Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Psychiatric Research 39 (2005) 415–420
JOURNAL OF PSYCHIATRIC RESEARCH
www.elsevier.com/locate/jpsychires

Stability of DSM-IV criterion symptoms for major depressive disorder

Kelly L. Minor, Jennifer E. Champion, Ian H. Gotlib *

Department of Psychology, Stanford University, Jordan Hall, Building 420, Stanford, CA 94305-2130, USA

Received 14 July 2004; received in revised form 2 November 2004; accepted 10 November 2004

Abstract

Given the chronic and recurrent nature of major depressive disorder (MDD), it is important to understand whether specific symptoms are stable over time or vary over the course of the disorder. This is the first longitudinal investigation examining the stability of the nine criterion symptoms of depression, as specified in the DSM-IV, among diagnosed depressed adults who were not recovered at follow-up. In this study, participants were assessed twice, ten months apart, with the structured clinical interview for DSM-IV, and stability of the nine criterion symptoms of MDD was examined. Findings indicate strong stability in individuals' symptom profiles. Among individuals who were clinically depressed at both assessments, there were no statistically significant fluctuations in specific symptoms endorsed. Changes in symptom endorsement among individuals who no longer met diagnostic criteria for MDD at Time 2 were attributable to reduced severity (i.e., number of symptoms) rather than to inconsistency of symptom endorsement. These results indicate that depressed individuals experience essentially the same pattern of specific symptoms over the course of a year. Variation in clinical course is likely to be attributable more to fluctuations in overall severity than to changes in specific symptoms of depression.

Keywords: Major depressive disorder; Recurrent depression; Course; Symptoms; Longitudinal; Reliability

1. Stability of criterion symptoms for major depressive disorder

Although the prevalence (Kessler et al., 2003; Regier et al., 1998) and recurrence (Belsher and Costello, 1988; Boland and Keller, 2002; Keller et al., 1992; Keller and Shapiro, 1981; Mueller et al., 1996) of major depressive disorder (MDD) have been established, the stability of specific symptoms experienced by an individual over the course of the illness is unknown. Based on the DSM-IV criteria (American Psychiatric Association, 1994), there are over 100 unique combinations of symptoms that warrant a diagnosis of MDD. The heterogeneity of symptom profiles *among* individuals is well documented (Blazer et al., 1988); it is uncertain, however, whether there is also significant variability of symptoms *within* individuals over time. It is possible, for example, that patients demonstrate variability in the specific symptoms that they experience over the course of a major depressive episode (MDE), or from one depressive episode to the next, even when controlling for episode severity. It is equally plausible, of course, that an individual's propensity to manifest particular symptoms of depression remains stable over the course of the disorder.

Change in individual symptoms of depression across separate assessments has been assessed in three reports. In the first report, Young et al. (1990) examined symptom concordance in 108 adults diagnosed with recurrent unipolar depression across two discrete MDEs over the

* Corresponding author. Tel.: +1 650 725 9216; fax: +1 650 725 5699.
E-mail address: gotlib@psych.stanford.edu (I.H. Gotlib).

0022-3956/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jpsychires.2004.11.001

course of five years. Symptom concordance, as measured by the presence or absence of each of 12 depressive symptoms [1] at the two time points, was found to be no greater than chance. Importantly, concordance was considerably increased when episode severity was controlled for by selecting a subsample of the 18 participants who experienced the same number of symptoms at both episodes. In a second study, Roberts et al. (1995) used data from the Oregon Adolescent Depression Project (OADP) to measure concordance of depressive symptoms in adolescents across their first and second MDEs. To control for severity of depression, 29 participants who endorsed an equivalent number of DSM-III-R symptoms at both episodes were examined.In contrast to Young et al.'s (1990) findings with adults, concordance was low for seven of the nine criterion symptoms (concordance was higher for sad mood and anhedonia). In a final report, Lewinsohn et al. (2003) used data from the OADP sample to assess symptom stability across MDEs in adolescents during their transition to young adulthood. Results indicated low concordance for the nine criterion symptoms, both from the first to the second MDE ($n = 224$), and from the second to the third MDE ($n = 79$). Two points are noteworthy here. First, unlike prior reports (Roberts et al., 1995; Young et al., 1990), Lewinsohn et al.'s (2003) study did not control for depressive severity. Second, in contrast to the findings reported earlier (Roberts et al., 1995) with a subset from the OADP sample, the chance-corrected concordance of sad mood and anhedonia across episodes was low. Taken together, these findings suggest that symptom constellation and depressive severity are variable within individuals over the course of depression.

To date, no study has examined symptom stability within the course of a depressive episode. Given the chronic and recurrent nature of MDD, it is important to understand whether specific symptoms of depression are stable over time or, alternatively, whether they vary over the course of depression – an issue that has significant implications for both the pharmacological and psychological treatment of this disorder. This is the first longitudinal investigation to examine the stability of the nine criterion symptoms of depression, as specified in the DSM-IV, among adults ($N = 71$) who met diagnostic criteria for MDD at initial assessment (Time 1) and who had not fully recovered at reassessment (Time 2) approximately ten months later. At Time 2, 36 participants maintained diagnoses of MDD, while 35 participants met criteria for partial remission from depression (MDD-PR). At Time 2 the participants in this latter group did not meet either diagnostic criteria for MDD or criteria for remission from depression. Certainly, the MDD-PR participants will, by definition, have fewer symptoms at Time 2 than they did when they met criteria for MDD at Time 1. Nevertheless, the inclusion of these two outcome groups in the present study affords an opportunity to examine the stability of symptoms as a function of the severity of depression.

2. Method

2.1. Participants

Participants were 71 adults diagnosed at Time 1 with major depressive disorder (MDD) who had not remitted at follow-up (Time 2) assessment. All participants were recruited from three outpatient clinics at Stanford University or were self-referred from the community in response to flyers posted in the community or newspaper and on-line media advertisements. The sample consisted of fluent English speakers, aged 18–60. At Time 1, participants met full diagnostic criteria for MDD using the structured clinical interview for DSM-IV (First et al., 1996). Exclusion criteria were lifetime history of psychotic symptomatology, current diagnosis of panic disorder or social phobia, history of substance or alcohol abuse/dependence (past six months), impaired mental status or mental retardation. All participants gave informed consent and were paid for their participation. This study complied with the APA ethical standard for treatment of human subjects and was approved by the Institutional Review Board at Stanford University.

At Time 2, approximately 10 months ($M = 9.6$; $SD = 3.3$) after the initial assessment, 36 participants maintained diagnoses of MDD. The remainder ($n = 35$) met criteria for major depressive disorder in partial remission (MDD-PR; Frank et al., 1991). More specifically, they no longer met diagnostic criteria for MDD, yet reported experiencing more than two symptoms at a moderate degree or greater when questioned week-by-week about the presence of all nine DSM-IV depression symptoms during the eight weeks prior to the interview. The two groups of participants did not differ significantly with respect to the length of the follow-up interval, $t(69) = 1.02$, $p > .30$.

2.2. Materials

Structured clinical interview for DSM-IV (SCID-I). This highly reliable (Skre et al., 1991; Williams et al., 1992) structured interview facilitates diagnosis of psychiatric patients according to DSM-IV criteria. Interviewers underwent extensive training (First et al., 1995) before starting data collection, yielding κ's ranging from

[1] Items that were endorsed by nearly all participants (e.g., depressed mood) were automatically discarded by Young et al. (1990) because they do not inform the probabilistic model. "Pervasive loss of interest or pleasure" was rephrased so that pervasiveness was not necessary for endorsement; this item demonstrated high concordance ($\kappa = 1.0$) upon correcting for chance and severity among unipolar MDD patients.

.92 to 1.0 for a diagnosis of MDD (Gotlib et al., 2004; Kasch et al., 2002; Rottenberg et al., 2002a,b,c).

The *Global assessment of functioning scale* (GAF, Axis V, DSM-IV) was used to assess overall level of psychological, social and occupational functioning (American Psychiatric Association, 1994). Ratings are made on the basis of information obtained during the SCID-I interview. For reliability purposes, an independent blinded rater randomly selected 14 taped interviews and made GAF ratings. Interrater reliability for the GAF was high, $r = .92$.

2.3. Procedure

Clinical assessment. All participants were administered the SCID-I twice. All interviews were video- or audiotaped, reviewed for discrepancies shortly after completion, and discussed at weekly consensus meetings in order to reduce reliability drift.

Statistical analysis. Responses to the SCID-I (Module A) were coded dichotomously to reflect endorsement of each of the nine DSM-IV symptoms of depression. We used several methods to assess symptom change. First, categorical variables were based on symptom endorsement at Time 1 relative to Time 2 that indicated symptom gain (i.e., absent at Time 1, present at Time 2), symptom loss (i.e., present at Time 1, absent at Time 2), or symptom maintenance (i.e., present at both Time 1 and Time 2; absent at both Time 1 and Time 2). Second, to summarize the amount of change across all symptoms for each participant, *proportion change* was calculated as the total number of symptoms demonstrating change divided by the number of different symptoms endorsed across test sessions (maximum of 9). Finally, to assess the stability of each participant's symptom profile, *concordance* was calculated as the sum total of symptoms showing stability at Time 1 relative to Time 2, divided by 9. Concordance reflects all symptoms that are stable, including symptoms that are present or absent at both Time 1 and Time 2.

Nonparametric statistical analyses for two related dichotomous variables (i.e., McNemar tests) were used to detect statistically significant changes in symptom endorsement from Time 1 to Time 2 using the chi-square and binomial distributions where appropriate. Data regarding concordance and proportion change were analyzed using *t*-tests.

3. Results

3.1. Demographic and clinical characteristics

Characteristics at the initial assessment are presented in Table 1 for all participants, and separately for the two outcome groups (MDD and MDD-PR). At Time 1 the MDD outcome group was older than the MDD-PR outcome group, $t(69) = 2.76$, and had lower GAF Scores, $t(69) = 3.42$, both $ps < .01$, although the scores of both groups were in the 50–60 range.

3.2. Descriptive statistics regarding symptom severity

To assess severity of symptoms beyond diagnostic categorization, we conducted separate paired-samples *t*-tests for each outcome group assessing change from Time 1 to Time 2 with respect to general indicators of symptom severity: GAF Scores, sum total of endorsed

Table 1
Participant characteristics

Variable	Time 1	Time 2 diagnosis	
	MDD ($N = 71$)	MDD ($n = 36$)	MDD-PR ($n = 35$)
Female, *N* (%)	49 (69)	23 (63.9)	26 (74.3)
Caucasian, *N* (%)	50 (70.4)	29 (80.6)	21 (60)
Single marital status, *N* (%)	47 (66.2)	22 (61.1)	25 (71.4)
Age, *M* (SD)[a,*]	35.2 (11.1)	38.6 (11.6)	31.7 (9.4)
Income level, *M* (SD)[b]	3.2 (1.6)	3.5 (1.6)	2.9 (1.6)
Education attainment, *M*(SD)[c]	6.7 (1.4)	6.6 (1.5)	6.8 (1.2)
Follow-up interval (months), *M* (SD)	9.6 (3.3)	9.2 (2.9)	10.0 (3.6)
In treatment *N* (%)[d]	42 (59.2)	23 (63.9)	19 (54.3)
MDD symptom count at Time 1, *M* (SD)	6.4 (1.2)	6.4 (1.3)	6.3 (1.2)
GAF score at Time 1, *M* (SD)[*]	53.8 (7.7)	50.9 (7.7)	56.8 (6.7)
Age at initial onset of dep. symptoms, *M* (SD)	17.7 (9.6)	18.1 (11.5)	17.5 (7.1)
Recurrent MDEs, *N* (%)	61 (85.9)	31 (86.1)	30 (85.7)

[a] Range = 18–58 (MDD-PR) and 18–60 (MDD).
[b] Equivalent to categorical income of $25–50K annually.
[c] An average education level of 6.7, on our ordinal scale of 1–8, is equivalent to having completed some college.
[d] Although this study is naturalistic in design, information about treatment was obtained from each participant.
[*] $p < .01$ for *t*-tests comparing MDD and MDD-PR outcome groups.

Table 2
Symptom frequencies and global assessment of functioning scores at Time 1 and Time 2 according to diagnostic outcome at Time 2

SCID criterion symptom	Time 2 diagnosis			
	MDD (n = 36)		MDD-PR (n = 35)	
	Time 1	Time 2	Time 1	Time 2
Depressed mood	0.86 (0.35)	0.81 (0.40)	0.86 (0.36)	0.06 (0.24)*
Anhedonia	0.86 (0.35)	0.81 (0.40)	0.86 (0.36)	0.26 (0.44)*
Weight/appetite	0.64 (0.49)	0.58 (0.50)	0.63 (0.49)	0.37 (0.49)
Sleep disturbance	0.67 (0.48)	0.64 (0.49)	0.66 (0.48)	0.29 (0.46)*
Psychomotor	0.42 (0.50)	0.56 (0.50)	0.46 (0.51)	0.09 (0.28)*
Energy loss	0.81 (0.40)	0.75 (0.44)	0.91 (0.28)	0.37 (0.49)*
Worthlessness/guilt	0.67 (0.48)	0.67 (0.48)	0.66 (0.48)	0.17 (0.38)*
Concentration/decision	0.83 (0.38)	0.75 (0.44)	0.94 (0.24)	0.46 (0.51)*
Suicidal ideation	0.67 (0.48)	0.58 (0.50)	0.31 (0.47)	0.11 (0.32)*
# of Symptoms, M (SD)	6.42 (1.25)	6.14 (1.69)	6.29 (1.20)	2.17 (1.87)***
GAF score, M (SD)	51.03 (7.86)	50.56 (10.87)	56.77 (6.73)	63.46 (13.16)**

Note. GAF = Global assessment of functioning scale. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ for within-group t-tests comparing Time 1 with Time 2 data.

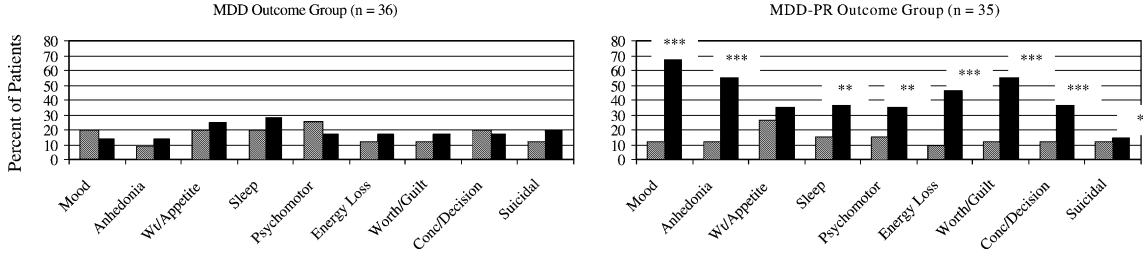

Fig. 1. Percent of patients endorsing clinically significant symptom change from Time 1 to Time 2. Histogram bars represent symptoms gained (i.e., striped) versus symptoms lost (i.e., solid black) at Time 2. All patients were diagnosed with MDD at initial assessment; at follow-up, half of the sample maintained diagnoses of MDD (left figure), while the remainder met criteria for MDD in partial remission (MDD-PR; right figure). *$p < .05$, **$p < .01$, ***$p < .001$.

symptoms per participant, and symptom frequencies (i.e., proportion of participants per group endorsing each symptom). Means and standard deviations for these variables are presented in Table 2. The MDD outcome group showed no statistically significant changes in these variables from Time 1 to Time 2 (all $ps > .05$). As expected, the MDD-PR group changed significantly from Time 1 to Time 2 with respect to mean number of symptoms, $t(34) = 10.10$, $p < .001$, and GAF Scores, $t(34) = 2.79$, $p < .01$.

3.3. Change at the symptom level

Results for the full sample indicated that depressed individuals demonstrate significant symptom change across a ten-month interval for seven of the nine criterion symptoms: depressed mood, $\chi^2(1, N = 71) = 21.03$, $p < .001$; anhedonia, $\chi^2(1, N = 71) = 15.61$, $p < .001$; sleep disturbance, $\chi^2(1, N = 71) = 4.69$, $p < .05$; energy loss, $\chi^2(1, N = 71) = 13.79$, $p < .001$; worthlessness/guilt, $\chi^2(1, N = 71) = 7.76$, $p < .01$; problems with concentration/decision making, $\chi^2(1, N = 71) = 12.03$, $p < .001$; and suicidality, exact significance [2] (2-tailed), $p = .04$. Two symptoms did not show statistically significant variability in endorsement over time: change in weight/appetite, χ^2 $(1, N = 71) = 2.70$; and psychomotor disturbance, $\chi^2(1, N = 71) = 1.53$, both $ps > .05$. Because the number of symptoms present at Time 2 differed significantly between outcome groups, we conducted separate analyses for these groups. As presented in Fig. 1, there were no statistically significant fluctuations in any of the nine criterion symptoms at Time 2 for the MDD group, all $ps > .30$. In contrast, the MDD-PR group showed statistically significant change scores for all symptoms except for change in weight/appetite. As is evident in Fig. 1, and as would be expected by the criteria used to assign participants to this group, this variation is attributable to loss of symptoms.

[2] This item did not satisfy the assumptions necessary for the χ^2 test of significance. The exact test is considered to be a reliable indicator of significance regardless of the size, balance, or distribution of the data.

3.4. Proportion change

The average rates of proportion change were .40 for the MDD group (Mdn = .38, SD = .23, SEM = .04) and .74 for the MDD-PR group (Mdn = .72, SD = .23, SEM = .04). Again, as expected, the MDD-PR group demonstrated significantly greater proportion change than did the MDD group, $t(69) = 6.41$, $p < .001$.

3.5. Concordance

The average concordance was .65 for the MDD group (Mdn = .67, SD = .21, SEM = .03) and .44 for the MDD-PR group (Mdn = .44, SD = .21, SEM = .04). Not surprisingly given the findings reported above, the MDD group exhibited significantly greater concordance of their symptoms than did the MDD-PR group, $t(69) = 4.26$, $p < .001$.

4. Discussion

This study examined the stability over time of the nine criterion symptoms of major depressive disorder. In a sample of 71 individuals who were diagnosed with MDD and who had not remitted approximately 10 months later, we found significant fluctuations for seven of the nine criterion symptoms of major depression. A more refined analysis of the data, however, yielded a different pattern of results, and helps to elucidate the conditions under which symptoms of depression are more and less stable. Symptom stability was significantly higher when participants were grouped at Time 2 according to the severity of their depressive symptoms. More specifically, the 36 participants who were diagnosed with MDD at both assessments exhibited striking stability with respect to specific symptoms endorsed, proportion of DSM-IV symptoms endorsed, symptom concordance, and depressive severity (i.e., mean number of symptoms, GAF Scores). In contrast, there was much less stability of depressive symptoms among MDD-PR participants at Time 2. As expected, despite the fact that these individuals were experiencing a number and severity of depressive symptoms sufficient to prevent them from being classified as fully remitted from depression, symptom fluctuations in this group were attributable to improvement (i.e., symptom loss) at Time 2. This finding is consistent with results reported recently in a sample of adolescents during their transition to adulthood: fluctuations in symptoms in that sample were primarily a function of low symptom concordance among those participants with the fewest depressive symptoms.

Our findings represent strong evidence that variation in clinical course is more likely to be attributable to fluctuations in the severity of depressive symptoms than to changes in the specific pattern of symptoms experienced by the patient. Given the limited, and mixed, findings from previous studies regarding the stability of symptoms of depression, the results of the present investigation provide important information concerning the stability of symptoms of depression for any given patient. The present results indicate that depressed individuals are likely to experience similar symptoms over the long-term course of a depressive episode.

The present findings also suggest that individuals with relatively chronic depression have a stable predisposition to manifest a specific profile of symptoms over the course of their affective disorder. Several lines of investigation assessing vulnerability factors for depression have focused on elucidating aspects of this disorder that may be stable as individuals experience, and then recover from, depressive episodes. For example, whereas level of interpersonal dysfunction has been found to be relatively independent of depressive state (Gotlib and Lee, 1989; Joiner, 2002), some characteristics of cognitive and biological functioning appear to ebb and flow in concert with severity of depression (Gotlib and Cane, 1987; Ingram and Siegle, 2002; Thase et al., 2002). It remains for future research to begin to examine the nature of the relation between various vulnerability factors for depression and specific symptom profiles or configurations of this disorder.

The current study is unique in several respects. First, we assessed symptom stability more comprehensively than has been the case in previous studies. Second, we carefully controlled for symptom severity and did so across a wider spectrum (i.e., MDD, MDD-PR) than has been reported in previous investigations. Third, the current sample is significantly larger than those included in earlier studies that have examined symptom stability controlling for severity (e.g., Roberts et al., 1995; Young et al., 1990). And finally, rather than relying on participants' retrospective reports of symptom endorsement, the current sample is composed of only those participants who were experiencing significant symptoms of depression at both diagnostic assessments.

We should comment here on two potential limitations of the current study that might have led to inflated estimates of symptom concordance among the participants. First, some of the interviewers were not blind to the participants' diagnostic status at Time 1, although in this context it is important to note that inter-rater reliability was high. Second, the test–retest interval of ten months was arguably short. Whether greater fluctuation in symptoms would have been found had a longer follow-up interval been used is a question for future investigation. More specifically, prospective studies with longer follow-up assessments are required to determine whether, controlling for

severity, there is symptom stability both within and across major depressive episodes. Nevertheless, the current findings strongly indicate that depressed adults experience essentially the same symptoms over a ten-month course of depression, and underscore the fact that depressive severity requires careful consideration with regard to examining and treating specific symptoms of depression over the course of the disorder.

References

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. Washington (DC): American Psychiatric Publishing; 1994.

Belsher G, Costello CG. Relapse after recovery from unipolar depression: a critical review. Psychological Bulletin 1988;104:84–96.

Blazer D, Swartz M, Woodbury M, Manton KG, et al. Depressive symptoms and depressive diagnoses in a community population: Use of a new procedure for analysis of psychiatric classification. Archives of General Psychiatry 1988;45:1078–84.

Boland RJ, Keller MB. Course and outcome of depression. In: Gotlib IH, Hammen CL, editors. Handbook of depression. New York (NY): Guilford Press; 2002.

First MB, Gibbon M, Spitzer RL, Williams JBW. User's guide for the structured clinical interview for DSM IV Axis I disorders (SCID I, Version 2.0, October 1995, Final Version).

First MB, Gibbon M, Spitzer RL, Williams JBW. Structured clinical interview for DSM-IV. Washington (DC): American Psychiatric Association; 1996.

Frank E, Prien RF, Jarrett RB, Keller MB, et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder: Remission, recovery, relapse, and recurrence. Archives of General Psychiatry 1991;48:851–5.

Gotlib IH, Cane DB. Construct accessibility and clinical depression: a longitudinal investigation. Journal of Abnormal Psychology 1987;96:199–204.

Gotlib IH, Krasnoperova E, Yue DN, Joormann J. Attentional biases for negative interpersonal stimuli in clinical depression. Journal of Abnormal Psychology 2004;113:121–35.

Gotlib IH, Lee CM. The social functioning of depressed patients: a longitudinal assessment. Journal of Social and Clinical Psychology 1989;8:223–37.

Ingram RE, Siegle GJ. Contemporary methodological issues in the study of depression: not your father's Oldsmobile. In: Gotlib IH, Hammen CL, editors. Handbook of depression. New York (NY): Guilford Press; 2002.

Joiner Jr TE. Depression in its interpersonal context. In: Gotlib IH, Hammen CL, editors. Handbook of depression. New York (NY): Guilford Press; 2002.

Kasch KL, Rottenberg J, Arnow BA, Gotlib IH. Behavioral activation and inhibition systems and the severity and course of depression. Journal of Abnormal Psychology 2002;111:589–97.

Keller MB, Lavori PW, Mueller TI, Endicott J, et al. Time to recovery, chronicity, and levels of psychopathology in major depression: a 5-year prospective follow-up of 431 subjects. Archives of General Psychiatry 1992;49:809–16.

Keller MB, Shapiro RW. Major depressive disorder. Initial results from a one-year prospective naturalistic follow-up study. Journal of Nervous Mental Disease 1981;169:761–8.

Kessler RC, Berglund P, Demler O, et al. The epidemiology of major depressive disorder: results from the national comorbidity survey replication (NCS-R). Journal of the American Medical Association 2003;289:3095–105.

Lewinsohn PM, Petit JW, Joiner Jr TE, Seeley JR. The symptomatic expression of major depressive disorder in adolescents and young adults. Journal of Abnormal Psychology 2003;112:244–52.

Mueller TI, Keller MB, Leon AC, et al. Recovery after 5 years of unremitting major depressive disorder. Archives of General Psychiatry 1996;53:794–9.

Regier DA, Rae DS, Narrow WE, Kaelber CT, Schatzberg AF. Prevalence of anxiety disorders and their comorbidity with mood and addictive disorders. British Journal of Psychiatry 1998;173:24–8.

Roberts RE, Lewinsohn PM, Seeley JR. Symptoms of DSM-III-R major depression in adolescence: evidence from an epidemiological survey. Journal of the American Academy of Child and Adolescent Psychiatry 1995;34:1608–17.

Rottenberg J, Gross JJ, Wilhelm FH, Najmi S, Gotlib IH. Crying threshold and intensity in major depressive disorder. Journal of Abnormal Psychology 2002a;111:302–12.

Rottenberg J, Kasch KL, Gross JJ, Gotlib IH. Sadness and amusement reactivity differentially predict concurrent and prospective functioning in major depressive disorder. Emotion 2002b;2:135–46.

Rottenberg J, Wilhelm FH, Gross JJ, Gotlib IH. Respiratory sinus arrhythmia as a predictor of outcome in major depressive disorder. Journal of Affective Disorders 2002c;71:265–72.

Skre I, Onstad S, Torgersen S, Kringlen E. High interrater reliability for the structured clinical interview for DSM-III–R axis I (SCID-I). Acta Psychiatrica Scandinavica 1991;84:167–73.

Thase ME, Jindal R, Howland RH. Biological aspects of depression. In: Gotlib IH, Hammen CL, editors. Handbook of depression. New York (NY): Guilford Press; 2002.

Williams JB, Gibbon M, First MB, Spitzer RL, et al. The structured clinical interview for DSM-III–R (SCID): II multisite test–retest reliability. Archives of General Psychiatry 1992;49:630–6.

Young MA, Fogg LF, Scheftner WA, Fawcett JA. Concordance of symptoms in recurrent depressive episodes. Journal of Affective Disorders 1990;20:79–85.

Jennifer Champion is now at the Department of Psychology, Vanderbilt University. This research was supported by Grant MH 59259 from the National Institute of Mental Health to Ian H. Gotlib. The authors are grateful to the research participants and wish to extend their gratitude to Ryan Bogdan and Kathryn Dingman for their assistance in data collection and entry.