

Does the DSM-IV Clinical Significance Criterion for Major Depression Reduce False Positives? Evidence From the National Comorbidity Survey Replication

Jerome C. Wakefield, Ph.D.,
D.S.W.

Mark F. Schmitz, Ph.D.

Judith C. Baer, Ph.D.

Objective: To reduce false positive diagnoses, DSM-IV added a clinical significance criterion to many diagnostic criteria sets requiring that symptoms cause significant distress or impairment. The DSM-V Task Force is considering whether clinical significance should remain a diagnostic threshold or become a separate dimension, as it is in ICD. Yet, the criterion's effectiveness in validly reducing the prevalence of specific disorders remains unclear. Critics have argued that for some categories, notably major depression, the criterion is redundant with symptoms, which are inherently distressing or impairing. The authors empirically evaluated the criterion's effect on the prevalence of major depression in the community. This report also considers more broadly the relationship of symptoms to impairment in diagnosis.

Method: Subjects were respondents, aged 18 to 54 years, who participated in the National Comorbidity Survey Replication (N=6,707). The effect of the clinical significance criterion's distress and impairment components on major depression was assessed in this sample. Distress questions were administered to all respondents reporting persistent sadness

(≥2 weeks) or the equivalent. Questions pertaining to role impairment were asked of all respondents satisfying major depression symptom-duration criteria.

Results: Of 2,071 individuals reporting persistent sadness or the equivalent, 97.2% (N=2,016) satisfied criteria for distress. Of 1,542 individuals satisfying depression symptom-duration criteria, 96.2% (N=1,487) satisfied criteria for impairment.

Conclusion: These findings support the redundancy thesis. Distress is virtually redundant with symptoms of persistent sadness, even in the absence of major depression, and impairment is almost always entailed by major depression-level symptoms. Thus, the clinical significance criterion does not substantially reduce the prevalence of major depression in the community. The DSM-V Task Force should consider eliminating the criterion and explore alternative ways to identify false positives in the diagnosis of depression. The criterion's status for other disorders should be evaluated on a disorder-by-disorder basis because the diagnostic relationship between symptoms and impairment varies across categories.

(*Am J Psychiatry* 2010; 167:298–304)

To reduce false positive diagnoses of conditions satisfying diagnostic criteria but insufficiently harmful to be classified as disorders, DSM-IV added the following clinical significance criterion to most diagnostic criteria sets, including those for major depression: "The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning" (1, p. 356). The goal of introducing this criterion was to help "establish the threshold for the diagnosis of a disorder in those situations in which the symptomatic presentation by itself (particularly in its milder forms) is not inherently pathological and may be encountered in individuals for whom a diagnosis of mental disorder would be inappropriate" (1, p. 8). The criterion was added strictly on conceptual grounds. DSM-IV field trials did not test its effectiveness in eliminating false positives (2).

The DSM-IV clinical significance criterion is broad and does not require "extreme" distress or "a lot" of impairment, only "some" (i.e., enough to be clinically significant). As in physical medicine, mental disorders can be mild or moderate. Thus requiring higher levels of distress or impairment in diagnoses increases the risk for substantial false negatives (3, 4). "Some" impairment resulting from a clear mental dysfunction certainly is sufficient to warrant diagnosis of a disorder, as reflected in DSM-IV's depression severity index, which allows for "some" impairment in the classification of moderate disorders.

The clinical significance criterion takes time to administer, and thus incurs clinician, patient, and epidemiologic interview costs. Yet, the criterion's usefulness and validity as a supplement to symptom criteria have been questioned on several grounds. It has been argued that symptom cri-

teria already ensure role impairment or distress, making the criterion redundant, and that symptoms alone are sufficient to indicate a disordered condition. Additionally, it has been objected that substantial distress or impairment may occur in normal reactions to loss or stress, and therefore the criterion's elimination of mild conditions does not address a central "false positives" challenge. Thus, the criterion's clinical and epidemiologic utility deserves scrutiny.

The criterion's disjunctive logical structure shapes its evaluation. If there is either significant distress or significant role impairment, then the overall criterion is satisfied. Consequently, the criterion is only as powerful as its weakest disjunct. If one disjunct is readily satisfied by nondisorders, then the overall criterion will not eliminate false positives.

The DSM-V Task Force is reportedly re-examining the clinical significance criterion, considering whether to remove it from the diagnostic criteria and make it a separate dimension to address problems that arise from linking symptoms and impairment in diagnosis (5). For example, when diagnosing comorbid conditions, clinicians may have difficulty differentiating whether each relevant symptom set causes distress or impairment.

The criterion is also a major point of divergence between DSM and ICD. Because role norms vary considerably across cultures, ICD avoids reference to role impairment in diagnostic criteria. The current ICD revision will further pursue such separation, with a recent proposal that "no functioning or disability should appear as part of the threshold of the diagnosis" of any disorder (6). Eliminating the DSM-IV clinical significance criterion would promote coordination between the two diagnostic manuals.

Added generically across DSM-IV categories, the criterion received only minimal disorder-by-disorder analysis. However, its utility and validity might vary according to each disorder's logic (3, 7). Our examination of the criterion focuses on one important category: major depression.

False positives are of particular concern in community prevalence studies (8, 9). Epidemiologists had hoped that the clinical significance criterion might correct what seemed to be inflated community prevalence rates reported in major studies for some disorders, including major depression (10, 11), but whether this hope was realized has not been adequately evaluated. In the present study, we empirically assessed whether the clinical significance criterion validly reduces the prevalence of major depression in the community. Reflecting the criterion's disjunctive nature and constraints of the data, each disjunct was examined separately.

Clinical Significance and Major Depression: The Redundancy Thesis

A decade ago, Spitzer and Wakefield (3 [also see references 12, 13]), using major depression as a primary example, argued, first, that distress is common to both normal reac-

tions (e.g., acute grief) and disordered conditions: "Since most of these symptoms are either intrinsically distressing (e.g., depressed mood, psychomotor agitation, fatigue) or are almost invariably accompanied by distress about having the symptom (e.g., diminished interest or pleasure, weight loss or gain, hypersomnia, psychomotor retardation, thoughts of death), it is highly unlikely that one could satisfy the criteria and not be significantly distressed" (3, p. 1862). Second, they argued that many symptoms (for depression, e.g., distraction, fatigue, psychomotor retardation) are "inherently associated with significant impairment, so the clinical significance criterion is redundant" (3, p. 1856). They predicted that "if distress and impairment are interpreted broadly, the clinical significance criterion is pragmatically redundant," whereas "if the clinical significance criterion is interpreted more narrowly, false negatives become a problem" (3, p. 1862). They suggested focusing on the validity of symptom indicators rather than clinical significance to screen out false positives.

Zimmerman et al.'s study (14) of 1,500 outpatients strongly confirmed the redundancy prediction in clinical populations. Trained diagnosticians judged whether case subjects met the DSM-IV clinical significance requirement. Results demonstrated that "no patient who met the symptom criteria for current major depressive disorder... failed to meet the clinical significance criterion" (14, p. 1400). The same 0% elimination rate was found for lifetime depression.

Community Prevalence and Clinical Significance

However, the true test of the redundancy thesis lies in community studies, where false positives are most likely to occur. Moreover, the current movement toward mass screening using DSM criteria makes the validity of community diagnosis more clinically relevant.

Community studies requiring clinical significance have sometimes yielded reductions in the rates of prevalence, but not always. Studying the prevalence of major depression in the National Comorbidity Survey (NCS) (11), Mojtabai (15) ignored distress and used "interference with life or activities" to approximate the criterion's impairment component. He observed that "very few individuals reported no impairment (weighted number=29)" (15, p. 207), suggesting that a full DSM-style clinical significance criterion would be redundant.

In contrast, Narrow et al. (9) imposed a clinical significance criterion on NCS data, resulting in a reduction in 1-year depression rates from 8.9% to 5.4%. However, their criterion required either outpatient service contact or a lot of interference with daily life or activities, which was more restrictive than the DSM-IV criterion. Moreover, service contact is conceptually unrelated to either the DSM-IV criterion or disorder status (16, 17). Slade et al. (18), using an Australian national probability sample,

found that requiring clinical significance reduced 1-year prevalence rates by 19%, but their criterion ignored distress and required that symptoms “seriously interfere” with—rather than just significantly impair—role functioning. Thus, reduced prevalence in the Narrow et al. and Slade et al. studies resulted from the use of narrower criteria than that of DSM-IV.

In their study of a community sample of Native Americans, Beals et al. (4) explored the effect of different versions of the clinical significance criterion on prevalence. The version closest to that of DSM-IV, which required some or a lot of either distress (“How much did X ever bother or upset you?”) or impairment (“How much did X ever interfere with your life or activities?”), had little effect, only reducing the prevalence of lifetime major depressive episodes from 10.5% to 10.2%, whereas narrower criteria reduced prevalence more saliently. Beals et al. reported the following conclusions: “Spitzer and Wakefield anticipated... little influence on the false positives. These findings support their hypothesis....The CS [clinical significance] criterion...demonstrates little effectiveness in increasing the validity of diagnoses” (4, p. 1197). However, the sample in the Beals et al. study was not nationally representative, and they examined the broader category of major depressive episode, not major depressive disorder.

In summary, substantial reductions in the rates of prevalence in the community have occurred using clinical significance criteria narrower than that of DSM-IV. However, there has been no clear test of the redundancy of the DSM-IV criterion using a nationally representative community sample.

Our redundancy hypothesis was that neither distress nor impairment criteria would have substantial effect on prevalence.

Method

Sample

The National Comorbidity Survey Replication (NCS-R) (19) is a community-based epidemiological survey of a nationally representative U.S. sample. The survey was administered between February 2001 and December 2002 to 9,282 persons aged 18 years or older. In the present study, we restricted our analyses to the sample respondents ages 18 to 54 years ($N=6,707$). Demographic and other survey data are available elsewhere (19).

DSM-IV assesses clinical significance in those individuals whose conditions satisfy the symptom-duration criterion A. The NCS-R followed this approach for questions regarding impairment, but questions pertaining to distress were asked earlier in the interview of any respondent who reported adequately persistent sadness or the equivalent. This complicated our analysis because, given different subsamples resulting from skipped interview questions, the distress and impairment components had to be analyzed separately. However, examining distress in the broader persistent-sadness group was advantageous because it posed a more difficult challenge to the redundancy prediction. This group would be expected to include many individuals reporting milder, less symptomatic experiences that might not be distressful.

Diagnostic Assessments

The NCS-R major depression diagnosis required 2 weeks of five or more DSM-IV-based symptoms, which had to include symptoms of sadness/emptiness/depression or of being discouraged or uninterested (not as the result of organic causes), and no lifetime mania or hypomania. In multiple-episode cases, symptoms were assessed for a “target” episode identified as the most severe. Mixed-episode and psychotic disorder exclusions were not operationalized. The survey’s clinical significance requirement for major depression used multiple measures for distress and impairment, reflecting the broad DSM-IV criterion.

The NCS-R did not operationalize DSM-IV’s bereavement exclusion. Thus, it allowed false positive diagnoses over and above the usual ones. These diagnoses were excluded in the NCS (20) at known rates, offering a comparison point for clinical significance in the replicated version of the survey.

The major depression portion of the interview was administered only if the respondent answered “yes” to one or more screener questions regarding persistent sadness (e.g., “Have you ever in your life had a period lasting several days or longer when most of the day you felt sad, empty or depressed?”), with additional screeners confirming that the symptom occurred at least 1 hour nearly every day for at least 2 weeks or multiple 3-day periods over 1 year. The distress questions followed this screening for persistent sadness. Those who reported distress were further evaluated on whether at least five relevant symptoms occurred during the same 2-week period (satisfying DSM-IV criterion A). Those reporting no distress were eliminated from the interview. Questions regarding impairment and various exclusions followed questions pertaining to symptoms.

Clinical Significance Criterion

A positive response to any one of 10 distress and impairment questions satisfied the survey’s major depression clinical significance criterion.

Distress. NCS-R distress questions were asked of everyone who experienced 2 weeks of sadness or the equivalent. Distress was operationalized using the following four questions: 1) “How severe was your emotional distress” (mild, moderate, severe, or very severe)?; 2) “How often was your emotional distress so severe that nothing could cheer you up” (often, sometimes, rarely, or never)?; 3) “How often was your emotional distress so severe that you could not carry out your daily activities” (often, sometimes, rarely, or never)?; 4) “Did you feel so sad that nothing could cheer you up nearly every day” (yes or no)? With the exception of the daily activities question, which had a positive response threshold of “rarely,” responses of “mild” and “rarely” pertaining to distress were not considered clinically significant, and responses of “sometimes” and “moderate” were thresholds.

Impairment. The following two main questions regarding impairment, which referred to the effects of the entire syndrome, were asked of all respondents who satisfied the major depression five-symptom and 2-week duration A criteria: 1) “How much did these problems interfere with either your work, your social life, or your personal relationships during the episode” (not at all, a little, some, a lot, or extremely)? 2) “How often during that episode were you unable to carry out your daily activities” (often, sometimes, rarely, or never)? Respondents answering the first impairment question with “not at all” skipped the second question and did not qualify for impairment. Respondents who answered with “some,” “a lot,” or “extremely” qualified for impairment. Thus, the second question potentially affected impairment results only for respondents who answered the first question nonqualifyingly as “a little.” Responses of “rarely” or above were considered positive for impairment on the second question.

TABLE 1. Respondents Satisfying the Distress and Impairment Components of the National Comorbidity Survey Replication (NCS-R) Clinical Significance Criterion for Major Depression^a

Variable	N	% ^b
DSM-IV clinical significance criterion component		
Distress		
Reported episodes of sadness	2,071	30.3
Reported episodes of sadness with significant distress (number who reported episodes of sadness who also reported significant distress)	2,016	97.2
Reported episodes of non-major depressive disorder sadness with significant distress (number who reported episodes of sadness that did not satisfy criteria for major depressive disorder [N=817] and also reported significant distress)	765	93.5
Impairment		
Satisfied criterion A for diagnosis of major depressive disorder (five symptoms; 2-week duration)	1,542	22.4
Reported significant impairment (of those satisfying requirements for criterion A)	1,487	96.2
Major depressive disorder adult prevalence rate		
NCS-R (with clinical significance criterion)	1,254	18.3
NCS (without clinical significance criterion)	1,235	15.2

^a Data represent the sample of survey respondents ages 18–54 years (N=6,707).

^b Weighted percentages were determined.

Later in the interview, of those respondents who met criterion A requirements, a subsample of potential 1-year case subjects (N=654) were asked the following four additional questions from the Sheehan Disability Scale (21) about the impairing effects of the stem symptom: “How much did your sadness/discouragement/lack of interest interfere with 1) your home management, like cleaning, shopping, and taking care of the house/apartment; 2) your ability to work; 3) your ability to form and maintain close relationships with other people; or 4) your social life?” The Sheehan Disability Scale uses 0 to 10 visual analogue scales for answers, with response options as follows: none (0), mild (1 to 3), moderate (4 to 6), severe (7 to 9), and very severe (10). Responses of ≥ 4 were considered to satisfy the impairment criterion.

Data Analysis

Statistical analyses were performed using the survey estimation procedures in Stata10 (22), which calculate weighted coefficients and use Taylor series linearization to determine standard errors to correct for the sampling design.

Results

Prevalence

Despite lack of a clinical significance criterion in the initial, DSM-III-R-based NCS and the addition of one in the DSM-IV-based NCS-R, the prevalence rates of adult depression in the 18- to 54-year-old age range rose substantially from 15.2% to 18.3%, respectively (Table 1).

Distress

The redundancy hypothesis for distress was strongly confirmed. Of the 2,071 respondents who experienced at least one mood symptom for 2 weeks (or multiple 3-day periods), 2,016 (97.2%) satisfied the distress clinical significance criterion. Only 55 individuals (2.8%) reported a mood symptom that did not satisfy the distress criterion.

Among the 2,071 case subjects evaluated for distress, 1,254 (60.5%) went on to qualify for a diagnosis of major depression, whereas 817 (39.5%) did not qualify for various reasons, mostly as a result of being subthreshold. Of

those who did not qualify, 765 (93.5%) satisfied NCS-R distress criteria. Thus, distress was an exceptionally poor indicator of diagnostic status.

Minor variations in distress criteria did not make a substantial difference to prevalence. For example, the threshold for the response of “rarely” for the third distress question seemed low. However, raising the threshold to “sometimes” had almost no effect (10 subjects with depression were eliminated).

Impairment

NCS-R impairment questions were asked of respondents satisfying symptom-duration criterion A requirements for depression. The redundancy prediction was strongly confirmed for impairment. Of the 1,542 individuals who satisfied criterion A requirements, 1,487 (96.2%) also satisfied clinical significance requirements for impairment (Table 1). The four Sheehan Disability Scale questions regarding impairment had little effect, adding only 19 impairment case subjects to those who already qualified on the basis of the interference with life and activities questions.

Effect of Severe Distress or Impairment

Despite false negative concerns, we recalculated the prevalence rates after raising the response thresholds for the distress and impairment questions to “severe,” “often,” and “a lot.” Raising the threshold for distress to “severe” eliminated some subjects but left many subthreshold sadness cases satisfying the criterion. Of 2,071 respondents reporting sadness episodes, 1,346 (64.9%) reported severe distress. Notably, of 817 respondents who reported non-major depression sadness, 461 (56.5%) reported severe distress. Using severe impairment alone for clinical significance, our results are consistent with those of previous studies. Of 1,542 subjects satisfying criterion A, 1,277 (82.7%) reported severe impairment. The overall effect on major depression was minimal. Of 1,254

subjects qualifying for the diagnosis of major depression, 1,168 (92.9%) satisfied “severe” clinical significance.

Discussion

Redundancy Thesis and Utility of the Depression Clinical Significance Criterion

The goal of the depression clinical significance criterion is to distinguish depressive disorder from symptomatically similar normal reactions. We found that significant distress accompanied almost all NCS–R cases of persistent sadness, and significant impairment accompanied almost all symptomatic cases satisfying criterion A. A disjunction of the two would cover even more cases.

The regular association of distress with sadness renders the distress criterion all but useless in distinguishing normal sadness from major depression. The clinical significance criterion’s disjunctive character means that the distress component’s weakness undermines the criterion’s overall validity, perhaps explaining why many researchers forgo the distress component (9, 15, 18). We found that only 4% of subjects satisfying the requirements for NCS–R criterion A were not impaired, which is less than the percentage of case subjects with uncomplicated bereavement who were excluded from NCS but would have likely represented false positive diagnoses satisfying NCS–R criterion A requirements in the present study.

Thus, the belief that clinical significance effectively addresses false positive diagnoses for major depression and reduces rates of community prevalence appears to be unsubstantiated. Strengthening the criterion is not necessarily a solution. Our analyses indicated that modestly raising low thresholds has little effect. Substantially raising severity thresholds reduces prevalence (4, 9, 18), but, as we demonstrated, does not eliminate subthreshold false positives and only minimally reduces the prevalence of major depression, whereas the number of false negatives may substantially increase (4).

Our findings raise questions about the common use of clinical significance to justify diagnosis of subthreshold depression under the category of “mood disorder not otherwise specified.” Such diagnoses, while allowing individuals who do not satisfy major depression criteria to get needed help, may be made without attention to whether an individual’s symptoms are caused by a dysfunction, as required by DSM-IV’s definition of disorder. The fact that virtually all individuals reporting extended sadness also reported significant distress suggests that the use of clinical significance to justify subthreshold diagnosis increases the risk of massive false positives.

Consideration of Contrary Claims

Our finding is contrary to claims made by some NCS–R researchers. Kessler et al. (19) noted that “previous methodological research...found that the NCS CIDI [Composite International Diagnostic Interview] overdiagnosed

MDD [major depressive disorder] because of false positive assessments of dysphoria and anhedonia” (19, p. 3097) and that the NCS–R’s clinical significance requirement for depression was an attempt to correct these problems. However, our results indicate that the NCS–R clinical significance criterion eliminates almost no case subjects reporting persistent dysphoria and anhedonia because distress is almost always present. Thus, any progress with false positives must be a result of other NCS–R features.

Was there such progress? Kessler et al. (19) reported, as evidence for false positive reduction, a lower NCS–R major depression prevalence rate relative to that of NCS: “The NCS–R MDD [major depressive disorder] prevalence estimates are intermediate between the ECA [Epidemiologic Catchment Area] and NCS estimates....The lower CIDI prevalence estimates than those in the NCS are consistent with the fact that these modifications operated largely by reducing false positive assessments” (19, p. 3103). This might apply to 1-year prevalence rates. However, our results indicate that the NCS–R adult rate of prevalence for lifetime major depression (18.3%) is higher than the NCS prevalence rate (15.2%). Published rates for lifetime NCS and NCS–R major depression are 14.9% (23) and 16.2% (19), respectively. This increase occurred despite changes that made the replicated survey symptom criteria more demanding than that of the initial survey (e.g., requiring observed psychomotor retardation and feelings of complete worthlessness). Thus, the claim of improved lifetime depression false positive reduction remains unsupported.

DSM-V Reassessment of the Clinical Significance Criterion: Generic or Disorder-by-Disorder?

How should DSM-V’s re-evaluation of the clinical significance criterion proceed? Should we accept the ICD’s generic rationale that diagnostic symptom/dysfunction assessment should be separated from considerations of role impairment?

To the contrary, the need for disorder-by-disorder evaluation is suggested by the ICD’s own continued use of social role functioning in some criteria sets. For example, conduct disorder requires “major violations of age-appropriate social expectations,” and reading disorder must “significantly interfere with academic achievement or activities of daily living that require reading skills.” These exceptions suggest that clinical significance is diagnostically useful in some categories. For some disorders, the only harm caused by the symptoms might lie in social role impairment, as in conduct disorder. Moreover, some role capacities are biologically shaped, and thus role impairment implies biological dysfunction. For example, if social phobia when interacting with family members impairs basic role functions such as parenting and sexual interaction, this is also biological dysfunction. Finally, sometimes role failure, even if it is not itself a biological dysfunction, is the only way to infer an underlying dysfunction. For example, failure to learn

to read, an invented social function, despite opportunity may imply an underlying dysfunction, and because the inferred dysfunction has no other known harmful effects the social impairment is important to diagnosis.

Even concerning redundancy, there are cross-category nuances. One might expect that a specific phobia entails distress in the form of object-triggered or anticipatory anxiety (in successful avoidance). The specific phobia criterion avoids redundancy by requiring role impairment or marked distress about having the condition. However, this solution creates a new problem. Distress about having a condition is of questionable validity in discriminating disorder from nondisorder (as in “ego-dystonic homosexuality”). Regarding role impairment, it might be argued that many specific phobias are instances of biologically prepared, normally distributed fears and thus are pathological only when they subvert other basic, biologically shaped functions, manifested in role impairment. Thus, within this category, impairment, if properly qualified, is diagnostically relevant.

These examples suggest that rather than wholesale removal that might yield unintended consequences, a disorder-by-disorder re-evaluation of the clinical significance criterion is preferable.

Conclusions

The NCS-R major depression clinical significance criterion does not substantially reduce prevalence rates, nor does it address presumed false positives in a community sample where false positives are most likely to occur, confirming the redundancy prediction. This evidence suggests that little or nothing will be lost if the criterion is eliminated in DSM-V, a possibility that should be considered by the Mood Disorders Work Group. New approaches to identifying false positive diagnoses for depression should be explored. Our findings suggest that the criterion might not be needed in some other categories as well.

In considering the future of the clinical significance criterion across categories, a disorder-by-disorder analysis is preferable to a generic approach because the relationship among diagnosis, symptoms, and impairment varies across disorders. Distress and impairment must remain important dimensions of assessment, as in ICD, irrespective of disorder status, and in some instances may continue to have diagnostic relevance, although clinical significance cannot substitute for evaluation of whether symptoms indicate mental dysfunction.

Finally, diagnosis and the need for treatment are not the same. Intense normal reactions to loss and stress can include distress, role impairment, and other deviations from homeostasis that can transiently resemble disorder (8). Access to professional intervention in such cases is desirable, even if no disorder is present.

Received April 22, 2009; revision received July 22, 2009; accepted Sept. 10, 2009 (doi: 10.1176/appi.ajp.2009.09040553). From the School of Social Work, Department of Psychiatry, and InSPIRES (Institute for Social and Psychiatric Initiatives—Research, Education and Services), New York University, New York; School of Social Administration, Temple University, Philadelphia; School of Social Work, Rutgers, The State University of New Jersey, New Brunswick, N.J. Address correspondence and reprint requests to Dr. Wakefield, 309 W., 104 St. #9C, New York, NY 10025; jerome.wakefield@nyu.edu (e-mail).

The authors report no financial relationships with commercial interests.

References

1. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 4th ed—Text Revision. Washington, DC, American Psychiatric Publishing, 2000
2. Widiger TA, Frances AJ, Pincus HA, Ross R, First MB, Davis WW, Kline M: DSM-IV Sourcebook, vol 4. Washington, DC, American Psychiatric Publishing, 1998
3. Spitzer RL, Wakefield JC: DSM-IV diagnostic criterion for clinical significance: Does it help solve the false positives problem? *Am J Psychiatry* 1999; 156:1856–1864
4. Beals J, Novins DK, Spicer P, Orton HD, Mitchell CM, Baron AE, Manson SM, Big Crow CK, Buchwald D, Chambers B, Christensen ML, Dillard DA, DuBray K, Espinoza PA, Flemming CM, Frederick AW, Gurley D, Jervis LL, Jim SM, Kaufman CE, Keane EM, Klein SA, Lee D, McNutly MC, Middlebrook DL, Moore LA, Nez TD, Norton IM, Randall CJ, Sam A, Shore JH, Simpson SG, Yazzie LL; AI-SUPERPFP Team: Challenges in operationalizing the DSM-IV clinical significance criterion. *Arch Gen Psychiatry* 2004; 61:1197–1207
5. Schneider ME: DSM-V Task Force weighing changes. *Clin Psychiatry News* 2009; 3:1–4
6. Ustun B, Kennedy C: What is “functional impairment”? Disentangling disability from clinical significance. *World Psychiatry* 2009; 8:82–85
7. Wakefield JC: Disability and diagnosis: Should role impairment be eliminated from DSM/ICD diagnostic criteria? *World Psychiatry* 2009; 8:87–88
8. Regier DA, Kaelber CT, Rae DS, Farmer ME, Knauper B, Kessler RC, Norquist GS: Limitations of diagnostic criteria and assessment instruments for mental disorders: implications for research and policy. *Arch Gen Psychiatry* 1998; 55:109–115
9. Narrow WE, Rae DS, Robins LN, Regier DA: Revised prevalence estimates of mental disorders in the United States: using a clinical significance criterion to reconcile 2 survey’s estimates. *Arch Gen Psychiatry* 2002; 59:115–123
10. Robins LN, Regier DA: *Psychiatric Disorders in America*. New York, Free Press, 1991
11. Kessler RC, McGonagle KA, Zhao SN, Nelson CB, Hughes M, Eshelman S, Wittchen HU, Kendler KS: Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: results from the National Comorbidity Survey. *Arch Gen Psychiatry* 1994; 51:8–19
12. Wakefield JC: DSM-IV: Are we making diagnostic progress? *Contemp Psychol* 1996; 41:646–652
13. Wakefield JC: Diagnosing DSM-IV, part 1: DSM-IV and the concept of mental disorder. *Behav Res Ther* 1997; 35:633–650
14. Zimmerman M, Chelminski I, Young D: On the threshold of disorder: a study of the impact of the DSM-IV clinical significance criterion on diagnosing depressive and anxiety disorders in clinical practice. *J Clin Psychiatry* 2004; 65:1400–1405
15. Mojtabai R: Impairment in major depression: implications for diagnosis. *Compr Psychiatry* 2001; 42:206–212

16. Wakefield JC, Spitzer RL: Lowered estimates—but of what? *Arch Gen Psychiatry* 2002; 59:129–130
17. Spitzer RL: Diagnosis and need for treatment are not the same. *Arch Gen Psychiatry* 1998; 55:120
18. Slade G, Andrews T, Issakedis C: Deconstructing current comorbidity: data from the Australian National Survey on Mental Health and Well-Being. *Br J Psychiatry* 2002; 181:306–314
19. Kessler RC, Berglund P, Demler O, Jen R, Koretz DS, Merikangas KR, Rush AJ, Walters EE, Wang PS; National Comorbidity Survey Replication Team: The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS–R). *JAMA* 2003; 289:3095–3105
20. Wakefield JC, Schmitz MF, First MB, Horwitz AV: Extending the bereavement exclusion for major depression to other losses: evidence from the National Comorbidity Survey. *Arch Gen Psychiatry* 2007; 65:433–440
21. Leon AC, Olfson M, Portera L, Farber L, Sheehan DV: Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. *Int J Psychiatry Med* 1997; 27:93–105
22. Stata Corp: Stata Statistical Software, Release 10. College Station, Tex, Stata Corp, 2007
23. Kessler RC, Nelson CB, McGonagle KA, Lin J, Swartz M, Beazer DG: Comorbidity of DSM-III-R major depressive disorder in the general population: results from the United States National Comorbidity Survey. *Br J Psychiatry Suppl* 1996; 30:17–30