Published in final edited form as:
Acta Psychiatr Scand. 2004 May ; 109(5): 345. doi:10.1111/j.1600-0447.2003.00284.x.

Personal characteristics associated with consistency of recall of depressed or anhedonic mood in the 13-year follow-up of the Baltimore Epidemiologic Catchment Area Survey

R. Thompson[1], **H. R. Bogner**[2], **J. C. Coyne**[1,2], **J. J. Gallo**[1,2], and **W. W. Eaton**[3]

[1] Department of Psychiatry, School of Medicine, University of Pennsylvania, Philadelphia, PA

[2] Department of Family Practice and Community Medicine, School of Medicine, University of Pennsylvania, Philadelphia, PA

[3] Department of Mental Hygiene, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA

Abstract

Objective—Our study addressed two primary questions: (1) How reliable is long-term recall of lifetime history of episodes of depressed mood? (2) What characteristics are associated with consistent recall of this history?

Method—Psychiatric symptoms were assessed in a population-based longitudinal survey of 1498 persons twice, in 1981 and 1994. Respondents whose reports of history of depressed affect were discordant after a 13-year follow-up interval were compared with those whose reports were concordant.

Results—Absence of a reported history of episode of depressed mood was more consistently recalled than presence of such an episode. The kappa of reported lifetime history of episode of depressed mood was 0.32. Several personal characteristics predicted consistency of recall.

Conclusion—If assessment of past episodes of depressed mood is used as guide for identifying cases at risk for depression, account must be taken of the personal factors that might influence recall.

Keywords

depression; longitudinal studies; mental recall; inteviews

Introduction

Depression is a chronic, highly recurrent condition. A history of depression is a marker of risk for future episodes (1–3). Asking about a lifetime history of depression may be a simple method to identify persons who may be at risk for recurrence of depression and who should be evaluated and monitored with periodic assessment of mood. Indeed, many psychosocial predictors of depression lose their significance once prior episodes of depression are taken into account (4). In this study, we wished to examine how reliably a history of depressed mood can be recalled after a long follow-up interval, and what personal characteristics were associated with reliability.

Richard Thompson, Research and Training Institute, Juvenile Protective Association, 1165 North Clark, Suite 305, Chicago, IL 60610, USA, rthompson@juvenile.org.

Depression is highly prevalent as measured in surveys of the community (5,6) and in primary health care settings (7,8), and results in substantial impairment in functioning and quality of life (9–11). Despite the availability of efficacious treatment for depression, most depressed persons do not receive adequate treatment or treatment is delayed due to patient, physician, or system factors (7,12–17). Thus, identifying those at risk of, or experiencing, an episode of depression has a potential to reduce much of the impact of depression. One way to do so may be by identifying those with a history of depression.

Previous studies have examined the correspondence of brief questions regarding a lifetime history of depression with more comprehensive detailed assessments regarding the lifetime history of depression (1). The validity of these brief self-reports of lifetime history of depression appears to depend on the context in which the questions are asked (18). Specifically, asking about a lifetime history of depression in the context of a general clinical interview, where patients are sensitized to think about their medical history, typically produces more accurate responses than when such questions are embedded in more general interviews (1). As well, assessments of the validity of such brief questions presume that these questions are reliable.

Personal characteristics have been shown to influence reliability of recall of lifetime history of episodes of major depression. Persons who have experienced an episode of depression that was severe or that was associated with more symptoms are more likely to consistently recall a lifetime history of depressed mood than are persons with less severe episodes of depression and fewer symptoms (19,20). Treated episodes of depression were more likely to be recalled than episodes not marked by treatment (21). Current depression is another factor that makes recall of past episodes of depression more likely (21,22). Current stressful life events also make recall of a lifetime history of depression more likely (22). A number of demographic factors appear to influence reliability of history of depression. In particular, less consistent recall has been found in older persons (21), ethnic minorities (22), and among persons with low educational attainment (22). Gender has been inconsistently associated with recall of history of episodes of depression, with some studies finding no association (e.g. 21), and others finding more consistent recall among women (23).

Few studies have evaluated how well a history of major depression can be recalled after a significant lapse of time. Most studies focus on the relatively short-term reliability of recall of depression with typical follow-up periods of 2 years or less (e.g. 19,20). Long-term studies of recall (e.g. 24) have focused on clinical samples of psychiatric patients. Such samples may be biased toward more severe episodes, and recall of episodes of depression may be facilitated by recall of associated treatment. There may be important differences in recall of depression when patients in specialty settings are compared with patients in primary care settings for whom past episodes of depression may have been less severe and less likely to be punctuated by treatment in specialty settings. As well, the recall of an episode of depression is relatively complex, depending on the reliability of recall of each requisite symptom, and of the relative time frame in which each symptom occurred. Finally, the assessments typically used are relatively time-consuming, and may not be well-received in usual care.

Aims of the study

We hypothesized that personal factors such as age, level of educational attainment, and cognitive and psychological status might be associated with variation in ability to recall past lifetime history of episodes of depressed mood. We have also focused specifically on depressed and anhedonic mood, rather than a diagnosis of depression derived from recall of a host of symptoms.

Material and methods

The epidemiologic catchment area program

The Epidemiologic Catchment Area (ECA) program was a survey of psychiatric disorders in the general population between 1980 and 1984 at five university-based sites in the USA. The target population for the Baltimore site of the ECA was the household residents of East Baltimore, a population of 175 221 adult residents. A random sample of blocks was initially selected, based on the expected number of households. Next, a random sample of households was selected, and from each household, one individual was randomly selected to respond to the survey. In addition, to oversample the elderly at the Baltimore site, all persons aged 65 years and older in a household were interviewed. Each interview was conducted in a private place, usually the respondent's home. Details of the design of the ECA have been published elsewhere (25).

The Baltimore ECA follow-up

At all ECA sites, including Baltimore, the initial cohort was re-interviewed 1 year later. At the Baltimore site in 1981, 4238 residents aged 18 and over were probabilistically designated; of these, 3481 (82%) were interviewed in 1981 and became the target population for the 1994 follow-up. During the follow-up period, 848 respondents died, 415 could not be located, 2218 were located, and of these, 1920 (87%) participated in follow-up interviews (26,27). The participants had given permission for future follow-up at the baseline interview, and the protocol was reviewed and approved by the Committee on Human Research of the Johns Hopkins University School of Hygiene and Public Health. For this investigation, we focused on respondents who had answered questions about lifetime history of episodes of depressed mood at both 1981 baseline and 1994 follow-up interviews (n = 1739). Persons who did not have complete data at both interviews were more likely to be non-white and to have less than 12 years of education. Depression status in 1981 was not strongly associated with loss to follow-up (26). Additionally, because of the possibility that some people would have a first episode of depressed or anhedonic affect between baseline and follow-up, we excluded the 241 respondents who reported a first episode of depressed or anhedonic affect between baseline and follow-up in the 1994 interview. Our rationale for excluding persons with a reported first onset of depressed or anhedonic affect between 1981 and 1994 was that including these persons would unrealistically inflate the number of disagreements between the 1981 and 1994 reports. In other words, a person who had a first episode between 1981 and 1994 who responded consistently to the questions about mood would be classified as inconsistent, biasing estimates of the accuracy of reporting. After excluding 241 persons with report of onset of depression between 1981 and 1994, 1498 respondents were left for the analyses reported here.

Measurement strategy

Lifetime history of episodes of dysphoria and anhedonia—Reported lifetime history of episodes of dysphoria and anhedonia was assessed using items from the Diagnostic Interview Schedule (DIS) (28). The DIS is a lay-administered structured interview designed to identify respondents who met the criteria for selected psychiatric disturbances defined in the *Diagnostic and Statistical Manual for Psychiatric Disorders* (DSM-IV) (29). Detailed discussions of the reliability and validity of the DIS have been published elsewhere; it has good reliability and moderate agreement with physician diagnosis (28,30,31). In 1981, there was a single question assessing history of episodes of depressed affect and/or anhedonia. This question read, 'In your lifetime, have you ever had two weeks or more during which you felt sad, blue, depressed, or when you lost all interest or pleasure in things that you usually cared about or enjoyed?' In 1994, separate questions assessed dysphoria and anhedonia. In order to be comparable with the interview assessment in 1981, responses to the two separate questions were combined in analyses. In other words, respondents were categorized reporting lifetime history of episodes

of depressed mood in 1994 if they endorsed either anhedonia or dysphoria. The question regarding depressed affect read, 'In your lifetime, have you ever had two weeks or more when nearly every day you felt sad, blue, depressed?' Anhedonia was assessed with the question, 'Has there ever been two weeks or longer when you lost all interest in things like work or hobbies or things you usually liked to do for fun?'.

One-year psychiatric diagnoses—One-year psychiatric diagnoses were assessed using the DIS (28). In each case, as a follow-up to the identification of a lifetime incidence of each psychiatric symptom, respondents were asked how recently they had experienced these symptoms. For the purposes of our analyses, respondents were classified as meeting DSM-IV criteria for psychiatric disturbance if they reported that each of the requisite symptoms had occurred within 1 year of the follow-up (1994) interview. Participants were also classified as to whether they had an anxiety disorder other than simple phobia (generalized anxiety disorder, social phobia, agorophobia, or obsessive-compulsive disorder) within a year of the interview.

Psychological distress—Psychological distress was assessed at both baseline and follow-up using the General Health Questionnaire (GHQ) (32,33). The GHQ is a widely used tool for assessing psychological distress in clinical and community settings. The GHQ as used in the ECA study consisted of 20 items, and has been described in detail elsewhere (34). In the conventional scoring of the GHQ, a score of 4 or more indicates psychological 'caseness' (e.g. 10,34,35). In these analyses, we treated GHQ scores as continuous variables.

Medical comorbidity—The participants were asked if they had ever had the following conditions: diabetes, heart trouble, arthritis, stroke, and cancer. A positive response to any of these conditions in the 1994 interview was considered a medical comorbidity.

Functional status—Functional status was assessed at follow-up using standard survey items on both activities of daily living (ADLs) and instrumental activities of daily living (IADLs). Items assessing ADLs included getting to bed by oneself, dressing and undressing, taking a bath or shower, using the toilet, and using a knife or fork to cut up food. Items assessing IADLs included preparing meals, cleaning house, using the telephone, keeping track of money and bills, and being able to get together with friends. Consistent with previous ECA reports using ADLs and IADLs (e.g. 10,35), individuals were characterized as having ADL or IADL impairment if they reported being unable to perform at least one of the respective activities without help.

Cognitive status—Cognitive functioning at follow-up was assessed using the Mini-Mental State Examination (MMSE) (36). The MMSE is a short standardized mental status examination that has been widely employed for clinical and research purposes for global assessment of cognitive functioning. The MMSE has been extensively studied, as reviewed by Tombaugh and McIntyre (37) and by Crum and her colleagues (38). The MMSE assesses orientation to time and place, registration, memory, attention and concentration, praxis, and constructional and language capacity. These components were identified as separate factors by Jones and Gallo (39). We used a conventional cut-point of 23 of 30 (36,37,40) to dichotomize MMSE scores.

Use of mental health services reported in 1981—Participants were asked about their use of mental health services in the 6 months prior to the 1981 interview, consistent with other studies from the ECA (41,42).

Analytical strategy

The analytic plan consisted of two phases. In the first phase, the overall level of agreement between history of episodes of depressed affect assessed at baseline and at follow-up was examined. Conventional measures of consistency (kappa, sensitivity, specificity, and positive and negative predictive value) were calculated (43,44). We examined the consistency of the 1994 reports of lifetime history of episodes of depressed mood with the 1981 assessments.

In the second phase of analysis, the respondents were sorted into four groups, based on the agreement between baseline and follow-up reports of a lifetime history of episodes of depressed mood: (i) no history of episodes of depressed mood at baseline or at follow-up (interviews concordant for absence of depressed mood); (ii) no history of episodes of depressed mood at baseline, but lifetime history of episodes of depressed mood reported at follow-up interview (interviews discordant for no history of depressed mood at baseline); (iii) history of episodes of depressed mood at baseline, but no lifetime history of episodes of depressed mood reported at follow-up (interviews discordant for history of depressed mood at baseline); and (iv) history of episodes of depressed mood reported at baseline and at follow-up (interviews concordant for presence of depressed mood).

We then carried out two sets of comparisons. First, we conducted a logistic regression, with personal characteristics as independent variables, and membership in group 2 (discordant for absence of depressed mood at baseline) vs. group 1 (concordant for absence of depressed mood) as the dichotomous dependent variable. In other words, among those who reported no lifetime history of episodes of depressed mood at baseline, we examined what characteristics were associated with reporting a lifetime history of episodes of depressed affect at follow-up. Finally, we employed multivariate analyses to adjust simultaneously for all other predictors.

Second, we conducted a similar logistic regression comparing the characteristics of group 3 (discordant for presence of depressed mood reported at baseline) vs. group 4 (concordant for presence of depressed mood). In this case, the comparison highlights the characteristics associated with failing to report a lifetime history of episodes of depressed mood that had been reported 13 years earlier. Finally, we employed multivariate analyses to adjust simultaneously for all other covariates. For these analyses we have employed a level of statistical significance set at $\alpha = 0.05$, recognizing that tests of statistical significance are approximations that serve as aids to interpretation and inference.

Results

Baseline characteristics

Table 1 presents the sample characteristics. The table shows that the sample was ethnically diverse, with a mean age of 53.5 years. Only 2.3% met criteria for current MDD at follow-up, and only 1.3% had ADL impairments. At least one medical condition was reported by 40.2% of the sample.

Consistency of reports of a lifetime history of episodes of depressed mood

Table 2 presents the cross-tabulation between lifetime history of episodes of depressed mood reported at baseline and at follow-up for the entire sample. Overlap between the two interviews in terms of who reported a lifetime history of episodes of depressed mood and who did not was modest, but highly statistically significant [χ^2 (1, n = 1498) = 159.61, $P < 0.001$, $\kappa = 0.32$]. Specificity and negative predictive value were relatively high, while sensitivity and positive predictive value were substantially lower. Stated a different way, although around 90% of those not reporting a lifetime history of depressed mood at baseline still did not report a history at

follow-up, 40% of those reporting a lifetime history of depressed mood at baseline did not do so at follow-up.

Predictors of reporting history of episodes of depressed mood in 1994 when none was reported in 1981

Among those who had reported no lifetime history of episodes of depressed mood at baseline, we compared those who reported a lifetime history of episodes of depressed mood at follow-up (group 2, discordant for no lifetime history of depressed affect at baseline) with those who did not report a lifetime history of episodes of depressed mood at follow-up (group 1, interviews concordant for absence of depressed affect). This comparison is presented in Table 3. Persons with 'new' depressed affect reported at follow-up were more likely to be female, have less than 12 years of education, currently meet criteria for major depression, be anxious, be psychologically distressed, report medical comorbidities, report ADL or IADL impairments, and have had a mental health visit within 6 months of the baseline interview in 1981 or follow-up interview in 1994.

We then conducted multivariate analyses, examining predictors of reporting a lifetime history of depressed mood at follow-up. Multicollinearity analyses revealed no tolerance below 0.73 (values less than 0.4 are considered indicative of possible collinearity in logistic regressions) (45). We compared characteristics of persons who reported a lifetime history of episodes of depressed mood at follow-up (group 2) with those who did not report a lifetime history of episodes of depressed mood (group 1), among those who had reported no lifetime history of episodes of depressed mood at baseline, adjusting for all possible covariates. When terms for all possible univariate covariates were entered into the model simultaneously, gender, current depression, current anxiety, medical comorbidity, and baseline and follow-up mental health contact were independently associated with reporting a lifetime history of episodes of depressed mood (see left-most column of Table 3).

Predictors of not reporting a lifetime history of episodes of depressed mood in 1994 when reported in 1981

Among those who reported a lifetime history of episodes of depressed mood at baseline, we compared those who no longer reported a lifetime history of episodes of depressed mood at follow-up (group 3, discordant for lifetime history of depressed affect reported at baseline) with those who reported a lifetime history of episodes of depressed mood at follow-up (group 4, interviews concordant for presence of depressed affect). Results are presented in Table 4. Persons who had failed to report a previously reported lifetime history of episodes of depressed mood were more likely to be non-white, not currently meeting criteria for major depression, not psychologically distressed, and cognitively impaired as measured by the MMSE. When terms for all univariate covariates were entered into the model simultaneously, multicollinearity analyses revealed no tolerance below 0.68. In this model, ethnicity, currently meeting criteria for major depression, and current distress were independently associated with reporting a lifetime history of episodes of depressed mood.

Discussion

Given the potential importance of assessing patients for a lifetime history of depression, we employed prospectively-gathered community responses spanning 13 years of follow-up to evaluate how well a lifetime history of episodes of depressed mood was recalled. Our study differs from prior work because we have investigated responses from a large, ethnically diverse community sample with standardized psychiatric assessments 13 years apart.

In this sample, respondents were moderately consistent in recalling a lifetime history of episodes of depressed mood after 13 years of follow-up; namely, the κ for agreement between baseline and 13-year follow-up reports was 0.32. Estimates of reliability of a lifetime diagnosis of depression tend to be somewhat higher (κ = 0.41: 19; κ = 0.48: 21), but these are based on a much briefer follow-up period (18–19 months). Those reporting no lifetime history of episodes of depressed mood were likely to have been consistent, about half of the persons with a lifetime history of episodes of depressed mood as reported in 1981 denied a lifetime history of episodes of depressed mood in 1994. Women, those meeting diagnostic criteria for major depression or an anxiety disorder, those experiencing psychological distress, medical comorbidities, functional impairment, and having reported a mental health visit within 6 months of the 1981 or 1994 interview were associated with recalling a lifetime history of episodes of depressed mood in 1994 when none was reported in 1981. Persons who did not recall a lifetime history of episodes of depressed mood in 1994 despite having reported a history of episodes of depressed mood in 1981 tended to self-identify as an ethnic minority and to be cognitively impaired. In addition, not meeting a current diagnosis of major depression and not being distressed as measured by a standardized questionnaire were also associated with failure to report depressed mood that had been reported earlier.

Before further discussing our findings and placing them within the context of other research, potential limitations of the current study should be acknowledged. First, there is the possibility of sources of error associated with self-report interview data. In particular, imperfect recall and response bias (especially socially desirable responding) may have influenced the results. Depression itself may influence reporting of mood states and functioning. Second, although the initial study was based on a community sample, the follow-up data consisted of persons who could be found and interviewed after a considerable follow-up period. Studies of the ECA follow-up have shown that baseline depression and other psychiatric disturbances (26,46) were not important characteristics associated with follow-up. Nevertheless, other variables may have been related to participation at follow-up and could influence our results in unknown ways. In our analysis, we excluded persons who reported a first episode of major depression between baseline and follow-up so that the discordance between baseline and follow-up would not be attributed to new onset of depression. However, it is still possible that participants were inaccurate in reporting when their first episode of depression had occurred. Third, our study represents an extreme test of recall for depressed mood because of the length of time between the baseline and follow-up interviews. The follow-up period was 13 years during which important secular trends occurred which may have affected how people define depression. A related issue is that we excluded persons based on reports of the timing of previous episodes of depressed or anhedonic mood. If the timing of onset of symptoms was remembered inaccurately, our findings may have been altered in unknown ways. As well, although this study used a large sample, some characteristics (such as cognitive impairment) were relatively rare, widening confidence intervals and reducing the likelihood of finding an association. A related point is that the two sets of comparisons (group 1 vs. group 2 and group 3 vs. group 4) also did not have equal power. Thus, interpretations of the failure to find statistically significant results should be made with caution. Finally, readers should keep in mind that this study focused on specific key symptoms of depression-sadness and anhedonia. We did not examine the consistency of depression diagnosis because, in the ECA, a history of depression is based on an algorithm applied to a series of symptom-level items, rather than a single question that asks about prior episodes of depression.

Another possible limitation of the study is the fact that the data collection was completed in 1994. Certainly, there have been a number of changes in popular attitudes about depression in the past decade, with the disorder becoming more accepted and more widely understood. This may change how people reconstruct their personal histories of depression, making them more sensitive to its occurrence, and more willing to seek treatment (47).

Despite this and the other limitations, our study has the advantage of using a large prospective dataset to assess personal characteristics associated with accuracy in recall of a lifetime history of episodes of depressed mood. In contrast, most other studies have been limited by small sample size (e.g. 24), unrepresentative samples (e.g. 24), brief follow-up periods (e.g. 20), or focus on the reliability of syndromes rather symptoms (e.g. 22). As well, it is important to note that the ECA dataset continues to be a rich source for recent investigation of emerging questions (e.g. 26,35).

Our results highlight some demographic and clinical characteristics that make history of depressed mood more difficult to ascertain reliably. First, the absence of current depression or distress makes it less likely that a previously reported episode of depressed mood will be reported. The association of absence of distress with decreased likelihood of reporting a past episode previously reported may explain why samples of patients from specialized mental health settings have more consistent recall for history of depression (e.g. 24). In any case, it appears that depressed patients may be cognitively predisposed to recall previous instances of depression, whether the mechanism involved is state-dependent learning (48) or mood-congruent recall (49). Patients in mental health settings (compared with primary care or general population samples) have more depressive episodes to report, and episodes that are more severe and likely to be marked by treatment (7,50). Patients in mental health settings are cued and sensitized by the clinical context of assessment. Depressed persons' reports of having been depressed may be influenced by current psychological state. The possibility that present state influences reporting was supported by our finding that persons currently experiencing distress or depression, and those with a history of mental health treatment, were also more likely to recall a history of depression, when none was provided in the earlier interview.

Other characteristics also influenced consistency of recall. Those who self-identify as ethnic minorities were less likely to report a lifetime history of episodes of depressed mood that had previously been reported. Younger adults and women were more likely to recall a lifetime history of episodes of depressed mood when none had been provided at the earlier interview. As we noted earlier, a number of secular trends occurred during the period between baseline assessment and follow-up that may have influenced our results. Specifically, during the 1980s and early 1990s, depression underwent considerable destigmatization, and treatment for depression became more accessible and more widespread (23). Public awareness of depression increased since 1981. This may have resulted in individuals reinterpreting earlier experiences based on increased awareness of depression. Thus, a discordance many not necessarily reflect an inaccuracy in either the follow-up or baseline report. This possibility was reduced in our sample by our reliance on a simple recall of 2 weeks of uninterrupted negative mood, rather than asking simply whether the respondent had ever been depressed. It is also difficult to understand how such secular trends would influence people to attribute an experience that they had earlier labeled a depressive episode as a non-episode.

If we were to use a single question about depressed mood to identify persons who require increased surveillance, our findings suggest that we would miss half of the persons who might earlier have reported history of episodes of depressed mood. The episodes of depressed mood that were reported in 1981 but not recalled in 1994 may not have been clinically significant. In other words, whether the 50% sensitivity is adequate for targeting a group of people for more intense follow-up depends on whether the persons who fail to recall their previously reported depressed affect are at any more increased risk of future episodes of depression than others.

Excluding persons who had a first episode of depression in the interim, the reliability of recall of a 2 week period of depressed mood over 13 years was moderate, and slightly lower than that for syndomes assessed over shorter follow-up periods. The results are encouraging for the

use of such brief questions for clinical and research purposes, particularly as the first stage of a two-stage assessment process. Clinical and demographic correlates of instability in reporting a history of depression need to taken into account in how the history of depression is used to target persons for further follow-up. Our results may serve to stimulate further interest in a promising strategy to identify increased risk of depression; namely, the prior history. Future research will have to assess how reliability can be improved with prompts and follow-up probes.

Acknowledgments

Data gathering in 1981 was supported by the Epidemiologic Catchment Area (ECA) Program of the National Institute of Mental Health Division of Biometry and Epidemiology (MH33870). Data gathering in the Baltimore ECA Follow-up (1993–1996) was supported by the National Institute of Mental Health (MH47447). Dr Bogner was supported by a Health Resources and Services Administration Faculty Development Program in Primary Care training grant and an Advanced Research Training Award from the American Academy of Family Physicians.

References

1. Coyne JC, Thompson R, Racioppo MW. Validity and efficiency of screening for history of depression by self-report. Psychol Assess 2001;13:163–170. [PubMed: 11433792]
2. Frank E, Thase ME, Frank E, Thase ME. Natural history and preventative treatment of recurrent mood disorders. Annu Rev Med 1999;50:453–468. [PubMed: 10073289]
3. Judd LL, Akiskal HS, Maser JD, et al. A prospective 12-year study of subsyndromal and syndromal dperessive symptoms in unipolar major depressive disorders. Arch Gen Psychiatry 1998;55:694–700. [PubMed: 9707379]
4. Kessler RC, Magee WJ. Childhood family violence and adult recurrent depression. J Health Soc Behav 1994;35:13–27. [PubMed: 8014427]
5. Kessler RC, Mcgonagle KA, Zhao S. Lifetime and 12-month prevalence of DSM-II-R psychiatric disorders in the United States. Arch Gen Psychiatry 1994;51:8–19. [PubMed: 8279933]
6. Regier DA, Boyd JH, Burke JDJ, et al. One-month prevalence of mental disorders in the United States: based on five Epidemiologic Catchment Area sites. Arch Gen Psych 1988;45:977–986.
7. Coyne JC, Fechner-Bates S, Schwenk TL. The prevalence, nature and co-morbidity of depressive disorders in primary care. Gen Hosp Psychiatry 1994;16:267–276. [PubMed: 7926703]
8. Katon W. Epidemiology of depression in medical care. Int J Psychiatry Med 1987;17:93–112. [PubMed: 3294691]
9. Broadhead WE, Blazer DG, George LK, Tse CK. Depression, disability days, and days lost from work in a prospective epidemiologic survey. JAMA 1990;264:2524–2528. [PubMed: 2146410]
10. Gallo JJ, Rabins PV, Lyketsos CG, Tien AY, Anthony JC. Depression without sadness: Functional outcomes of nondysphoric depression in later life. J Amer Geriatr Soc 1997;45:570–578. [PubMed: 9158577]
11. Beekman AT, Penninx BW, Deeg DJ, de Beurs E, Geer Omg SW, van Tilburg W. The impact of depression on the well-being, disability and use of services in older adults: a longitudinal perspective. Acta Psychiatr Scand 2002;105:20–27. [PubMed: 12086221]
12. Depression Guidelines Panel. Detection and diagnosis: clinical practice guideline. AHCPR Publication No. 93–0550. Vol. 1. US Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research; Rockville, MD: 1993. Depression in primary care.
13. Depression Guidelines Panel. Treatment of major depression: clinical practice guideline. AHCPR Publication No. 93–0551. Vol. 2. US Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research; Rockville, MD: 1993. Depression in primary care.
14. Gerber PD, Barrett J, Manheimer E, Whiting R, Smith R. Recognition of depression by internists in primary care: a comparison of internist and ‘gold standard’ psychiatric assessments. J Gen Intern Med 1989;4:7–13. [PubMed: 2915277]

15. Monroe SM, Simons AD, Thase ME, Monroe SM, Simons AD, Thase ME. Onset of depression and time to treatment entry: roles of life stress. J Consult Clin Psychol 1991;59:566–573. [PubMed: 1918561]

16. Luutonen S, Holm H, Salminen JK, Risla A, Salokangas RK. Inadequate treatment of depression after myocardial infarction. Acta Psychiatr Scand 2002;106:434–439. [PubMed: 12392486]

17. Williams JW Jr, Mulrow CD, Chiquette E, Noel PH, Aguilar C, Cornell J. A systematic review of newer pharmacotherapies for depression in adults: evidence report summary. Ann Intern Med 2000;132:743–756. [PubMed: 10787370]

18. Howe A, Bath P, Goudie F, et al. Getting the questions right: an example of loss of validity during transfer of a brief screening approach for depression in the elderly. Int J Geriatr 2000;15:650–655.

19. Bromet EJ, Dunn LO, Connell MM, Dew MA, Schulberg HC. Long-term reliability of diagnosing major depression in a community sample. Arch Gen Psych 1988;43:435–440.

20. Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ. The lifetime history of major depression in women: reliability of diagnosis and heritability. Arch Gen Psych 1993;50:863–870.

21. Kendler KS, Gardner CO, Prescott CA. Are there sex differences in the reliability of a lifetime history of major depression and its predictors? Psych Med 2001;31:617–625.

22. Aneshensel CS, Estrada AL, Hansell MJ, Clark VA. Social psychological aspects of reporting behavior: lifetime depressive episode reports. J Health Soc Behav 1987;28:232–246. [PubMed: 3680917]

23. Warshaw MG, Klerman GL, Lavori PW. Are secular trends in major depression an artifact of recall? J Psychiatr Res 1991;25:141–151. [PubMed: 1941710]

24. Andrews G, Anstey K, Brodaty H, Issakidis C, Luscombe G. Recall of depressive episode 25 years previously. Psych Med 1999;29:787–791.

25. Robins, LN.; Regier, DA. Psychiatric disorders in America: the Epidemiologic Catchment Area Study. New York: Free Press; 1991.

26. Badawi MA, Eaton WW, Myllyluoma J, Weimer LG, Gallo J. Psychopathology and attrition in the Baltimore ECA 15-year follow-up 1981–1996. Soc Psychiatry Psychiatr Epidemiol 1999;34:91–98. [PubMed: 10189815]

27. Eaton WW, Anthony JC, Gallo JJ, et al. Natural history of Diagnostic Interview Schedule/DSM-IV major depression: the Baltimore Epidemiologic Catchment Area follow-up. Arch Gen Psych 1997;54:993–999.

28. Eaton, WW.; Kessler, LG. Epidemiologic field methods in psychiatry: the NIMH epidemiologic catchment area program. New York: Academic Press; 1985.

29. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, DSM-IV. 4. Washington D.C: American Psychiatric Association; 1994.

30. Anthony JC, Folstein M, Romanoski AJ, et al. Comparison of the lay Diagnostic Interview Schedule and a standardized psychiatric diagnosis: experience in eastern Baltimore. Arch Gen Psych 1985;42:667–675.

31. Helzer JE, Robins LN, McEvoy LT, et al. A comparison of clinical and diagnostic interview schedule diagnoses: physician examination of lay-interviewed cases in the general population. Arch Gen Psych 1985;42:657–666.

32. Goldberg, DP. The detection of psychiatric illness by questionnaire. London: Oxford Univeristy Press; 1972.

33. Goldberg DP, Rickels K, Downing R, Hesbacher P. A comparison of two psychiatric screening tests. Br J Psychiatry 1976;129:61–67. [PubMed: 938806]

34. Samuels JF, Nestadt G, Anthony JC, Romanoski AJ. The detection of mental disorders in the community setting using a 20-item interview version of the General Health Questionnaire. Acta Psychiatr Scand 1994;89:14–20. [PubMed: 8140902]

35. Bogner HR, Gallo JJ, Sammel MD, Ford DE, Armenian HK, Eaton WW. Urinary incontinence and psychological distress among community-dwelling older adults. J Am Geriatr Soc 2002;50:489–495. [PubMed: 11943045]

36. Folstein MF, Folstein S, McHugh PR. 'Mini-Mental State': a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–198. [PubMed: 1202204]

37. Tombaugh TN, Mcintyre NJ. The Mini-Mental State Examination: a comprehensive review. J Am Geriatr Soc 1992;40:922–935. [PubMed: 1512391]
38. Crum RM, Anthony JC, Bassett SS, Folstein MF. Population-based norms for the Mini-Mental State Examination by age and educational level. JAMA 1993;269:2386–2391. [PubMed: 8479064]
39. Jones RN, Gallo JJ. Factor analysis of the Mini-Mental State Examination in a community-dwelling sample of older adults. Psych Med 2000;30:603–613.
40. Anthony JC, Leresche L, Niaz U, von Korff M, Folstein MF. Limits of the 'Mini-Mental State' as a screening test for dementia and delirium among hospital patients. Psych Med 1982;12:397–408.
41. Gallo JJ, Marino S, Ford D, Anthony JC. Filters on the pathway to mental health care: II. Sociodemographic factors. Psych Med 1995;25:1149–1160.
42. Marino S, Gallo JJ, Ford D, Anthony JC. Filters on the pathway to mental health care: I. Incident mental disorders. Psych Med 1995;25:1135–1148.
43. Fleiss, J. Statistical methods for rates and proportions. 2. New York: John Wiley & Sons; 1981.
44. Kraemer, HC. Evaluating medical tests: objective and quantitative guidelines. Newbury Park, CA: Sage Publications; 1992.
45. Allison, PD. Logistic regression using the SAS system. SAS Institute; Cary, NC: 1999.
46. Eaton WW, Anthony JC, Tepper S, Dryman A. Psychopathology and attrition in the epidemiologic catchment area surveys. Am J Epidemiol 1992;135:1051–1059. [PubMed: 1595691]
47. Olfson M, Marcus SC, Druss B, Elinson L, Tanielian T, Pincus HA. National trends in the outpatient treatment of depression. JAMA 2002;287:203–209. [PubMed: 11779262]
48. Bower GH. Mood and memory. Am Psychol 1981;36:129–148. [PubMed: 7224324]
49. Dalgleish T, Watts FN. Biases of attention and memory in disorders of anxiety and depression. Clin Psychol Rev 1990;10:589–604.
50. Coyne JC, Schwenk TL, Fechner-Bates S. Nondetection of depression by primary care physicians reconsidered. Gen Hosp Psychiatry 1995;17:3–12. [PubMed: 7737492]

Table 1

Description of the sample. Data gathered from the Baltimore, Maryland Epidemiologic Catchment Area Program follow-up, 1981–1994

	Mean (SD) or proportion
Sociodemographic characteristics at follow-up	
Age (years)	53.5 (16.1)
Gender (% female)	62.5
Marital status (% married)	49.8
Ethnicity (% white)	60.7
Educational attainment (years)	11.3 (2.9)
Psychological status at follow-up	
Major depression (% within 1 year of interview)	2.3
Anxiety disorder excluding phobias (% within 1 year of interview)	18.7
GHQ score	1.8 (3.1)
Cognitive status at follow-up	
Mini-Mental State Examination score below cutpoint (%)	5.5
Medical and functional status at follow-up	
Any medical comorbidity (%)	40.2
ADL impairment (%)	1.3
IADL impairment (%)	4.7
Use of health services at baseline	
Mental health visit within 6 months of 1981 interview (%)	3.9

GHQ, General Health Questionnaire; ADL, activities of daily living; IADL, instrumental activities of daily living.

Table 2

Relationship between history of episodes of depressed mood reported in 1981 and in 1994. Data gathered from the Baltimore, Maryland Epidemiologic Catchment Area Program follow-up, 1981–1994

	No history of depressed affect in 1981	History of depressed affect reported in 1981	Total	
No history of depressed affect in 1994	970	259	1229	NPV = 78.9% 95% CI, (77.5, 80.3)
History of depressed affect reported in 1994	109	160	269	PPV = 59.5% 95% CI, (54.1, 64.7)
Total	1079	419	1498	Kappa = 0.32 95% CI, (0.26, 0.37)
	Specificity = 89.9% 95% CI, (88.6, 91.2)	Sensitivity = 38.2% 95% CI, (34.7, 41.5)		

CI, confidence interval; NPV, ‘negative predictive value’; PPV, ‘positive predictive value’.

Table 3

Factors associated with providing a lifetime history of episodes of depressed mood in 1994 among persons who denied a lifetime history of depressed mood in 1981. Odds ratios represent estimates of likelihood of recalling lifetime history of depressed mood at follow-up relative to likelihood of not recalling history. Data gathered from the Baltimore, Maryland Epidemiologic Catchment Area Program follow-up, 1981–1994

	Univariate OR	Adjusted OR
Sociodemographic characteristics in 1994		
Age (above 60 years)	0.77 (0.50–1.20)	0.69 (0.39–1.20)
Gender (female)	2.31 (1.49–3.66)**	2.00 (1.20–3.30)*
Marital status (married)	0.68 (0.46–1.01)	0.96 (0.60–1.54)
Ethnicity (white)	0.93 (0.62–1.38)	1.23 (0.76–2.01)
Educational attainment (at least 12 years)	0.64 (0.43–0.96)*	0.71 (0.44–1.14)
Psychological status in 1994		
Major depression (within 1 year of interview in 1994)	125.00 (15.85–1000)**	80.25 (9.46–681.07)**
Anxiety disorder excluding simple phobia (within 1 year of interview in 1994)	3.02 (1.95–4.70)**	1.80 (1.07–3.05)*
GHQ score (4 and over)	2.58 (1.65–4.70)**	1.22 (0.66–2.20)*
Cognitive status in 1994		
Mini-Mental State Examination score (less than 23)	0.45 (0.14–1.47)	0.43 (0.12–1.60)
Medical and functional status in 1994		
Any medical comorbidity	1.64 (1.10–2.45)*	1.76 (1.07–2.88)*
ADL impairment	5.24 (1.51–18.18)*	2.34 (0.44–12.36)
IADL impairment	2.91 (1.44–5.90)**	2.08 (0.79–5.46)
Use of health services		
Mental health visit within 6 months of interview in 1981	8.25 (2.93–23.22)**	8.91 (2.95–26.93)**
Mental health visit within 6 months of interview in 1994	18.11 (1.63–201.41)**	47.00 (3.48–635.30)*

OR, odds ratio; CI, confidence interval; GHQ, General Health Questionnaire; ADL, activities of daily living; IADL, instrumental activities of daily living.

* $P < 0.05$;

** $P < 0.005$.

Table 4

Factors associated with failing to report a history of depressed mood in 1994 among persons who reported a lifetime history of depressed mood in 1981. Odds ratios represent estimates of likelihood of not recalling lifetime history of depressed mood at follow-up relative to likelihood of recalling history. Data gathered from the Baltimore, Maryland Epidemiologic Catchment Area Program follow-up, 1981–1994

	Univariate OR	CI
Sociodemographic characteristics in 1994		
Age (above 60 years)	1.41 (0.90–2.18)	1.31 (0.74–2.31)
Gender (female)	1.00 (0.66–1.54)	1.06 (0.65–1.73)
Marital status (married)	1.13 (0.76–1.69)	1.23 (0.77–1.97)
Ethnicity (white)	0.52 (0.34–0.80)*	0.42 (0.25–0.71)**
Educational attainment (at least 12 years)	0.70 (0.49–1.01)	0.82 (0.49–1.36)
Psychological status in 1994		
Major depression (within 1 year of interview in 1994)	0.03 (0.00–0.19)**	0.03 (0.00–0.26)**
Anxiety disorder excluding simple phobia (within 1 year of interview in 1994)	0.86 (0.55–1.34)	1.28 (0.74–2.23)
GHQ score (4 and over)	0.43 (0.28–0.67)**	0.58 (0.34–0.97)**
Cognitive status in 1994		
Mini-Mental State Examination score (less than 23)	3.11 (0.89–10.93)	3.61 (0.60–21.63)
Medical and functional status in 1994		
Any medical comorbidity	0.92 (0.62–1.36)	0.88 (0.54–1.44)
ADL impairment	1.24 (0.31–5.03)	0.59 (0.10–3.45)
IADL impairment	0.79 (0.34–1.85)	0.75 (0.33–1.43)
Use of health services		
Mental health visit within 6 months of interview in 1981	0.61 (0.33–1.16)	0.69 (0.33–1.43)
Mental health visit within 6 months of interview in 1994	n/a	n/a

OR, odds ratio; CI, confidence interval; GHQ, General Health Questionnaire; ADL, activities of daily living; IADL, instrumental activities of daily living.

* $P < 0.05$;

** $P < 0.005$;

n/a, one group had no persons with a recent mental health visit in 1994 so the odds ratios could not be calculated.