When does depression become a disorder? Using recurrence rates to evaluate the validity of proposed changes in major depression diagnostic thresholds

JEROME C. WAKEFIELD¹⁻⁴, MARK F. SCHMITZ^{3,5}

¹Silver School of Social Work, New York University, 1 Washington Square North, New York, NY; ²Department of Psychiatry, School of Medicine, New York University, New York, NY; ³InSPIRES (Institute for Social and Psychiatric Initiatives–Research, Education and Services), New York University, New York, NY; ⁴Division of Clinical Phenomenology, Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, NY; ⁵School of Social Work, Temple University, Philadelphia, PA, USA

High community prevalence estimates of DSM-defined major depressive disorder (MDD) have led to proposals to raise MDD's diagnostic threshold to more validly distinguish pathology from normal-range distress. However, such proposals lack empirical validation. We used MDD recurrence rates in the longitudinal 2-wave Epidemiologic Catchment Area Study to test the predictive validity of three proposals to narrow MDD diagnosis: a) excluding "uncomplicated" episodes (i.e., episodes that last no longer than 2 months and do not include suicidal ideation, psychotic ideation, psychomotor retardation, or feelings of worthlessness); b) excluding mild episodes (i.e., episodes with only five to six symptoms); and c) excluding nonmelancholic episodes. For each proposal, we used lifetime MDD diagnoses at wave 1 to distinguish the group proposed for exclusion, other MDD, and those with no MDD history. We then compared these groups' 1-year MDD rates at wave 2. A proposal was considered strongly supported if at wave 2 the excluded group's MDD rate was not only significantly lower than the rate for other MDD but also not significantly greater than the no-MDD-history group. Results indicated that all three excluded groups had significantly lower recurrence rates than other MDD (uncomplicated vs. complicated, 3.4% vs. 14.6%; mild vs. severe, 9.6% vs. 20.7%; nonmelancholic vs. melancholic, 10.6% vs. 19.2%, respectively). However, only uncomplicated MDD's recurrence rate was also not significantly greater than the MDD occurrence rate for the no-MDD-history group (3.4% vs. 1.7%, respectively). This low recurrence rate resulted from an interaction between uncomplicated duration and symptom criteria. Multiple-episode uncomplicated MDD did not entail significantly elevated recurrence over single-episode cases (3.7% vs. 3.0%, respectively). Uncomplicated MDD's general-distress symptoms, transient duration, and lack of elevated recurrence suggest it may generally represent nonpathologic intense sadness that should be addressed in treatment guidelines and considered for exclusion from MDD diagnosis to increase the validity of the MDD/normal sadness boundary.

Key words: Major depression, recurrence, uncomplicated depression, diagnosis, validity, severity, melancholic depression

(*World Psychiatry* 2013;12:44–52)

Establishing a valid diagnostic boundary between major depressive disorder (MDD) versus intense normal sadness or mild adjustment disorder that generally does not require intervention has proven challenging. The problem is that nonpathological reactions to major losses and stressors possess many of the same general-distress symptoms as depressive disorder (1,2).

The DSM's five-out-of-nine-symptoms-for-2-weeks descriptive MDD diagnostic criteria were originally formulated to distinguish depressive conditions from other mental disorders among severely ill populations for research purposes (3). No empirically grounded rationale exists for the current boundary in distinguishing between disorder and normal distress in community populations (4,5). Professional and public concerns about the validity of MDD criteria have been provoked by dramatically rising community prevalence estimates of DSM-specified MDD as the methodology of epidemiologic studies improves, from 2–3% some decades ago (6) to current estimates of half or more of all individuals meeting criteria over a lifetime (7–9). This diagnostic conundrum takes on additional importance at a time of increased depression screening in schools and physicians' offices, making the validity of diagnostic criteria relevant to everyone in the community, not just those presenting for treatment.

In light of these concerns, some depression researchers argue that the diagnostic threshold currently set by DSM criteria for MDD is too low, yielding false-positive diagnoses that misclassify considerable numbers of normally distressed individuals as having MDD (2,10–12). However, specific proposals for raising the MDD diagnostic threshold remain unsupported by persuasive empirical evidence regarding validity.

Crystallizing the need for a scientific approach to the MDD boundary question, Maj (13,14) recently challenged depression researchers to undertake the difficult task of establishing an empirically based answer to the question, When does depression become a mental disorder? We address Maj's challenge by empirically evaluating the predictive validity of three suggestions derived from the literature for narrowing the category of major depression: a) extend the DSM-IV's (15) "bereavement exclusion" to all MDD, thus excluding from MDD diagnosis all "uncomplicated" episodes (i.e., episodes lasting no longer than 2 months and lacking suicidal ideation, psychotic symptoms, psychomotor retardation, and feelings of worthlessness); b) exclude all mild episodes (i.e., episodes having only five or six symptoms) from MDD diagnosis; and c) equate MDD with melancholic depression, thus excluding from diagnosis all depression that is nonmelancholic.

To evaluate these proposals, we examined recurrence rate as a predictive validator of disorder. No single outcome variable can entirely resolve the question of where to place the diagnostic threshold between normal sadness and MDD to optimally prevent false positives while avoiding false negatives. However, substantially elevated recurrence rates are characteristic of MDD in both clinical and community samples (16-23), justifying the standard conceptualization of MDD as an episodic and chronic condition (24-26) for which prevention of recurrence is a major treatment goal (25,27-29). Moreover, recurrence is often interpreted as the prime evidence allowing one to infer that there exists a latent dysfunction that repeatedly generates the MDD symptoms, thus verifying the presence of a disorder rather than a normal emotional response (16,30).

In this study, we examine two progressively stronger indicators that a group's recurrence diverges from standard MDD. First, we examine whether subgroups of MDD proposed for exclusion have lower recurrence rates than other MDD, suggesting a difference in kind of condition worthy of further investigation. Second – and much more stringently – we examine whether the groups have rates that are not significantly elevated over the occurrence of MDD in the general population, calling into question the presence of pathology. Many variables are known to predict differentially lower MDD recurrence rates (16-18,20). However, no study has identified a substantial subset of MDD that does not have an elevated risk for recurrence above general population risk levels, with the one limited exception of uncomplicated bereavement-related depression (23,31). The lack of elevated recurrence runs contrary to the standard characterization of MDD as a highly recurrent condition in which the end of a depressive episode means “recovery from the episode, not from the illness *per se*” (30). Quite aside from diagnostic boundary issues, variables indicating lack of elevated recurrence rates or substantially lower recurrence rates would be of potential importance to clinicians in judging prognosis, evaluating the risk/reward ratio of treatments having potential negative side effects, and deciding about the timing of treatment termination.

METHODS

Sample

To evaluate the proposals for revising the MDD threshold, we used longitudinal community data from the 2-wave Epidemiologic Catchment Area Study (ECA, 32). We selected the ECA dataset for this analysis because it is the only currently publicly available nationally representative American longitudinal dataset with the information needed to perform the study's analyses.

The ECA's sampling and data collection are described extensively in other publications (32). Briefly, respond-

ents (ages 18 to 98 years) at five sampled sites (Baltimore, Durham, Los Angeles, St. Louis, and New Haven) were interviewed face-to-face twice, approximately 12 months apart (waves 1 and 2), between 1980 and 1985. Lifetime and 1-year DSM-III (33) diagnostic data were obtained at both the waves. Our analytic sample includes only cases with valid responses at both waves ($n=18,943$). Data were weighted to account for selection and nonresponse effects and to match age, sex, and race distributions in the 1990 US census, in order to provide nationally representative estimates.

Major depression

The present study compares ECA respondents with no history of MDD at wave 1 ($n=18,239$; 52.1% female, 68.1% white, 68.2% high school degree or more, average age 43.0 years old) to various subgroups of those with histories of MDD at wave 1 ($n=704$; 73.3% female, 77.9% white, 81.5% high school degree or more, average age 37.0 years).

The ECA's DSM-III-based diagnostic criteria for MDD required ever having 2 weeks or more of sadness accompanied by at least four other symptoms. Symptoms due to physical injury or use of drugs, medication, or alcohol were excluded. Psychotic and organic episodes were also excluded. In addition, the ECA ensures clinical significance by requiring that each symptom must be adequately severe, defined as either told a doctor or another professional about the symptom, took medication more than once for the symptom, or the symptom interfered with life or activities a lot. This requirement, which relies heavily on professional help-seeking, is more demanding than the DSM-IV clinical significance requirement of distress or role impairment, which is known to have little impact on prevalence (34-36). Thus, one can expect the ECA sample to be more severe than a usual DSM-defined sample.

The ECA also imposed a syndromal severity requirement that is essentially identical to the symptom requirement in relying largely on help-seeking but applied to the overall syndrome. We suspended the syndrome-severity requirement, because it was largely redundant with the already demanding symptom-severity requirement, and it pushes the sample even further in a direction of greater severity. The suspension of the syndrome-severity criterion modestly increased the MDD sample size from 945 to 1065.

Previous studies that have attempted to examine the trajectory of depression have sometimes failed to distinguish between recurrence and chronicity as explanations of MDD at the follow-up interview, due to lack of information about the trajectory between interviews (16,37). Because our interest was specifically in measuring recurrence of MDD at wave 2 in those with remitted lifetime MDD at wave 1, we ensured an appropriate sample by eliminating from the MDD analytic sample all

those cases in which the respondent reported experiencing an ongoing episode at the time of the wave 1 interview. The elimination of wave 1 ongoing cases reduced the size of our primary MDD sample from 1065 to 704.

Primary exclusion proposals to be tested

The primary ideas for reducing the scope of MDD that are tested in the present study were: a) excluding “uncomplicated” episodes (i.e., episodes that last no longer than 2 months and do not include suicidal ideation, psychotic ideation, psychomotor retardation or feelings of worthlessness); b) excluding mild episodes (i.e., episodes with only five to six symptoms); and c) excluding nonmelancholic episodes.

In evaluating each of these proposals, multiple episode cases were classified in accordance with the classification of the individual’s worst episode, for which we had detailed data. This was based on the assumption that if the worst episode was uncomplicated, or mild, or nonmelancholic, then it is unlikely that the individual’s other episodes would be complicated, severe, or melancholic, respectively, because those would generally be judged worse.

Uncomplicated depression

This proposal extends to all potential MDD the exclusion for “uncomplicated” episodes that the DSM currently applies only to bereavement-related depression in the “bereavement exclusion” (criterion E of major depressive episode). Exclusion as an uncomplicated episode requires brief duration (no more than 2 months) and the absence of four specific symptoms (psychomotor retardation, suicidal ideation, psychotic ideation, sense of worthlessness) identified by the DSM as “uncharacteristic” of normal sadness.

Uncomplicated MDD episodes thus include only “general distress” symptoms that are common in normal reactions to stress, such as sadness, insomnia, loss of interest in usual activities, decrease in appetite, and difficulty in concentrating. The “uncomplicated” criteria were created to distinguish normal range from disordered depressive episodes during bereavement (38,39), and the proposal to be evaluated is that the distinction should be extended to MDD in general (40).

The DSM-IV also specifies that, in order to be excluded as uncomplicated, a case must lack marked impairment. We were unable to include this feature in our criteria due to lack of appropriate data, because of the way the ECA severity questions were structured with skip-outs. This deviation from DSM-IV increases the difficulty of supporting the “uncomplicated” hypothesis, because it allows episodes with more marked impairment – and thus more likelihood of disorder and recurrence – to qualify as excluded. However, this limitation is consistent with both the ICD’s and

DSM’s aspiration to separate role impairment from disorder diagnosis.

Mild (low-symptom) MDD

It has been suggested that the MDD threshold should be raised to seven symptoms from the current five (41,42) “to lower the risk of ‘over-diagnosis’ or pathologizing ‘normal’ depression or sadness” (42). This proposal is supported by findings over the years suggesting a discontinuity between MDD episodes with five to six and seven to nine symptoms (4,5,43,44). As Maj (13) notes, “when a point of rarity has been reported, it has usually corresponded to a threshold higher than that fixed by the DSM-IV”. Moreover, previous studies suggest that severity in the form of number of symptoms predicts recurrence (16,18).

We considered that a narrower exclusion might yield better results. Thus, we also tested a variant proposal to exclude only cases with five symptoms, setting the MDD threshold at six symptoms.

Nonmelancholic MDD

The third proposal is to restrict MDD to the traditional category of melancholic depression, long considered the quintessential form of depressive disorder, and frequently targeted in depression scale development and research (45). The rest of what we presently call depression would be conceived as various other forms of dysphoria (e.g., “In reality, nonmelancholia is a mixture of dysphoria, anxiety, and depressive character”(45)). Maj (13) reviews some research suggesting that there is “a subtype [of MDD], grossly corresponding to DSM-IV major depression with melancholia, which may be qualitatively different from normal sadness” and argues that melancholia warrants further examination in reconsidering the normal/disordered boundary.

The ECA did not define melancholic depression and did not contain all the symptom questions relevant to diagnosis of DSM melancholic depression. As an approximation, we defined melancholic depression as episodes that included three or more from a list of five melancholic-like ECA symptom items: fatigue (“felt tired out all the time”), psychomotor retardation (“talked or moved more slowly than is normal for you”); agitation (“had to be moving all the time, that is, you could not sit still and paced up and down?”); decreased interest in sex (“your interest in sex was a lot less than usual”); and slowed thinking (“your thoughts came slower than usual or seemed mixed up”). Nonmelancholic depression was defined as any MDD cases that had less than three melancholic symptoms.

We considered that a narrower exclusion might yield better results. Thus, we also tested a variant proposal to exclude only cases with less than two melancholic

symptoms, setting the threshold for melancholic depression at two melancholic symptoms rather than three.

Hypothesis testing procedure

Testing of primary hypotheses

Using lifetime MDD diagnoses at wave 1, for each of our five proposals (i.e., the three primary proposals plus the two narrower variants), we identified the subgroup of MDD proposed for exclusion, all other MDD, and those with no history of MDD. We then compared 1-year MDD recurrence/occurrence rates at wave 2 (1 year later) for these three wave 1 groups and tested differences in rates for significance.

We considered a proposal for excluding a subgroup from MDD to be disconfirmed if the subgroup's wave 2 recurrence rate was not significantly different from the rate for other MDD; to be weakly confirmed as worth further exploration if the subgroup's recurrence rate was significantly and substantially lower than the rate for other MDD; and to be strongly supported if the subgroup had a recurrence rate that was not only significantly lower than the recurrence rate for other MDD but also not significantly higher than the 1-year wave 2 MDD occurrence rate for those who had no prior history of MDD at wave 1.

Component analysis of "uncomplicated depression"

In a post-hoc analysis, we examined whether the results for "uncomplicated depression" were due to just one of the two criteria (brief, no uncharacteristic symptoms) or to an interaction between the two. To examine this question, we recalculated recurrence outcomes separately for the two components.

Increasing the "uncomplicated" duration limit to 6 months

Recent evidence regarding bereavement-related depression (31,46,47) suggests that the current 2-month DSM-IV duration threshold for uncomplicated episodes may be too low and that more valid cut-points may exist at higher durations. In a post hoc analysis, we tested whether our results for uncomplicated MDD were preserved if the duration threshold was lengthened to 6 months.

Reanalysis using the unadjusted ECA MDD sample

As noted, we made two changes to ECA diagnostic criteria in deriving our sample: we eliminated cases that reported ongoing episodes at wave 1 and we suspended the syndrome-level severity requirement. Concerned that these changes inadvertently might have introduced a

bias into the results, we reanalyzed the data using the unadjusted ECA MDD sample, applying the ECA's syndrome-severity requirement and including MDD cases with ongoing episodes at wave 1. This increased our MDD sample from 704 to 945 cases. Both the changes would be expected to inflate the recurrence rates of proposed-for-exclusion subgroups, increasing the difficulty of confirming the exclusion proposals.

Recurrence in multiple-episode versus single-episode cases

We examined whether the frequent finding in the literature that multiple-episode cases predict higher recurrence held for our sample and was similar in uncomplicated and other MDD cases.

Statistical analysis

All data were weighted and corrected for sampling design to approximate a representative sample of the American population. Statistical analyses used Stata 12 survey estimation procedures, which calculate weighted coefficients and used Taylor series linearization to calculate standard errors (48).

As we hypothesized specific directions of the tested relationships, the analyses included one-tailed t-tests of statistical significance in mean differences for the percentage of respondents reporting a 1-year MDD at wave 2, between the different categories of cases.

RESULTS

Evaluation of five proposals for narrowing MDD's scope

We evaluated five proposals for narrowing the scope of MDD, using recurrence as a validator. Only one proposal, to exclude uncomplicated MDD, passed our most stringent test. Among those with lifetime uncomplicated MDD at wave 1, the likelihood of MDD recurrence at wave 2 was 3.4%, which was not significantly greater than the likelihood of 1-year occurrence of MDD at wave 2 among those with no history of MDD at wave 1 (1.7%) and was significantly lower than the recurrence rate at wave 2 of all other MDD (14.6%) (Table 1).

The proposals to exclude mild cases (defined as MDD with five to six symptoms or in a narrower variant as five-symptom MDD) and nonmelancholic cases (defined as having no more than two melancholic symptoms or in a narrower variant as having not more than 1) failed to pass this strong test (Table 1). This was because the recurrence rates for these proposed excluded groups (ranging from 8.8% to 10.6%) were in every case

Table 1 Recurrence rates of three subgroups proposed for exclusion from MDD, compared to occurrence rates of those with no history of MDD and recurrence rates of other (nonexcluded) MDD

MDD subgroup proposed for exclusion	Subgroup's recurrence rate	Is the subgroup's recurrence rate different from the No History of MDD group's rate=1.7% (1.4%, 1.9%), [n=18,239]	Recurrence rate for All Other MDD (except the subgroup)	Is the subgroup's recurrence rate different from the recurrence rate for All Other MDD?
Uncomplicated (brief and no psychomotor retardation, suicidal ideation, or felt worthless/guilty)	3.4 (−0.4, 7.1) [n=88]	Not different from No History	14.6 (11.3, 17.9) [n=616]	Different from Other MDD
Low-symptom (five to six MDD symptom groups)	9.6 (6.6, 12.6) [n=487]	Different from No History	20.7 (14.1, 27.3) [n=217]	Different from Other MDD
Narrowed low-symptom (five MDD symptom groups)	8.8 (5.6, 11.9) [n=310]	Different from No History	16.2 (11.9, 20.4) [n=394]	Different from Other MDD
Nonmelancholic (at most two melancholic symptoms)	10.6 (7.3, 13.9) [n=503]	Different from No History	19.2 (12.8, 25.6) [n=201]	Different from Other MDD
Narrowed nonmelancholic (at most one melancholic symptom)	10.4 (5.2, 15.7) [n=279]	Different from No History	14.9 (11.3, 18.4) [n=425]	Different from Other MDD

Significant differences between groups tested using one-tailed t-test, $p < 0.05$

significantly higher than the occurrence rate for those with no MDD history (1.7%). However, the recurrence rates for these groups were in every case significantly and substantially lower than the recurrence rates for other MDD (ranging from 14.9% to 20.7%), and to this extent, these proposals warrant further examination for implications about the MDD boundary.

Component analysis of “uncomplicated depression”

We explored whether the “uncomplicated” hypothesis's confirmation was due to just one of the two criteria (brief duration, no uncharacteristic symptoms) or an interaction between the two components. Although uncomplicated MDD ($n=88$) had a recurrence rate of 3.4%, brief cases ($n=326$) had a recurrence rate of 12.6%, and cases lacking the uncharacteristic symptoms ($n=153$) had a recurrence rate of 7.3% (Table 2). Both of the components' rates were significantly higher than the no-MDD-history group. Clearly, the dramatic result for uncomplicated MDD in which its recurrence rate is no higher than the general population's MDD rate is due to an interaction, with both duration and symptom requirements contributing to this result.

Reanalysis of the “uncomplicated MDD” proposal with 6-month duration limit

On the basis of the earlier studies, we hypothesized that increasing the duration criterion for “uncomplicated episode” to 6 months would not alter the results of recurrence-rate comparisons. This prediction was confirmed. When uncomplicated duration was allowed to be up to 6 months, the number of uncomplicated cases

increased by about 38% to 121 cases, yet the results of recurrence comparisons remained the same. The recurrence rate for uncomplicated cases still did not differ significantly from no-MDD-history cases (5.7% vs. 1.7%), and was significantly lower than other MDD (14.7%).

Reanalysis using the unadjusted ECA MDD sample

To check whether the adjustments we made to the ECA MDD criteria might have inadvertently biased the results in our favor, we repeated the recurrence analyses using the unadjusted ECA MDD sample. The results of comparisons (not displayed) were identical to the results of our primary analyses. Once again, only uncomplicated MDD ($n=95$) passed our strongest test as a potential variant of normal sadness. Uncomplicated MDD had an unadjusted recurrence rate not significantly different from the no-MDD-history group's occurrence rate and significantly lower than the rate for other MDD (1.7%, 5.0%, and 17.3%, respectively). All other subgroups proposed for exclusion had recurrence rates (range 7.2% to 10.3%) significantly higher than no-MDD-history's 1.7% rate but significantly lower than other MDD cases (range 17.7% to 19.1%).

Because this analysis of wave 2 MDD recurrence includes cases that involved an ongoing MDD episode at wave 1, in effect it evaluates a combination of recurrence and chronicity. This expanded test still yields the same verdict that uncomplicated MDD is not different from no-MDD-history but is very different from other MDD.

Recurrence in single-episode versus multiple-episode cases

Many studies indicate that multiple-episode MDD cases predict substantially higher recurrence. We, thus,

Table 2 Follow-up analyses: recurrence rates of components for uncomplicated criteria compared to occurrence rates of those with no history of MDD and recurrence rates of other (nonexcluded) MDD

Component analysis of uncomplicated	Subgroup's recurrence rate	Is the subgroup's recurrence rate different from the No History of MDD group's rate=1.7% (1.4%, 1.9%), [n=18,239]	Recurrence rate for All Other MDD (except the subgroup)	Is the subgroup's recurrence rate different from the recurrence rate for All Other MDD?
Brief (2 months or less duration)	12.6 (8.9, 16.3) [n=326]	Different from No History	13.6 (9.2, 18.0) [n=378]	Not different from Other MDD
No uncharacteristic symptoms (psychomotor retardation, suicidal ideation, or felt worthless/guilty)	7.3 (2.6, 11.9) [n=154]	Different from No History	14.7 (11.2, 18.1) [n=550]	Different from Other MDD

Significant differences between groups tested using one-tailed t-test, $p < 0.05$

examined whether recurrence for uncomplicated MDD was predominantly due to multiple-episode cases. Although the low recurrence rate limited this analysis, we found recurrences similarly distributed between single- and multiple-episode cases. Of 43 single-episode uncomplicated wave 1 cases, two had wave 2 1-year MDD (3.0%; 95% CI: -1.2, 7.3); of 45 multiple-episode uncomplicated wave-1 cases (median episodes=3), again two cases had 1-year wave-2 MDD (3.7%; 95% CI: -2.2, 9.6). In each instance, one wave 2 case was complicated and one uncomplicated. This pattern of data is consistent with the situational determination of depressive episodes, in which no enduring tendency to recurrence is sustained despite multiple past episodes.

The situation was strikingly different with respect to other (complicated) MDD cases, in which multiple-episode wave 1 MDD (median episodes=3) had a significantly and substantially higher 1-year wave 2 recurrence rate (19.1%; 95% CI: 12.2, 24.1) than did single-episode cases (7.2%; 95% CI: 3.3, 11.2). The multiple-episode uncomplicated cases' substantially lower recurrence rate as compared to other multiple-episode cases offers a hitherto unnoticed clinical differentiation. In sum, uncomplicated MDD multiple-episode histories do not predict higher recurrence than single-episode histories, whereas other MDD multiple-history episodes do predict higher recurrence.

DISCUSSION

Recurrence is a cardinal feature of depressive disorder and can be used to evaluate the plausibility of proposed changes to the MDD diagnostic threshold. Our analysis identified a subset of current MDD – namely, uncomplicated MDD – for which recurrence was no greater than background occurrence of MDD in the general population who never had MDD, a criterion frequently men-

tioned for recognizing normal intense suffering versus depressive disorder (23,31,49).

Another frequently mentioned criterion is the “similarity thesis”, according to which the strongest evidence that a subgroup of MDD is a form of pathology is that it is similar on validators to other MDD (22,50). Our results strongly disconfirm the similarity thesis for uncomplicated versus other MDD in terms of recurrence rates. The lack of a difference in recurrence rate between multiple-episode and single-episode uncomplicated cases also marks uncomplicated MDD as importantly dissimilar from other MDD. The other hypotheses we examined displayed some differences from other MDD as well, but the differences were not as consistent or pronounced, so the status of these other proposals remains more ambiguous.

A series of additional analyses confirmed the stability of our results. We lowered the thresholds for the low-symptom and nonmelancholic proposals, reanalyzed all the proposals using the original ECA MDD sample rather than our adjusted sample, and tried a 6-month duration in the definition of “uncomplicated”, all without changing the essential results. We explored the source of our “uncomplicated” result by analyzing each definitional component separately, and found that support for the “uncomplicated” hypothesis resulted from an interaction between the two components. Our results in shifting the uncomplicated duration threshold to 6 months suggest that further research should explore the proper duration threshold for MDD and for any potential “uncomplicated MDD” exclusion.

Our results underscore the heterogeneity of current DSM MDD (12) and offer findings regarding recurrence risk that are clinically relevant irrespective of the issue of MDD diagnostic boundaries. Monroe and Harkness (20) express concern that “research focusing on depression as a recurrent condition has generally failed to reveal any useful early indicators of risk for recurrence”. Our results reveal several such indicators of differential recurrence risk.

An episode can be judged uncomplicated only if it does not persist past 2 months. This presents a challenge to the clinician who attempts to apply these results, because of the need to monitor the patient, using “watchful waiting” until the condition remits or the 2-month duration is reached. The challenge is, however, limited. About 78% of our MDD cases displayed one or more of the uncharacteristic symptoms and so would be diagnosed as MDD on that basis alone, without reference to duration (Table 2). Of the remaining 22% who displayed no uncharacteristic symptoms, about 60% (13% of the MDD sample) remitted by 2 months and qualified as “uncomplicated” cases, whereas the remainder persisted beyond the 2-month duration limitation for uncomplicated cases. Consequently, briefly monitoring such individuals may be crucial to optimal formulation of prognosis and diagnosis. In some instances, treatment for symptom relief may be advisable much sooner, whether the condition is judged normal or disordered. Further research should be aimed at identifying low-recurrence uncomplicated-type MDD cases with minimal use of the duration threshold, so that adequate assessment need not depend on monitoring over time.

Limitations of the ECA dataset and of our measures must be considered in interpreting our findings. A major limitation is the ECA’s 1-year window for assessing recurrence. Further studies are needed to see if our results are replicated at longer follow-up periods. Another limitation is that diagnoses were based on respondents’ fallible recollections of symptoms.

The ECA’s DSM-III-based MDD diagnostic algorithm lacked the DSM-IV’s clinical significance criterion, and the ECA’s structure also made it impossible to apply the DSM’s impairment criterion for uncomplicated MDD. However, the ECA’s demanding symptom severity criteria more than compensated. Indeed, the ECA sample is likely more severe on average than DSM MDD, so one would expect this study to offer an especially demanding test of the examined exclusion proposals.

The DSM criteria for uncomplicated MDD require that there be no morbid preoccupation with worthlessness – a more demanding symptom level than the ECA “felt worthless” symptom that we used as a proxy. Also, the ECA worthlessness item included feeling guilty or sinful as alternatives. Thus, the ECA worthlessness item likely excluded somewhat fewer cases from MDD diagnosis than the stricter DSM criterion. It seemed preferable to use this weaker version rather than to ignore the worthlessness criterion altogether, given that since Freud such feelings of low self-esteem have been considered a crucial indicator that normal sadness may have transformed into pathology.

Because ECA severity relied partly on service contact, but the severity responses were not encoded in the ECA data set, we could not analytically distinguish treatment effects from MDD’s inherent course. Thus, treatment

might have influenced the course of these depressions and their recurrences. However, previous studies indicate that recurrence remains high in treated MDD samples and is not reduced to general population levels (25,27-29). Thus, treatment effects would be unlikely to explain why the uncomplicated group’s recurrence is not only substantially less than other MDD but also not different from the general population. The severity indicator does reveal that uncomplicated MDD cases are often seen by professionals, thus that these uncomplicated cases are clinically relevant.

Recurrence as a validator has the advantage that it is logically independent of the criteria used to define uncomplicated, low-symptom, and nonmelancholic MDD, so its use to evaluate these proposals does not tautologically bias the results. Moreover, recurrence is relevant to distinguishing normality from disorder, whereas many other commonly used validating variables, from lack of social support to suppressed immunological function, correlate both with disorder and intense normal distress and so do not distinguish these hypotheses. Previous attempts to address perceived MDD overdiagnosis generally focused on enhanced clinical significance/impairment criteria to raise the MDD diagnostic threshold (51). This strategy proved both conceptually misguided and empirically ineffectual (34-36,52-55).

However, recurrence as a validator imposes some limitations. Strictly speaking, recurrence is neither necessary nor sufficient for disorder. It is not necessary because in principle there can be one-episode genuine depressive disorders that do not confer a liability to recurrence (20). It is not sufficient because recurrence is conceptually ambiguous. It may represent the reemergence of an underlying dysfunction and thus indicate disorder, but it also might indicate, for example, the repetition of an earlier normal reaction due to either a chronic stressor or the occurrence of a related new stressor (e.g., loss of a spouse may lead to financial problems that in turn cause loss of a home, with each new stressor triggering distress). Future studies would benefit from a more nuanced, contextualized approach to understanding recurrence.

An additional limitation is that, due to the nature of the ECA dataset, we were unable to distinguish stress-triggered from nontriggered “endogenous” depressive episodes, so “uncomplicated” was applied without context sensitivity. However, previous studies indicate that application of DSM MDD criteria to community samples yields cases in which the vast majority have episodes triggered by a stressor. For example, in the National Comorbidity Survey, only 5% of MDD cases reported no trigger for their depressive feelings (38). Moreover, previous analyses indicate that almost no endogenous cases satisfy “uncomplicated” criteria. Thus, this gap in ECA data likely did not substantially influence the present study’s results.

In conclusion, using elevated recurrence rates to predictively validate disorder, our results indicate that eliminating uncomplicated episodes would increase the validity of MDD diagnosis. The results also confirm that less-severe forms of MDD, such as low-symptom and non-melancholic MDD, have lower recurrence rates than other MDD and warrant further exploration regarding the proper threshold for distinguishing MDD from normal sadness.

Uncomplicated MDD includes only symptoms common in normal distress, is transient, and, in our ECA sample, does not raise recurrence rates significantly above population levels. If these findings are replicated, there is no empirical rationale for generally distinguishing uncomplicated MDD from intense normal sadness. Consideration should thus be given to eliminating uncomplicated episodes as a class from MDD diagnosis. Exceptional cases that appear pathological despite their uncomplicated status could then be diagnosed under mood disorders not otherwise specified. Alternatively, treatment guidelines should emphasize stepped treatment, in light of the dramatic differences in prognosis between uncomplicated and other MDD.

References

1. Clayton P, Desmarais L, Winokur G. A study of normal bereavement. *Am J Psychiatry* 1968;125:168-78.
2. Regier DA, Kaelber CT, Rae DS et al. Limitations of diagnostic criteria and assessment instruments for mental disorders: implications for research and policy. *Arch Gen Psychiatry* 1998;55:109-15.
3. Kendler KS, Munoz RA, Murphy G. The development of the Feighner criteria: a historical perspective. *Am J Psychiatry* 2010;167:134-42.
4. Kendler KS, Gardner CO. Boundaries of major depression: an evaluation of DSM-IV criteria. *Am J Psychiatry* 1998;155:172-7.
5. Kessler RC, Zhao SY, Blazer DG et al. Prevalence, correlates, and course of minor depression and major depression in the National Comorbidity Survey. *J Affect Disord* 1997;45:19-30.
6. Klein DF, Thase M. Medication versus psychotherapy for depression: progress notes. *Am Soc Clin Psychopharmacol* 1997;8:41-7.
7. Eaton WW, Anthony JC, Gallo J et al. Natural history of Diagnostic Interview Schedule/DSM-IV major depression: the Baltimore Epidemiologic Catchment Area follow-up. *Arch Gen Psychiatry* 1997;54:993-9.
8. Eaton WW, Neufeld K, Chen LS et al. A comparison of self-report and clinical diagnostic interviews for depression: Diagnostic Interview Schedule and Schedules for Clinical Assessment in Neuropsychiatry in the Baltimore Epidemiologic Catchment Area follow-up. *Arch Gen Psychiatry* 2000;57:217-22.
9. Moffitt TE, Caspi A, Taylor A et al. How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. *Psychol Med* 2010;40:899-909.
10. Tyrer P. Are general practitioners really unable to diagnose depression? *Lancet* 2009;374:589-90.
11. Fink M, Taylor MA. Resurrecting melancholia. *Acta Psychiatr Scand* 2007;115(Suppl. 433):14-20.
12. Goldberg D. The heterogeneity of "major depression". *World Psychiatry* 2011;10:226-8.
13. Maj M. When does depression become a mental disorder? *Br J Psychiatry* 2011;199:85-6.
14. Maj M. Refining the diagnostic criteria for major depression on the basis of empirical evidence. *Acta Psychiatr Scand* 2011;123:317.
15. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-IV-TR, 4th ed. Washington: American Psychiatric Association, 2000.
16. Burcusa SL, Iacono WG. Risk for recurrence in depression. *Clin Psychol Rev* 2007;27:959-85.
17. Colman I, Naicker K, Zeng Y et al. Predictors of long-term prognosis of depression. *Can Med Assoc J* 2011;183:1969-76.
18. Coryell W, Endicott J, Keller MB. Predictors of relapse into major depressive disorder in a nonclinical population. *Am J Psychiatry* 1991;148:1353-8.
19. Judd LL, Akiskal HS, Maser JD et al. A prospective 12-year study of subsyndromal and syndromal depressive symptoms in unipolar major depressive disorders. *Arch Gen Psychiatry* 1998;55:694-700.
20. Monroe SM, Harkness KL. Recurrence in major depression: a conceptual analysis. *Psychol Rev* 2011;118:655-74.
21. Mueller TI, Leon AC, Keller MB et al. Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. *Am J Psychiatry* 1999;156:1000-6.
22. Zisook S, Shear K, Kendler KS. Validity of the bereavement exclusion criterion for the diagnosis of major depressive disorder. *World Psychiatry* 2007;6:102-7.
23. Mojtabai R. Bereavement-related depressive episodes: characteristics, 3-year course, and implications for DSM-5. *Arch Gen Psychiatry* 2011;68:920-8.
24. Greden JF. The burden of recurrent depression: causes, consequences, and future prospects. *J Clin Psychiatry* 2001;62(Suppl. 22):5-9.
25. Nierenberg AA, Petersen TJ, Alpert JE. Prevention of relapse and recurrence in depression: the role of long-term pharmacotherapy and psychotherapy. *J Clin Psychiatry* 2003;64(Suppl. 15):13-7.
26. Solomon DA, Keller MB, Leon AC et al. Multiple recurrences of major depressive disorder. *Am J Psychiatry* 2000;157:229-33.
27. Bockting CLH, ten Doesschate MC, Spijker J et al. Continuation and maintenance use of antidepressants in recurrent depression. *Psychother Psychosom* 2008;77:17-26.
28. Clarke GN, Rohde P, Lewinsohn PM. Cognitive-behavioral treatment of adolescent depression: efficacy of acute group treatment and booster sessions. *J Am Acad Child Adolesc Psychiatry* 1999;38:272-9.
29. Geddes JR, Carney SM, Davies C et al. Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. *Lancet* 2003;361:653-61.
30. Frank E, Prien RF, Jarrett RB et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder: remission, recovery, relapse, and recurrence. *Arch Gen Psychiatry* 1991;48:851-5.
31. Wakefield JC, Schmitz MF. Recurrence of depression after bereavement-related depression: evidence for the validity of DSM-IV bereavement exclusion from the Epidemiologic Catchment Area Study. *J Nerv Ment Dis* 2012;200:480-5.
32. Robins LN, Regier DA. *Psychiatric disorders in America*. New York: Free Press, 1991.
33. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 3rd ed. Washington: American Psychiatric Association, 1980.
34. Spitzer RL, Wakefield JC. DSM-IV diagnostic criterion for clinical significance: does it help solve the false positives problem? *Am J Psychiatry* 1999;156:1856-64.

35. Wakefield JC, Schmitz MF, Baer JC. Does the DSM-IV clinical significance criterion for major depression reduce false positives? Evidence from the NCS-R. *Am J Psychiatry* 2010;167:298-304.
36. 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;65:1400-5.
37. Barkow K, Maier W, Ustun B et al. Risk factors for depression at 12-month follow-up in adult primary health care patients with major depression: an international prospective study. *J Affect Disord* 2003;76:157-69.
38. Maj M. Bereavement related depression in the DSM-5 and ICD-11. *World Psychiatry* 2012;11:1-2.
39. Wakefield JC, First MB. Validity of the bereavement exclusion to major depression: does the empirical evidence support the proposal to eliminate the exclusion in DSM-5? *World Psychiatry* 2012;11:11-5.
40. Wakefield JC, Schmitz MF, First MB et al. Should the bereavement exclusion for major depression be extended to other losses? Evidence from the National Comorbidity Survey. *Arch Gen Psychiatry* 2007;64:433-40.
41. Chouinard G, Chouinard VA, Corruble E. Beyond DSM-IV bereavement exclusion criterion E for major depressive disorder. *Psychother Psychosom* 2011;80:4-9.
42. Parker G. Classifying clinical depression: an operational proposal. *Acta Psychiatr Scand* 2011;123:314-6.
43. Broadhead WE, Blazer DG, George LK et al. Depression, disability days, and days lost from work in a prospective epidemiologic survey. *JAMA* 1990;264:2524-8.
44. Elkin I, Shea T, Watkins JT et al. National Institute of Mental Health Treatment of Depression Collaborative Research Program: general effectiveness of treatments. *Arch Gen Psychiatry* 1989;46:971-82.
45. Shorter E. The doctrine of the two depressions in historical perspective. *Acta Psychiatr Scand* 2007;115:5-13.
46. Wakefield JC, Schmitz MF, Baer JC. Did narrowing the major depression bereavement exclusion from DSM-III-R to DSM-IV increase validity? Evidence from the National Comorbidity Survey. *J Nerv Ment Dis* 2011;199:66-73.
47. Wakefield JC, Schmitz MF, Baer JC. Relation between duration and severity in bereavement-related depression. *Acta Psychiatr Scand* 2011;124:487-94.
48. Stata Corporation. Stata statistical software, Release 12 (Version 12). Texas: Stata Corporation, 2011.
49. Kendler KS, Myers J, Zisook S. Does bereavement-related major depression differ from major depression associated with other stressful life events? *Am J Psychiatry* 2008;165:1449-55.
50. Zisook S, Kendler KS. Is bereavement-related depression different than nonbereavement-related depression? *Psychol Med* 2007;37:779-94.
51. Narrow WE, Rae DS, Robins LN et al. 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;59:115-23.
52. Beals J, Novins DK, Spicer P et al. Challenges in operationalizing the DSM-IV clinical significance criterion. *Arch Gen Psychiatry* 2004;61:1197-207.
53. Ustun B, Kennedy C. What is "functional impairment"? Disentangling disability from clinical significance. *World Psychiatry* 2009;8:82-5.
54. Wakefield JC. Disability and diagnosis: should role impairment be eliminated from DSM/ICD diagnostic criteria? *World Psychiatry* 2009;8:87-8.
55. Wakefield JC, Spitzer RL. Lowered estimates – but of what? *Arch Gen Psychiatry* 2002;59:129-30.

DOI 10.1002/wps.20015