoL fluctuations and led to reductions in time spent "doing nothing." Patients' decisions to prolong treatment depended on clinical improvement, mQoL changes, and specific early side effects. At 18 weeks, remitted patients still showed deficits on ESM daily life measures relative to healthy controls, even though QoL had returned to normal on retrospective measures.

Conclusion: ESM provides new insights in the effects of antidepressant treatment on daily life experiences and should therefore be considered as a supplement to conventional instruments in clinical trials.

(J Clin Psychiatry 2002;63:477–485)

Received Jan. 8, 2001; accepted Nov. 29, 2001. From the Department of Psychiatry and Neuropsychology, Maastricht University, Maastricht, the Netherlands.

This study was supported by a research grant from Solvay Pharmaceuticals B.V., Weesp, the Netherlands.

In the spirit of full disclosure and in compliance with all ACCME Essential Areas and Policies, the faculty for this CME activity were asked to complete a full disclosure statement. The information received is as follows: Drs. Nicolson and Barge-Schaapveld have received research grant support from Solvay.

The following investigators participated in the multicenter clinical trial: M. F. J. Bemelmans, M.D.; R. H. J. Brouwers, M.D.; G. Costongs, M.D.; A. C. J. Schlösser, M.D.; J. J. M. Theunissen, M.D.; D. J. M. van Aubel, M.D.; D. J. M. van der Voort, M.D., Ph.D; and M. Weijts, M.D. The authors thank P. M. L. F. Versteeg-Noteborn, B.Sc.; C. J. M. van der Voort-Duindam, B.Sc.; G. H. M. Driesen, B.Sc.; and F. J. Delespaul-van Goethem, B.A., for assistance in data collection and processing and J. Berkhof, Ph.D., for statistical support. P. A. E. G. Delespaul, Ph.D., and M. W. deVries, M.D., provided advice on various aspects of the study.

Corresponding author and reprints: Nancy A. Nicolson, Ph.D., Department of Psychiatry and Neuropsychology, PAR45, Maastricht University, P.O. Box 616, 6200 MB Maastricht, the Netherlands (e-mail: n.nicolson@sp.unimaas.nl).

Quality of life (QoL) has become an accepted outcome measure in clinical drug trials, going beyond measures of symptom reduction. The acknowledged importance of QoL is reflected in an increasing trend over the past 2 decades to use QoL measures in clinical trials in all medical subdisciplines.[1] The same period has also seen a growing interest in the effects of drug treatments on QoL in depressive disorders. There are several reasons for this interest. First, depression has been recognized as a debilitating disorder with significant impairments in many areas of daily life,[2,3] and treatment-related improvements in these areas are not adequately reflected in clinical rating scales. Another reason is that available antidepressant treatments, although comparable in terms of efficacy, may have different effects on QoL by virtue of their divergent side effect profiles.[4,5] Finally, QoL evaluations assess the global impact of aspects related to side effect profile from the viewpoint of the patient and, hence, may provide better insight into the effects of different antidepressants than efficacy measures alone.

Using the approach taken by Sanders et al.,[1] we conducted a literature search over the period 1980 to 2000 and identified 42 randomized controlled trials (RCTs) of pharmacologic treatments in unipolar depressive disorders that

included QoL assessments. Although only a small minority of these studies included a placebo arm, the 9 studies that did reported significant differences between active and placebo groups on various dimensions of QoL, such as social relationships, life satisfaction, and leisure activities,[6,7] thus supporting the hypothesis that antidepressants improve QoL in depressed patients. A number of questions remain, however.

First, although side effects of antidepressants are widely believed to have a negative impact on QoL, research findings on this relationship are relatively meager.[8] Because QoL is generally assessed at the end of treatment, when most side effects have subsided, information on the influence of early-appearing side effects is particularly scarce. Physical complaints experienced during the first week or two of treatment can contribute to poor adherence and early termination of treatment,[9,10] and it is thus important to investigate the frequency, severity, and QoL impact of early side effects as they are experienced in daily life.

Second, certain features of the QoL measures typically used in such studies place limitations on their ability to capture the full impact of the illness and its treatment on QoL. All QoL measures, whether generic (e.g., the Sickness Impact Profile)[11] or depression-specific (e.g., the Quality of Life Enjoyment and Satisfaction Questionnaire),[12] assess patients' QoL retrospectively, usually over the past week. In general, but even more so in depression research,[13] retrospective measures are vulnerable to influences of current affective state, forgetting, cognitive reframing, and other sources of bias.[14] This problem is compounded by the fact that most studies rely on single pretreatment and posttreatment measures. Both retrospective bias and infrequent measurement can reduce the reliability of QoL assessments.

A related limitation of conventional QoL measures is that they summarize average levels of well-being and thus cannot reveal how QoL varies in relation to everyday contexts and experiences. QoL is influenced by affect,[15] physical symptoms,[16] and satisfaction with daily activities,[17] all of which vary within the course of a single day. In other words, QoL has state as well as trait properties. Moreover, there is evidence that variability in QoL, from moment to moment or from day to day, is heightened in individuals most vulnerable to depression.[18] Whether pharmacologic treatment stabilizes QoL is as yet unknown, but the level of stability could prove to be a useful outcome measure.

Finally, RCTs have not yet clarified the time course over which aspects of QoL recover to "normal" levels during antidepressant treatment. Six of the identified articles[6,19–23] compared posttreatment findings with normative data; results indicated, for example, that QoL levels remained lower in clinically improved depressed patients than in a normative sample in the United States[19] or Finland.[20] However, the treatment duration in these studies ranged from 6 to 12 weeks, which might have been too short for QoL aspects to normalize. In addition, in only 1 study[21] was a group of healthy subjects recruited for direct comparison with patients; in the other studies, differences in demographic characteristics, location, and time could have biased the comparisons with the controls.

The Current Study

To redress some of the limitations of earlier studies, the current study of the effects of pharmacologic treatment on QoL used an intensive time-sampling approach. The experience sampling method (ESM) was developed to measure subjective experience over time and across situations.[24] During their normal daily routines, subjects complete frequent self-reports with respect to their mood, activity, location, and social context in response to signals they receive at frequent but unpredictable intervals over a period of several days or a week. Advantages of the method include greater ecological validity and increased reliability, due to the repeated measurements and reduction in retrospective bias through time sampling. Among many other applications, ESM has been used in research on psychiatric and psychosomatic disorders,[25] including studies of QoL.[26–28] To date, however, applications of ESM in the context of a clinical trial have been limited to a handful of studies: our initial study in a small sample of depressed outpatients treated with amitriptyline or fluvoxamine,[29] a pilot investigation of the effect of relaxation training on asthma symptoms,[30] and 2 recent studies that used ESM to evaluate the effects of pharmacologic treatments on nicotine craving and withdrawal symptoms during smoking cessation.[31,32] Results obtained so far in ESM studies suggest that applying an intensive time-sampling approach in the context of a clinical trial may provide a better and more reliable understanding of the effects of treatment on the quality of daily life experience. In a previous analysis of ESM data collected in the same sample of depressed outpatients prior to treatment,[27] we found that self-reports of general well-being ("momentary QoL," or mQoL) showed considerable variation over the course of a day, with significantly greater within-person variability in depressed patients than in healthy controls. In both groups, positive and negative mood states, enjoyment of daily activities, and physical complaints had independent influences on mQoL. As expected, depressed patients had more negative (or less positive) scores on all of these variables than controls. In a different sample of depressed outpatients, we found that positive mood increased, negative mood decreased, and patterns of time use changed (time spent doing household chores increased and time spent in passive leisure decreased) following successful treatment with fluvoxamine or amitriptyline.[29]

The current study was designed to extend these findings in a number of ways: (1) Inclusion of assessments during the first week of treatment in the longitudinal design of the

study allowed investigation of the relationship between early side effects, mQoL, and subsequent termination of treatment. Patients with side effects were expected to experience lower levels of mQoL than patients without side effects (assuming that the positive effects of active treatment would not yet be apparent) and, as a consequence, be more likely to withdraw from treatment. (2) In contrast to an earlier ESM study,[29] the current study was placebo controlled, so that observed changes in daily measures of QoL during treatment could be attributed with more confidence to the active drug, in this case, imipramine, and not only to spontaneous remission, placebo effects, or the effects of the ESM procedure itself. Imipramine was chosen as the active treatment because imipramine and related tricyclics remain the "gold standard" for antidepressant efficacy.[33] (3) To determine whether mQoL stabilizes after active treatment, we compared intraindividual variability in mQoL from pretreatment to posttreatment in imipramine versus placebo groups. In the same sample, prior to treatment,[27] the magnitude of intraindividual variability was unrelated to the severity of depression, which would suggest that any decrease in variability found is more likely to reflect an actual change in a subject's experience of QoL than an increased reliability of the methodology in less depressed individuals. (4) To obtain a better picture of the effects of extended treatment, additional sampling was done at 18 weeks. Comparison of ESM measures in patients treated for 18 weeks with those of healthy controls provided a measure of the efficacy of antidepressant treatment in normalizing QoL.

METHOD

Study Design

The study was conducted in a primary care setting, in a sample of depressed outpatients presenting for treatment to their general practitioners (GPs). During an initial baseline week, participants received no treatment of any form. Thereafter, patients were randomly assigned to twice-daily, double-blind treatment with either imipramine (starting dose of 50 mg/day, increased to 200 mg/day over the first week of treatment) or placebo (starting with 1 capsule per day, increased to 4 capsules over the first week of treatment). In cases of intolerance, the dose could be decreased to either 100 mg/day of imipramine or 2 placebo capsules per day. After 6 weeks, a decision was made concerning prolongation of double-blind treatment to 18 weeks, based on consensus between physician and patient. The possibility that some patients might thereby continue placebo treatment for up to 18 weeks was considered acceptable for the following reasons: (1) informed consent, (2) relatively mild symptomatology, (3) intensive clinical monitoring, (4) freedom to drop out, and (5) poststudy access to standard treatment and continued follow-up. The study was approved by a medical ethics committee.

Subjects

Eighty-three patients with a DSM-III-R/DSM-IV diagnosis of current major depressive disorder were recruited in 8 primary care practices in the Netherlands (for details concerning diagnosis and screening, see Barge-Schaapveld et al.[27]). Inclusion criteria were age between 18 and 65 years, a score at entry of ≥ 18 on the 17-item Hamilton Rating Scale for Depression (HAM-D),[34] and a score ≥ 4 on the Clinical Global Impressions scale (CGI).[35] Exclusion criteria included current use of psychotropic medications and major medical disorders. All subjects gave written informed consent.

After exclusion of subjects who either did not meet all of the above criteria (N = 9) or had insufficient data during the ESM baseline sampling period (N = 11), 63 patients were included in the current analysis: 32 in the imipramine group and 31 in the placebo group (for details about exclusion criteria and differences between included and excluded subjects, see Barge-Schaapveld et al.[27]). One placebo patient withdrew consent and 3 imipramine patients dropped out due to adverse events during the first week of treatment. Of the 63 patients randomly assigned to treatment, 49 (77.8%) completed the first 6 weeks of treatment. Subsequently, 35 patients agreed to prolong treatment; 22 imipramine and 13 placebo subjects (95.7% vs. 50.0% of patients at week 6, respectively; $\chi^2 = 12.5$, $p < .001$). Eighty percent of these patients (N = 28; 17 imipramine and 11 placebo) completed the prolongation phase through 18 weeks.

A control group of 22 healthy individuals, similar to the patient groups in sociodemographic characteristics, was recruited to provide reference values within normal range for ESM measures (for details, see Barge-Schaapveld et al.[27]).

Assessments

Clinical monitoring: efficacy and tolerability. The HAM-D was administered by the treating physician (GP) at screening, baseline, and weeks 1, 2, 4, 6, 10, 14, and 18. At each visit, the GP asked whether the patient had experienced any unusual or unwanted signs or symptoms since the last visit. If so, start and stop dates were recorded as well as the severity of the symptom.

Retrospective QoL measures. Questionnaires completed with reference to the past week provided retrospective measures of QoL. At the end of each sampling period, subjects were asked to rate the quality of their life on a 100-mm visual analogue scale (QoL VAS)[36] and to indicate their global life satisfaction on the Satisfaction With Life Scale (SWLS),[37] which consists of 5 items (rated from 1, "strongly disagree," to 7, "strongly agree"). The validated Dutch version of the SF-36 Health Survey (SF-36)[38] was completed by a subset of 42 depressed subjects at some point during the baseline week.

ESM monitoring. After a briefing session in which a research nurse explained the procedure, patients completed

ESM reports in response to signals from a programmed wristwatch at 10 semirandom intervals per day, between 7:30 a.m. and 10:30 p.m. ESM was conducted during 6 consecutive days during the baseline week, again during the last 3 days of the first week of treatment, and for 6 consecutive days in the sixth week of treatment. Finally, patients who remained in the treatment prolongation phase completed 6 consecutive days of ESM in week 18. An ESM self-report was considered valid if completed within 15 minutes after the signal. The protocol allowed only patients with a minimum of 30 valid reports to be randomly assigned to treatment, so that by definition all subjects fulfilled this criterion in the baseline week. ESM compliance remained acceptable throughout the study, with more than 80% of subjects in each of the repeated sampling periods completing valid reports in response to half or more of all signals.

Subjects were asked to complete ESM items with reference to the moment at which they received a signal (a beep). At each beep, subjects rated their momentary QoL (mQoL) in response to the question "In general, how is it going with you right now?" Responses ranged from –3 ("very bad") to + 3 ("very good"). In addition, subjects rated current mood and enjoyment of the present activity on 7-point scales. Mood items were combined into separate scales for positive affective states (PA; items energetic, cheerful, satisfied, alert, calm, enthusiastic, strong, and happy) and negative affective states (NA; items hostile, depressed, tense, lonely, anxious, insecure, guilty, harried, and irritable). Physical complaints included 3 common side effects of imipramine: dry mouth, dizziness, and nausea.[39] In response to an open question, subjects provided descriptions of their current activities; these were later coded and collapsed into 8 categories. Only the category "doing nothing" was used in the current analysis. The ESM procedure, questionnaire, and derived measures are described in greater detail elsewhere.[27,28]

Definition of Measures

For each sampling week, ESM measures were averaged across all valid ESM records for each subject to obtain mean levels of mQoL, PA, NA, and enjoyment of activities. At baseline, mean mQoL was significantly associated with the other ESM measures (PA: $r = 0.55$, NA: $r = –0.67$, enjoyment of activities: $r = 0.46$; $N = 63$; all p values $< .001$, 1-tailed tests). Mean mQoL also showed positive associations with all retrospective measures (mQoL with QoL VAS: $r = 0.42$, $p < .001$; mQoL with SWLS: $r = 0.33$, $p < .01$; mQoL with SF-36 mental subscale: $r = 0.29$, $p < .05$; 1-tailed tests). These correlations support the construct validity of mQoL. Amount of time spent doing nothing was defined as the percentage of valid ESM records in which "doing nothing" was specifically reported as the current activity. Mean severity of each ESM complaint (i.e., dry mouth, dizziness, and nausea) was obtained by averaging across all ESM records for each subject per sampling week. The frequency of each ESM complaint was defined as the percentage of records in which the complaint was reported. Patients who showed an increase from pretreatment baseline in either the mean severity or the frequency of a complaint were considered to have experienced a side effect of treatment (ESM side effect). Side effects reported to the GP (GP side effects) as occurring in the same period during which ESM side effects were recorded were identified according to indicated start and stop dates.

Statistical Analyses

Descriptive statistics are presented as means and standard deviations. Differences in continuous measures between groups over time were tested with analysis of variance (ANOVA) for repeated measures. Comparisons between groups at specific points in time were performed with 2-sample t tests for continuous variables and with chi-square tests for categorical variables. The kappa statistic was calculated to assess agreement between GP and ESM side effect reports. Pearson correlations were used to assess the relationship between early-appearing side effects and changes in mQoL. To test the hypothesis that intraindividual variability in mQoL would stabilize after active treatment, we used multilevel regression analysis[40]; this method takes into account the 3-level hierarchical structure of the ESM dataset, in which beep level measures are nested within days, and days within subjects. We compared estimated within-subject variance components (beep level and day level)[27] for imipramine and placebo patients at baseline and at week 6, and for controls. Unless otherwise noted, statistical tests were 2-tailed, with an alpha level of .05.

RESULTS

Patient Characteristics

Patients ranged in age from 25 to 59 years (mean = 43.4 years). The majority were women (73% [N = 46]) and married (68% [N = 43]). Most had a regular job (44% [N = 28]) or were housewives (25% [N = 16]). There were no significant differences between the 2 treatment groups on either sociodemographic characteristics or initial HAM-D ratings (imipramine: 24.0 ± 3.5, placebo: 23.5 ± 2.6; $t = –0.6$, $df = 61$, NS). Patients randomly assigned to imipramine (N = 32) were less likely to have had a previous episode of depression than those randomly assigned to placebo (N = 31) (25.0% [N = 8] vs. 61.3% [N = 19], $\chi^2 = 8.5$, $p < .01$).

Side Effects of Treatment

Differences between GP and ESM reports. More than 75% of all side effects reported to the GP occurred in the first week of treatment. Among the side effects most fre-

Table 1. Percentages of Patients With Specific Side Effects Present in the First Week of Treatment[a]

	% of Patients With				% Agreement Between GP and ESM Measures (N = 59)
	GP Side Effects		ESM Side Effects		
Side Effect	Imipramine (N = 29)	Placebo (N = 30)	Imipramine (N = 29)	Placebo (N = 30)	
Dry mouth	41.4*	16.7	86.2**	46.7	59.3
Nausea	20.7*	3.3	41.4	36.7	62.7
Dizziness	13.8	10.0	72.4*	46.7	49.2

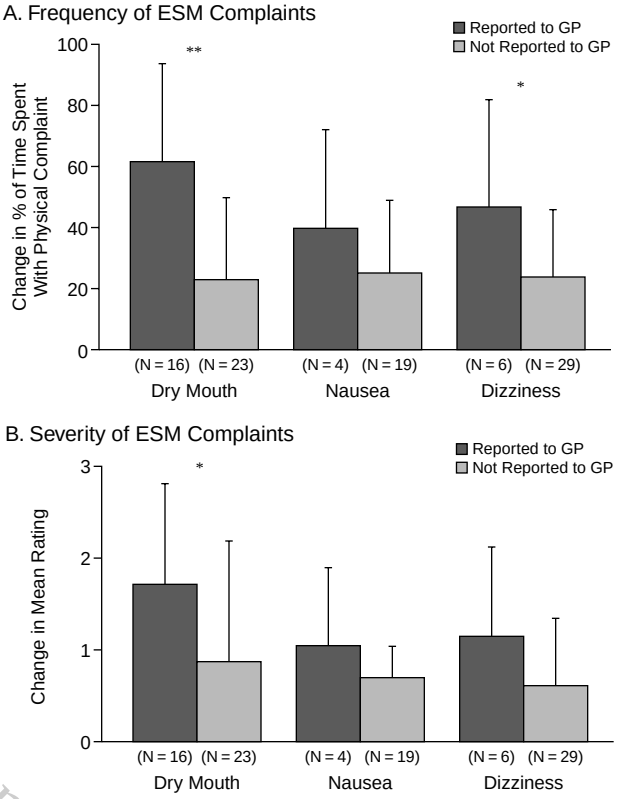
[a]Differences between imipramine and placebo groups (χ^2 tests): *p < .05, **p < .001. Abbreviations: ESM = experience sampling method, GP = general practitioner.

quently reported to the GP during the 3 days at the end of the first week of treatment (corresponding to the ESM sampling period) were dry mouth, nausea, and dizziness, with all except dizziness reported by significantly more imipramine than placebo patients (Table 1). Of the 3 complaints assessed with ESM, dry mouth and dizziness were reported by significantly more imipramine than placebo patients.

The overall degree of agreement between GP and ESM assessments of side effects was surprisingly low (kappa values < 0.30). However, in at least half of the cases (see last column of Table 1), GP ratings and ESM reports were in agreement on the presence or absence of side effects. In general, more patients reported ESM side effects than GP side effects, with increased dizziness, for example, experienced by 5 times as many subjects according to ESM reports (35 patients with ESM vs. 7 with GP side effects). In general, ESM side effects were also reported as GP side effects when an increase from baseline occurred in the percentage of time spent with a complaint (Figure 1A). Dry mouth was also more likely to be reported to the GP when a significant increase from baseline had occurred in the ESM-rated severity (Figure 1B). However, patients with both ESM and GP side effects showed no greater decreases in mQoL than patients with ESM side effects only (dry mouth: F = 1.8, nausea: F = 2.2, dizziness: F = 0.0, all NS); in other words, a decrease in mQoL was no reason for patients to report side effects to the GP.

ESM side effects, mQoL, and early termination. To clarify the clinical relevance of the ESM side effects in the first treatment week, we examined their relationship to subsequent treatment dropout. Of the 59 patients in the first week of treatment, 10 did not complete the entire 6-week treatment period (6 imipramine and 4 placebo subjects). Associations between mQoL and both frequency and severity of complaints were greater in the dropouts than in completers (Table 2). In addition, dropouts tended to report greater mean decreases in mQoL from baseline to week 1 than patients who subsequently completed 6 weeks of treatment (–0.3 ± 0.9 vs. 0.0 ± 0.6, F = 2.1, p = .15). This suggests that the impact of ESM side effects—including those not reported to the GP—on mQoL led patients to withdraw from treatment.

Figure 1. Change From Baseline to Week 1 in ESM Physical Complaints[a]



[a]Patients who experienced an increase in ESM complaints and reported the same side effect to the GP are contrasted to patients who did not report ESM complaints as side effects to the GP. Tests of group differences: *p < .05, **p < .001. Abbreviations: ESM = experience sampling method, GP = general practitioner.

Changes at 6 Weeks

Clinical efficacy, retrospective QoL, and mQoL. At 6 weeks, both the imipramine and placebo groups showed clinical improvement (repeated-measures ANOVA, time effect: F = 194.8, p < .001), with HAM-D scores declining to a mean of 8.9 ± 6.2 and 12.5 ± 6.3, respectively. The imipramine group showed significantly greater improvement than the placebo group (group-by-time effect: F = 4.9, p < .05). Despite greater clinical improvement, imipramine patients did not report larger QoL increases than placebo patients on either the QoL VAS (mean change: 27.3 ± 21.5 vs. 21.0 ± 28.9, F = 0.7, NS) or mQoL (mean change: 0.7 ± 0.7 vs. 0.5 ± 1.0, F = 0.7, NS). However, imipramine patients did show greater increases in life satisfaction than placebo patients (mean increase in SWLS score: 6.7 ± 7.8 vs. 1.6 ± 6.3, F = 6.3, p < .05).

mQoL variability. Active treatment was expected to decrease intraindividual fluctuations in mQoL. Results of the multilevel regression analysis supported this hypothesis. At week 6, variances in mQoL at both beep and day levels had significantly decreased from baseline in the

Table 2. Correlations at Week 6 Between ESM Side Effects (severity and frequency) and Changes in mQoL (from week 1) in Treatment Completers (N = 49) Versus Dropouts (N = 10)[a]

	Side Effect Severity		Side Effect Frequency	
Side Effect	Completers	Dropouts	Completers	Dropouts
Dry mouth	–0.16	–0.84*	0.02	–0.51
Nausea	–0.27	–0.22	–0.29*	–0.39
Dizziness	–0.16	–0.20	–0.08	–0.20

[a]Abbreviations: ESM = experience sampling method, mQoL = momentary quality of life.
*p < .05.

Table 3. Intraindividual Variance Components for mQoL in the Multilevel Regression Model[a]

	Variance Estimate	
Group	Beep Level	Day Level
Depressed		
Imipramine		
Baseline	0.57	0.30
Week 6	0.22	0.10
Placebo		
Baseline	0.58	0.31
Week 6	0.36	0.26
Controls	0.28	0.20

[a]For further explanation of how the model was estimated, see Barge-Schaapveld et al.[27] Abbreviation: mQoL = momentary quality of life.

combined patient groups (likelihood ratio tests, p < .0001). As shown in Table 3, active treatment had a stronger stabilizing effect, with greater decreases in beep level (p < .001) and day level (p < .05) variances in imipramine than in placebo patients. At week 6, mQoL remained relatively unstable in placebo patients, as evidenced by significantly greater beep level variance in placebo patients than in controls (p < .001). In contrast, mQoL in the imipramine group had stabilized by week 6, with both beep level (p < .001) and day level (p < .01) variance estimates even smaller than in healthy controls.

Other daily life measures. By week 6, significant changes from baseline in PA (repeated-measures ANOVA, time effect: F = 16.6, p < .001) and NA (F = 19.5, p < .001), but not in enjoyment of the current activity (F = 0.01, NS), were observed in both groups. The 2 treatment groups did not differ significantly from each other on any of these measures (group-by-time effect for PA: F = 0.33, NA: F = 0.21, and enjoyment of activities: F = 0.84; all NS).

Time spent "doing nothing." Previous ESM studies showed a link between inactivity and low mQoL.[26,27] The percentage of patients who reported "doing nothing" as their current activity at some point during ESM sampling decreased significantly from baseline to week 6 in the imipramine (87.5% [28/32] vs. 43.5% [10/23], respectively; $\chi^2 = 12.1$, p < .001) but not in the placebo group (74.2% [23/31] vs. 61.5% [16/26], respectively; $\chi^2 = 1.1$, NS). To compare the magnitude of these changes, we identified the percentage of patients per treatment group who reported some inactivity at baseline but no inactivity at week 6; this percentage was twice as high in the imipramine as in the placebo group (52.2% [12/23] vs. 26.9% [7/26], respectively; $\chi^2 = 3.3$, p = .07). In other words, more imipramine than placebo patients reduced their level of inactivity to 0% over the course of 6 weeks. There was, however, no significant difference in the reduction in percentage of time spent doing nothing between the imipramine and placebo groups (–2.1 ± 6.6% vs. –3.7 ± 6.0%; F = 0.8, NS).

Changes With Prolonged Treatment (18 weeks)

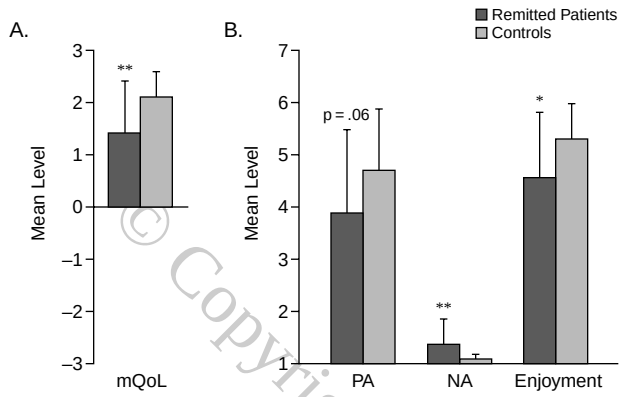
Factors related to prolongation. As expected, the 35 patients who entered the treatment prolongation phase had shown a greater decrease in mean HAM-D score than patients who stopped at 6 weeks (–15.7 ± 5.1 vs. –6.5 ± 6.5; F = 27.8, p < .001). Prolongers had also experienced a greater increase from baseline in levels of mQoL than nonprolongers (0.8 ± 0.8 vs. 0.1 ± 0.8; F = 7.9, p < .01). Contrary to expectation, prolongers had experienced greater increases than nonprolongers in the ESM complaint "dry mouth" during the first week of treatment (increase in frequency: 31.0 ± 39.2% vs. 7.4 ± 32.6%; F = 5.9, p < .05; increase in severity: 1.0 ± 1.4 vs. 0.2 ± 1.2; F = 4.7, p < .05). This was specifically true for prolongers relative to nonprolongers taking imipramine (frequency: F = 5.3, severity: F = 5.1; both p values < .05), but not those taking placebo (frequency: F = 0.4, severity: F = 0.1; NS). Experience of other side effects did not differentiate prolongers from nonprolongers. Taken together, these results suggest that degree of clinical improvement, mQoL changes, and specific side effects may all have influenced the decision to prolong treatment.

Normalization. There was no difference in clinical severity between the imipramine and placebo groups at week 18 (mean HAM-D score: 4.9 ± 6.7 vs. 3.8 ± 5.3; F = 0.03, NS). Of the 28 patients who completed treatment, 23 were considered to be clinically remitted (HAM-D score ≤ 7) at 18 weeks (imipramine: 14/17, placebo: 9/11; $\chi^2 = 0.0$, NS). To determine whether QoL measures achieved normal levels, measures in the remitted patients were contrasted with those in the healthy control group.

Remitted patients and controls had similar ratings on the QoL VAS (75.0 ± 19.1 vs. 76.5 ± 15.6; t = 0.3, df = 43, NS) and SWLS (25.6 ± 7.3 vs. 27.5 ± 6.8; t = 0.9, df = 43, NS). However, remitted patients still had significantly lower mQoL than controls (Figure 2). Only 10 of the 23 patients in remission (5 in each treatment arm) had mean mQoL levels at or above the lower bound (mQoL = 1.86) of the 95% confidence interval for healthy controls. Mean levels of other daily life measures in remitted patients were also significantly different from those of healthy controls (see Figure 2).

Prior to treatment, a greater percentage of patients than controls had reported at times to be doing nothing. In ad-

Figure 2. Mean Levels of the ESM Variables (A) mQoL and (B) PA, NA, and Enjoyment of Activities in Remitted Patients (N = 23) and Control Subjects (N = 22) at Week 18[a]



[a]mQoL: t = 2.8, df = 31. PA: t = 1.9, df = 43. NA: t = –2.8, df = 24. Enjoyment: t = 2.3, df = 35. Abbreviations: ESM = experience sampling method, mQoL = momentary quality of life, NA = negative affective states, PA = positive affective states.
*p < .05.
**p < .01.

dition, depressed subjects had spent a greater percentage of time doing nothing than controls. At 18 weeks, a similar percentage of remitted patients and controls reported doing nothing at some point during the sampling period (47.8% [11/23] vs. 50.0% [11/22], respectively; $\chi^2 = 0.2$, NS). Moreover, the percentage of time spent doing nothing no longer differed between the 2 groups (4.9 ± 9.0% vs. 2.2 ± 3.3%, respectively; t = –1.3, df = 28, NS).

Depressed patients had reported both greater frequency and intensity of physical complaints, even before treatment, than healthy controls. At 18 weeks, significantly more remitted patients than healthy controls still reported dizziness (30.4% [7/23] vs. 4.5% [1/22]; $\chi^2 = 5.2$, $p < .05$) or dry mouth (65.2% [15/23] vs. 27.3% [6/22]; $\chi^2 = 6.5$, $p < .01$) at some point during ESM sampling. None of these remitted patients had reported dizziness, and only 6 (26.1%) had reported dry mouth as a GP side effect in that week. Dry mouth and dizziness also occurred with a higher frequency (dry mouth: 42.3 ± 46.4 vs. 2.2 ± 6.1; t = –4.1, df = 23, $p < .001$; dizziness: 12.5 ± 26.4 vs. 0.4 ± 2.0; t = –2.2, df = 22, $p < .05$) and were rated as more severe (dry mouth: 2.0 ± 1.3 vs. 1.0 ± 0.1; t = –3.6, df = 22, $p < .01$; dizziness: 1.2 ± 0.5 vs. 1.0 ± 0.0; t = –2.1, df = 22, $p < .05$) in the remitted patients than in the control group.

In summary, despite normalized QoL on global retrospective measures, remitted patients still differed from healthy controls in most aspects of daily QoL.

DISCUSSION

Most antidepressant drug studies rely primarily on clinician measures such as the HAM-D. However, self-rating scales can provide important additional information for therapy evaluation as they reflect the patient's personal experience of illness and recovery.[41] In the current study, intensive ESM monitoring in the context of daily life revealed effects of depression and antidepressant treatment on well-being, mood states, physical complaints, enjoyment of activities, and patterns of time use—information that could not have been obtained with conventional instruments. More specifically, this study provides new information concerning early side effects, changes in daily experience associated with treatment and with clinical improvement, intraindividual variability in the state of well-being, and normalization of daily experience during sustained treatment.

As expected, more imipramine than placebo patients reported side effects in the first week of treatment; imipramine side effects were also more frequent and more severe. Only a small percentage of patients who showed an increase in specific physical complaints on ESM measures reported these as side effects in the same period to the GP. ESM side effects were associated with decrements in mQoL, and patients who showed strong negative associations were overrepresented among subsequent treatment dropouts. It is important to note that the study was not designed to determine whether ESM-reported side effects might better predict treatment dropout than reports to the GP. According to the standard protocol for assessing adverse events in clinical trials, side effects that lead to treatment discontinuation must be registered as GP side effects. This means, by definition, that there is a coupling between GP side effects and subsequent dropout. Meaningful comparison of the 2 side effect measures is therefore difficult. However, the lack of any association between mQoL and side effects reported to the GP suggests that clinicians were unaware of side effects of treatment that had a negative impact on patients' daily QoL.

Whether or not patients report side effects to the GP also depends to some extent on factors present at the time of the assessment[42] and personality characteristics of the patient.[43] On the basis of the current results, we suggest that such problems can be reduced by obtaining self-reports from patients in real time, in their everyday environments, and by asking patients to rate physical complaints (often present prior to any treatment) instead of side effects. In future studies, ESM could be useful in describing the side effect profiles of different antidepressant drugs and their relative impacts on QoL.

The current findings also point to the potential usefulness of ESM measures in understanding treatment adherence. Since the decision to prolong treatment was left up to the patient and the GP, it is not surprising that prolongers showed greater clinical improvement and greater increases in mQoL than nonprolongers. It is noteworthy, however, that prolongers, specifically those taking imipramine, had experienced more frequent and severe

dry mouth (a typical imipramine side effect) in the first week of treatment than nonprolongers. Some authors have questioned the integrity of the double-blind procedure[44]: if patients can correctly guess the treatment they are receiving due to experienced side effects, they may be more likely to expect positive results. We speculate that these positive expectations might have led patients in our study to opt for a prolongation of treatment.

At 6 weeks, imipramine patients showed no greater improvements than placebo patients on the ESM measures of mQoL, mood measures, and enjoyment of activities, even though the active treatment did result in greater clinical improvement on the HAM-D. Imipramine was more effective than placebo in decreasing mQoL within-subject variability and time spent in inactivity. Both the relatively small number of patients in the sample (despite the large number of observations per subject) and the choice of antidepressant may have contributed to the subtle nature of the observed differences between imipramine and placebo treatments. Placebo-controlled studies that have reported significant effects of antidepressant treatment on QoL outcome measures[45–47] included at least 100 patients per treatment group. Furthermore, although imipramine remains a "gold standard" for clinical efficacy, other RCTs have been unable to demonstrate an advantage of this drug over placebo in QoL outcomes (see, for example, Philipp et al.[48]).

At 18 weeks, ESM measures still differed in clinically remitted patients from those of healthy individuals, even though QoL had returned to normal on retrospective measures. Normalization in daily life appears to require more time than results obtained with conventional measures would suggest. It is also possible that the remaining differences in the experience of daily life reflect either "scars" from the depressive episode or trait-like aspects, which may in turn predispose remitted patients to subsequent episodes. The investigation of daily QoL measures in future recurrence-prevention studies may shed more light on these different possibilities. Even though most of the completers were in clinical remission at 18 weeks, only a minority had achieved normal levels of daily QoL by that time, suggesting that the currently used criteria for response in depression may not be sufficient to provide evidence of normalization in daily life. These results underscore the need to rethink the current reliance on clinical scales as the sole measure of treatment success in depression.[49]

Aspects of the study design place some limits on the conclusions. First, we assessed momentary QoL with a single item. Although the observed moderate correlations between average mQoL scores and retrospective QoL measures (see Method) support the construct validity of mQoL, a multi-item measure would have psychometric advantages in future studies.[50] Furthermore, multilevel regression results as well as the pattern of correlations among ESM measures indicate that mQoL represents more than mood alone.[27] Second, patients provided ESM data during 4 discrete sampling periods throughout the treatment. Continuous ESM sampling could yield more conclusive information about improvement curves, but the research burden for subjects might become unacceptably high. Future studies should weigh the option of sampling less frequently per day over longer periods of time; evidence that continuous long-term sampling may be feasible comes from an ESM study of migraine patients in which ESM reports were obtained 6 times a day for 10 weeks.[51] Third, the decision-making process for prolonging treatment beyond 6 weeks (based on the consensus reached by GP and patient), although similar to real-life clinical practice, makes it difficult to interpret the findings at 18 weeks.

Experience to date shows that ESM is feasible in the context of a clinical trial. Depressed patients, with the possible exception of the most severely depressed,[52] are willing and able to comply with repeated sampling periods.[27,29] Although, in this particular study, the observed advantages of antidepressant treatment over placebo on QoL measures were subtle, the potential usefulness of ESM measures in understanding treatment compliance, termination, and outcome has been highlighted. In light of the growing interest in the role of QoL in the course and outcome of depressive disorders, we believe that more widespread use of time-sampling methods as a supplement to conventional approaches will prove useful for clinicians as well as researchers.

Drug names: amitriptyline (Elavil and others), fluvoxamine (Luvox and others).

Disclosure of off-label usage: The authors have determined that, to the best of their knowledge, no investigational information about pharmaceutical agents has been presented in this article that is outside U.S. Food and Drug Administration–approved labeling.

REFERENCES

1. Sanders C, Egger M, Donovan J, et al. Reporting on quality of life in randomised controlled trials: bibliographic study. BMJ 1998;317: 1191–1194
2. De Lisio G, Maremmani I, Perugi G, et al. Impairment of work and leisure in depressed outpatients: a preliminary communication. J Affect Disord 1986;10:79–84
3. Wells KB, Stewart A, Hays RD, et al. The functioning and well-being of depressed patients: results from the Medical Outcomes Study. JAMA 1989;262:914–919
4. Anderson IM, Tomenson BM. The efficacy of selective serotonin reuptake inhibitors in depression: a meta-analysis of studies against tricyclic antidepressants. J Psychopharmacol 1994;8:238–249
5. Song F, Freemantle N, Sheldon TA, et al. Selective serotonin reuptake inhibitors: meta-analysis of efficacy and acceptability. BMJ 1993;306: 683–687
6. Kocsis JH, Zisook S, Davidson J, et al. Double-blind comparison of sertraline, imipramine, and placebo in the treatment of dysthymia: psychosocial outcomes. Am J Psychiatry 1997;154:390–395
7. Lydiard RB. Sertraline vs amitriptyline and quality of life: a double-blind, placebo-controlled study [abstract]. Neuropsychopharmacology 1994;10 (suppl 3, pt 2):S163
8. Healy D. The three faces of the antidepressants: a critical commentary on the clinical-economic context of diagnosis. J Nerv Ment Disord 1999;187: 174–180

9. Blackwell B. Antidepressant drugs: side effects and compliance. J Clin Psychiatry 1982;43(11, sec 2):14–21
10. Cookson J. Side-effects of antidepressants. Br J Psychiatry 1993;163 (suppl 20):20–24
11. Sechter D, Troy S, Paternetti S, et al. A double-blind comparison of sertraline and fluoxetine in the treatment of major depressive episode in outpatients. Eur Psychiatry 1999;14:41–48
12. Leinonen E, Skarstein J, Behnke K, et al. Efficacy and tolerability of mirtazapine versus citalopram: a double-blind, randomized study in patients with major depressive disorder. Nordic Antidepressant Study Group. Int Clin Psychopharmacol 1999;14:329–337
13. Atkinson M, Zibin S, Chuang H. Characterizing quality of life among patients with chronic mental illness: a critical examination of the self-report methodology. Am J Psychiatry 1997;154:99–105
14. Diener E. Assessing subjective well-being: progress and opportunities. Soc Ind Res 1994;31:103–157
15. Corrigan PW, Buican B. The construct validity of subjective quality of life for the severely mentally ill. J Nerv Ment Dis 1995;183:281–285
16. Pyne JM, Patterson TL, Kaplan RM, et al. Assessment of the quality of life of patients with major depression. Psychiatr Serv 1997;48:224–230
17. Goethe JW, Fischer EH. Functional impairment in depressed inpatients. J Affect Disord 1995;33:23–29
18. Gable SL, Nezlek JB. Level and instability of day-to-day psychological well-being and risk for depression. J Pers Soc Psychol 1998;74: 129–138
19. Walker V, Streiner DL, Novosel S, et al. Health-related quality of life in patients with major depression who are treated with moclobemide. J Clin Psychopharmacol 1995;15(suppl 2):60S–67S
20. Lonnqvist J, Sintonen H, Syvalahti E, et al. Antidepressant efficacy and quality of life in depression: a double-blind study with moclobemide and fluoxetine. Acta Psychiatr Scand 1994;89:363–369
21. Ravindran AV, Anisman H, Merali Z, et al. Treatment of primary dysthymia with group cognitive therapy and pharmacotherapy: clinical symptoms and functional impairments. Am J Psychiatry 1999;156:1608–1617
22. Heiligenstein JH, Ware JE, Beusterien KM, et al. Acute effects of fluoxetine versus placebo on functional health and well-being in late-life depression. Int Psychogeriatr 1995;7(suppl):125–137
23. Miller IW, Keitner GI, Schatzberg AF, et al. The treatment of chronic depression, pt 3: psychosocial functioning before and after treatment with sertraline or imipramine. J Clin Psychiatry 1998;59:608–619
24. Csikszentmihalyi M, Larson R. Validity and reliability of the Experience Sampling Method. J Nerv Ment Dis 1987;175:526–537
25. deVries MW. The Experience of Psychopathology: Investigating Mental Disorders in Their Natural Settings. Cambridge, Mass: Cambridge University Press; 1992
26. Barge-Schaapveld DQCM, Nicolson NA, Delespaul PAEG, et al. Assessing daily quality of life with the Experience Sampling Method. In: Katschnig H, Freeman H, Sartorius N, eds. Quality of Life in Mental Disorders. Chichester, England: John Wiley & Sons; 1997:95–107
27. Barge-Schaapveld DQCM, Nicolson NA, Berkhof J, et al. Quality of life in depression: daily life determinants and variability. Psychiatr Res 1999; 88:173–189
28. deVries MW, Delespaul PAEG, Nicolson NA. Experience Sampling Method (ESM). In: Salek S, ed. Compendium of Quality of Life Instruments. Chichester, England: John Wiley & Sons; 1999:2L:6, 2L:6a
29. Barge-Schaapveld DQCM, Nicolson NA, Gerritsen van der Hoop R, et al. Changes in daily life experience associated with clinical improvement in depression. J Affect Disord 1995;34:139–154
30. Smyth JM, Soefer MH, Hurewitz A, et al. The effect of tape-recorded relaxation training on well-being, symptoms, and peak expiratory flow rate in adult asthmatics: a pilot study. Psychol Health 1999;14:487–501
31. Shiffman S, Johnston JA, Khayrallah M, et al. The effect of bupropion on nicotine craving and withdrawal. Psychopharmacology (Berl) 2000;148: 33–40
32. Shiffman S, Elash CA, Paton SM, et al. Comparative efficacy of 24-hour and 16-hour transdermal nicotine patches for relief of morning craving. Addiction 2000;95:1185–1195
33. Gumnick JF, Nemeroff CB. Problems with currently available antidepressants. J Clin Psychiatry 2000;61(suppl 10):5–15
34. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56–62
35. Guy W. ECDEU Assessment Manual for Psychopharmacology. US Dept Health, Education, and Welfare publication (ADM) 76-338. Rockville, Md: National Institute of Mental Health; 1976:59–70
36. Andrews FM, McKennell AC. Measures of self-reported well-being: their affective, cognitive, and other components. Soc Indicators Res 1980;8: 127–155
37. Diener E, Emmons RA, Larsen RJ, et al. The Satisfaction With Life Scale. J Pers Assess 1985;49:71–75
38. Aaronson NK, Acquadro C, Alonso L, et al. International Quality of Life Assessment (IQOLA) project. Qual Life Res 1992;1:349–351
39. Cassano GB, Toni C, Petracca A, et al. Adverse effects associated with the short-term treatment of panic disorder with imipramine, alprazolam or placebo. Eur Neuropsychopharmacol 1994;4:47–53
40. Snijders TAB, Bosker RJ. Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. London, England: SAGE Publications; 1999
41. Möller HJ. Rating depressed patients: observer- vs self-assessment. Eur Psychiatry 2000;15:160–172
42. Anderson RB, Testa MA. Symptom distress checklists as a component of quality of life measurement: comparing prompted reports by patient and physician with concurrent adverse event reports via the physician. Drug Info J 1994;28:89–114
43. Davis C, Ralevski E, Kennedy SH, et al. The role of personality factors in the reporting of side effect complaints to moclobemide and placebo: a study of healthy male and female volunteers. J Clin Psychopharmacol 1995;15:347–352
44. Antonuccio DO, Danton WG, DeNelsky GY, et al. Raising questions about antidepressants. Psychother Psychosom 1999;68:3–14
45. Lydiard RB, Stahl SM, Hertzman M, et al. A double-blind, placebo-controlled study comparing the effects of sertraline versus amitriptyline in the treatment of major depression. J Clin Psychiatry 1997;58:484–491
46. Dubini A, Bosc M, Polin V. Noradrenaline-selective versus serotonin-selective antidepressant therapy: differential effects on social functioning. J Psychopharmacol 1997;11:S17–S23
47. Lenderking WR, Tennen H, Nackley JF, et al. The effects of venlafaxine on social activity level in depressed outpatients. J Clin Psychiatry 1999; 60:157–163
48. Philipp M, Kohnen R, Hiller KO. Hypericum extract versus imipramine or placebo in patients with moderate depression: randomised multicentre study of treatment for eight weeks. BMJ 1999;319:1534–1538
49. Rush AJ. Assessing outcome in practice: a paradigm shift? Curr Opin Psychiatry 1996;9:1–2
50. Diener E. Subjective well-being. Psychol Bull 1984;95:542–575
51. Honkoop PC, Sorbi MJ, Godaert GL, et al. High-density assessment of the IHS classification criteria for migraine without aura: a prospective study. Cephalalgia 1999;19:201–206
52. Kraan H, Meertens H, Hilwig M, et al. Selecting measures, diagnostic validity and scaling in the study of depression. In: deVries MW, ed. The Experience of Psychopathology: Investigating Mental Disorders in Their Natural Settings. Cambridge, Mass: Cambridge University Press; 1992: 324–338

For the CME Posttest for this article, see pages 543–544.