



Published in final edited form as:

JAMA Psychiatry. 2013 June ; 70(6): 599–607. doi:10.1001/jamapsychiatry.2013.751.

Evidence for Multiple Genetic Factors Underlying *DSM-IV* Criteria for Major Depression

Dr. Kenneth S. Kendler, MD, Dr. Steven H. Aggen, PhD, and Dr. Michael C. Neale, PhD
Virginia Institute for Psychiatric and Behavioral Genetics (Drs Kendler, Aggen, and Neale) and
Departments of Psychiatry (Drs Kendler, Aggen, and Neale) and Human and Molecular Genetics
(Drs Kendler and Neale), Virginia Commonwealth University, Richmond

Abstract

Importance—We do not know whether the clinical criteria for major depression (MD) reflect a single or multiple dimensions of genetic risk.

Objective—To determine the structure of genetic and environmental risk factors for the 9 *DSM-IV* symptomatic MD criteria.

Design—Population-based twin registry.

Setting—Virginia.

Participants—Seven thousand five hundred members of adult twin pairs from the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders.

Main Outcome Measure—Symptoms of lifetime MD as assessed at personal interview.

Results—The best-fit twin model was multidimensional, requiring 3 genetic, 1 common environmental and 3 unique environmental common factors, and criterion-specific unique environmental factors. The first genetic factor was characterized by high loadings on cognitive and psychomotor depressive symptoms. The second and third genetic factors had strong loadings for mood and neurovegetative depressive symptoms, respectively. Genetic factor scores derived from these 3 factors differentially predicted patterns of comorbidity, other historical/ clinical features of MD, and demographic variables. These results suggested that the first genetic factor reflected a general liability to internalizing disorders, while the third genetic factor was more specific for melancholic MD. The 3 unique environmental common factors reflected, respectively, global depressive, core mood, and cognitive depressive symptoms.

Conclusions and Relevance—The *DSM-IV* syndrome of MD does not reflect a single dimension of genetic liability. Rather, these criteria reflect 3 underlying dimensions that index genetic risk for cognitive/psychomotor, mood, and neurovegetative symptoms. While in need of replication, these results, validated by predictions using estimated genetic factor scores, have implications for gene-finding efforts for MD.

The historical origins of the *DSM-IV* criteria for major depression (MD)¹ are traceable back through *DSM-III*² and Research Diagnostic and Feighner criteria^{3,4} to proposals made in the 1950s.^{5,6} Clinical judgment and not psychometrics guided the development of these criteria.

©2013 American Medical Association. All rights reserved.

Correspondence: Kenneth S. Kendler, MD, Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, PO Box 980126, Richmond, VA 23298 (kendler@vcu.edu).

Conflict of Interest Disclosures: None reported.

These criteria are diverse reflecting mood, hedonia, vegetative functions, psychomotor activity, and cognitive function and content. Perhaps surprisingly, factor analyses in both epidemiological^{7,8} and clinical⁹ studies suggest that these 9 criteria form a single dimension. However, such studies have been conducted solely at the phenotypic level.

Twin and family studies consistently show the etiological importance of genetic factors for the diagnosis of MD.^{10,11} However, we are unaware of any prior study applying multivariate genetic methods to determine whether the 9 *DSM* symptomatic criteria for MD reflect a single underlying genetic factor. We herein report such a study using personal interview-based assessments of lifetime MD in the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders.

METHODS

SAMPLE

Participants in this study derived from 2 interrelated Virginia Adult Twin Study of Psychiatric and Substance Use Disorders studies of white same-sex twin pairs.¹² Twins were ascertained from the birth certificate-based Virginia Twin Registry. Female-female (FF) twin pairs, born in 1934 to 1974, were eligible if both members responded to a mailed questionnaire in 1987 or 1988. Reports on symptoms of lifetime MD used in this report were collected at the third interview wave (FF3), conducted in 1992 to 1995, with cooperation in individual waves ranging from 88% to 93%.¹² Data on the male-male and male-female pairs (MMMF) came from a sample (birth years 1940–1974) initially ascertained, with a 72% cooperation rate, directly from registry records containing all twin births. The first interview (MMMF1) was completed largely by telephone in 1993 to 1996. The second wave of interviews (MMMF2), conducted in 1994 to 1998, had a response rate of 83%. Data on lifetime symptoms of MD were used from the MMMF2 wave.

Zygosity was determined by discriminate function analyses using standard twin questions validated against DNA genotyping in 496 pairs.¹³ The mean (SD) age and years of education of the twins were 35.1 (7.5) and 14.3 (2.2) at the FF4 interview and 37.0 (9.1) and 13.6 (2.6) at the MMMF2 interview.

These analyses used data from 7500 twins, including both members of 3084 pairs (503 monozygotic FF twin pairs, 346 dizygotic FF twin pairs, 703 monozygotic MM twin pairs, 485 dizygotic MM twin pairs, and 1047 opposite-sex dizygotic pairs) and 1325 twins without their co-twin. (These numbers do not sum because all possible pairings for triplet and quadruplet sets were included.)

The lifetime MD section of the interview was adapted from the *Structured Clinical Interview for DSM-III-R*¹⁴ modified to include *DSM-IV* criteria. Subjects had to respond positively to 1 or more lifetime episodes that included “feeling depressed or down” and/or “uninterested in things or unable to enjoy the things you used do” to continue in the section. Episodes that were judged to result from illness, medications, or normal grief were excluded. The worst lifetime episode was defined and individual questions were asked about the presence of each of the remaining *DSM-IV* MD symptomatic criteria for that time “when things were at their worst.”

The human subject committees at Virginia Commonwealth University approved this project. Written informed consent was obtained prior to face-to-face interviews and verbal consent, prior to telephone interviews. Interviewers had a master’s degree in a mental health-related field or a bachelor’s degree in this area plus 2 years of clinical experience. Members of a twin pair were always interviewed by different interviewers.

STATISTICAL METHODS

Twin models decompose the sources of individual differences in liability to MD into 3 components: additive genetic effects (A), shared family environment (C), and unique environment (E).¹² Shared environment reflects family and community experiences that increase similarity in siblings raised together. Unique environment includes random developmental effects, environmental experiences not shared by siblings, and random error.

Multivariate twin models estimate the degree to which genetic and environmental influences are shared across the 9 *DSM-IV* A criteria for MD¹ (which we term *common factors*) vs those specific to each individual criterion (termed *criterion specific*). This is done by including in the model genetic and environmental common factors that influence risk for all the criteria as well as criterion-specific influences.

Independent pathway structural equation twin models were fitted using the full information maximum likelihood method in Mx.¹⁵ We tested for both quantitative sex effects (ie, is the magnitude of genetic influences on MD the same in both sexes?) as well as qualitative sex effects (ie, are the same genetic factors influencing risk to MD in men vs women?). To obtain unbiased population-based parameter estimates, the model had to take into account the structure of the interview through which individuals were either selected into or skipped out of the MD section. That is, subjects who denied lifetime episodes of either sad mood (*DSM-IV* criterion A1 for MD) or loss of interest/ pleasure (criterion A2) could not meet diagnostic criteria for MD and were skipped out of the MD section. No data were recorded from them on the presence or absence of MD criteria A3 through A9.

The full information maximum likelihood method implemented in Mx makes use of all available twin information and can provide asymptotically unbiased parameters when selection items (here criteria A1 and A2) are included and missingness can be assumed to be “at random.”¹⁶ Mx optimizations were performed using both the try-hard option and different starting values to reduce the possibility that a solution found was a local rather than global minimum.

We began the sequence of model fitting with a basic 1-1-1 model with specifics (where the first, second, and third numbers reflect the number of genetic, shared environmental, and unique environmental common factors). Next, tests for quantitative and qualitative sex effects were carried out. Then, attempts were made to simplify the resulting model by deleting 1 by 1 all the common factors and then trying to delete the criterion-specific genetic and common environmental effects. (We did not attempt to eliminate the criteria-specific unique environmental effects as this unrealistically assumes that each criterion is assessed without error.) Next, we made the model progressively more complex systematically searching, at each step, for the “best-fit” model.

The primary goal was to find the model that reflected the optimal balance between parsimony and explanatory power. This goal was operationalized using the Akaike information criterion (AIC),^{17,18} which equals $2-2 df$ where df equals the difference in the df of the 2 models. We sought to minimize the AIC value. (We considered using the Bayesian information criterion, which has some desirable properties with complex models.¹⁹ However, probably because of the substantial missingness in our data, the Bayesian information criterion performed erratically and was dropped early in our model fitting.)

After determining a best-fitting model based on AIC, the Mx estimated factor loadings for this model were extracted and rotated in SAS²⁰ using a Varimax rotation criterion to improve interpretability. We considered as “prominent” loadings that were 0.31 or more (ie, accounted for 10% of the variance).

Maximum likelihood genetic factor scores were estimated by computing the conditional likelihood of the twin pairs' item responses, weighted by the joint likelihood of the factor score estimates. This is an application of Bayes' theorem, in that the joint likelihood $p(F,R)$ of the factor scores F and the item responses R is evaluated as $p(R|F)p(F)$. This factor score model was iteratively fitted, separately for each of the 5 different zygosity/sex groups, to each twin pair's raw data to estimate genetic factor scores for each individual. To validate the genetic factors found in our best-fit twin model, we predicted from these genetic factor scores a representative group of "validating" variables unrelated to the *DSM* MD criteria. These measures included 1 personality dimension known to be strongly associated with MD,²¹⁻²³ 2 representative internalizing and 1 representative externalizing psychiatric disorder known to be highly comorbid with MD²⁴; and 6 clinical and historical features of MD and 2 demographic variables (sex and years of education as an index of social class) known to be associated with MD.^{25,26}

To examine whether the genetic factor scores differed from each other in their prediction of the external validators, 2 regression analyses were performed. First, separate regressions were conducted to examine the pattern of differences in prediction for each validator. Second, a model constraining the 3 genetic regression coefficients to be equal within each validator variable was specified. Since the outcome variables were binary or ordinal (or rescaled to be ordinal), the robust weighted least squares mean and variance adjusted estimator in Mplus version 6.0²⁷ was used to optimize models. This model estimates probit regression coefficients for each of the genetic factor scores predicting the validator. Since the estimated genetic factor scores are calibrated on a uniform standard scale, the effect size units are more readily interpretable when comparing the coefficients.

RESULTS

SAMPLE

Of the sample of 7500 interviewed twins, 3829 (51.1%) denied lifetime episodes of sad mood and loss of interest/ pleasure, skipping out of the remainder of the MD section. Table 1 shows, in all subjects and in those screened into the MD section, the endorsement rates for the 9 criteria. In the entire sample, a period of 2 weeks or more of depressed mood was endorsed more frequently than a similar period of loss of interest or pleasure. Among the remaining 7 criteria reported for the time of the worse depressive episode, sleep problems was the most and suicidal ideation the least frequently endorsed.

MULTIVARIATE TWIN ANALYSIS OF THE SYMPTOMATIC CRITERIA

Model fitting started with basic model 1 (Table 2), which includes 1 genetic, 1 shared environmental, and 1 specific environmental common factor as well as criteria-specific genetic, shared environmental, and specific environmental effects. This is a 111_111 model where the first 3 numbers reflect the number, respectively, of the genetic, shared, and specific environmental common factors and the second set of numbers, the presence (designated by 1) or the absence (designated by 0) of criteria-specific genetic, shared environmental, and specific environmental factors.

Models 2, 3, and 4 modify the basic model, respectively, testing for quantitative sex effects, qualitative sex effects, and both quantitative and qualitative effects. Of these, only model 3 produced an improved fit index, suggesting the presence of qualitative sex effects.

Working from model 3, we simplified the model by dropping individual criteria-specific genetic (model 5), shared environmental (model 6), and both criteria-specific effects (model 7). All of these models fit better by AIC than model 3 but model 6 produced the best fit.

The next set of models (8, 9, and 10 in Table 2) eliminated, 1 at a time, the single genetic, shared environmental, and unique environmental common factors. None of these improved on the fit of model 6. Dropping the common unshared environmental factor (model 10) produced a large change in AIC relative to the other models, suggesting that unshared environmental factors are a prominent source of risk common to these symptomatic criteria for MD.

Additional models were then fit that added additional common factors to model 6. Models 11, 12, and 13 added, respectively, a second genetic, shared environmental, and unique environmental common factor. All of these models produced much better values of AIC than model 6, with the best result coming from model 11. Adding a second genetic factor improved the AIC by nearly 100 points, robust evidence for multiple genetic factors underlying the MD symptomatic criteria.

Trying to improve on these models, a fifth common factor was added to model 11. Models 14 to 16 explore the fit of a third genetic, a second shared environmental, or second unique environmental common factor. Of these, model 16 fit best. The substantial model improvement supports the existence of multiple unique environmental common factors for the depressive criteria.

Working from model 16, we then added, in models 17 to 19, a third genetic, second shared environmental, or third specific environmental common factor. Both models 17 and 19 produced an improved fit over that observed for model 16, but model 17 fit better. The improvement in fit was 12 AIC units.

Now, given model 17, models 20 to 22 added a fourth genetic, second shared environmental, or third specific environmental common factor. Only model 22 improved on the AIC value of model 17 but only modestly (0.5 AIC unit).

In models 23 to 25, models were further extended to include a fourth genetic, second shared environmental, or fourth specific environmental common factor. None of these models improved on the fit of model 22, making it our best-fit model.

BEST-FIT MODEL

Parameter estimates for the best-fit model 22 in Table 2 are depicted in the Figure for, respectively, the genetic and environmental common and criteria-specific factors.

The best-fit model included 3 orthogonal genetic factors that could be substantively interpreted. Factor 1 had prominent loadings on 4 A criteria: “psychomotor changes,” “feelings of worthlessness,” “difficulty concentrating,” and “suicidal ideation.” These criteria reflect the psychomotor/cognitive symptoms of MD. Symptom criteria with the highest loadings on factor 2 were the 2 stem criteria for MD, “sad mood” and “loss of interest/pleasure,” and the symptom of “worthlessness.” These reflect the core mood symptoms of MD. The third factor was defined primarily by neurovegetative-related symptoms, with prominent loadings on “sleep difficulties” and “fatigue.” This factor also had the highest loading for “changes in weight.” Only 1 symptom, fatigue, had a substantial genetic criterion-specific loading.

While significant in aggregate, shared environmental effects were modest for all the *DSM-IV*A criteria for MD with no particularly coherent pattern of loadings (Figure, B).

The 3 individual-specific environmental common factors were, by contrast, readily interpretable. The first factor had prominent loadings on 7 of the A criteria (all but “guilt” and “suicidal ideation”). This came closest to reflecting a general depressive symptom

factor. The second individual-specific environmental common factor had relatively strong loadings on both the screening items and “reduced concentration.” Like the second genetic factor, this unique environmental factor was defined predominantly by mood-related symptoms of MD. The third individual-specific environmental factor had robust factor loadings on 2 criteria that have traditionally reflected core cognitive symptoms of MD: “guilt” and “suicidal ideation.” The criterion-specific environmental factors (which include measurement error) varied from a low of +0.11 for anhedonia to a high of +0.74 for weight/appetite.

Table 3 presents the results of the best-fit model in a different but informative way. Noteworthy is the generally modest total heritability of most of the individual depressive symptoms, with estimates ranging from a low of 12% for weight/appetite changes to a high of 36% for fatigue. Collectively, for lifetime reports of the 9 *DSM-IV* MD criteria A symptoms, a large proportion of the risk liability was attributable to unique environmental sources.

VALIDATION OF GENETIC FACTORS

Based on the rotated results from the overall best-fit model, the level of genetic liability for each of the 3 identified factors was estimated as a factor score for each subject in our sample. These 3 genetic factor scores were then examined for their differential prediction of a representative set of external “validators.” As outlined in Table 4, the 3 factors differed significantly in the magnitude of their predictive effect sizes with 11 of the 12 variables. For 7 of the 12 analyses (neuroticism; lifetime history of generalized anxiety disorder, panic disorder, and alcoholism; age at onset; number of episodes; and treatment seeking), the effect was strongest with the first or cognitive/motor genetic factor. In these analyses, the second strongest loading varied between the second mood and the third neurovegetative genetic factor.

The neurovegetative factor was the strongest predictor for 2 variables: the melancholia subtype and 1 key “endogenous” symptom, early-morning awakening. The other key endogenous symptom (unreactive mood) was equally well predicted by the first and third genetic factors. The second or core mood genetic factor was the only factor that had a significant effect and differential prediction of sex and the most strongly associated with educational status. While the second genetic factor was associated with higher educational status, both the first and third genetic factors were associated with lower educational attainment.

COMMENT

The goal of this report was to investigate the dimensionality of the underlying structure of genetic and environmental risk factors for the *DSM-IV* symptomatic criteria for MD. We were particularly interested in determining whether genetic influences on these 9 criteria were homogeneous or heterogeneous. The sequence of model-fitting comparisons produced strong statistical evidence against the hypothesis of genetic homogeneity for the symptomatic criteria for MD. The best-fit model required 3 genetic factors. Based on the rotated profile of factor loadings, these 3 factors reflected the psychomotor/ cognitive, mood, and neurovegetative features of MD. We estimated maximum likelihood genetic factor scores from the parameters of this best-fit model and showed that the resulting scores were meaningfully differentiated in their patterns of predictions for selected validating criteria. The psychomotor/cognitive genetic factor was the strongest predictor of the personality trait of neuroticism and co-morbidity with anxiety disorders. This factor appears to reflect, at least in part, a nonspecific genetic liability to a broad range of internalizing disorders (as detected in multivariate analyses of 2 twin cohorts^{28,29}). By contrast, the third

genetic factor (neurovegetative) was most strongly predictive of the melancholic subtype and may reflect genetic risks more specific to MD.

Our model fitting produced evidence for qualitative sex effects on the genetic risk factors for depressive symptoms. These results are consistent with what has been seen in twin models for syndromal MD both in this sample³⁰ and in the large national Swedish twin sample.¹¹

Fatigue was the 1 MD criterion with substantial specific genetic effects. Consistent with prior studies,^{31,32} this suggests that considerable genetic variation in risk for fatigue exists in the population, much of which is independent of the vulnerability to depression.

We are unaware of any prior multivariate genetic study of *DSM* criteria for MD with which to compare our results. The most comparable prior study is that of Korszun et al,³³ who examined depressive symptom factors using items assessed by the Schedules for Clinical Assessment in Neuropsychiatry in 475 sibling pairs concordant for recurrent MD. They identified 4 factors, the first 3 of which (mood symptoms and psychomotor retardation, anxiety and psychomotor agitation, guilt, and suicidality) were significantly correlated in the sibling pairs. While the sample and assessment techniques and details of results differed from those of our study, both studies detected multiple, largely independent familial dimensions of depressive symptoms.

We also found that the latent common individual-specific environment was multidimensional and a prominent feature of the underlying risk liability structure for the MD symptom criteria. The 3 factors reflected general depressive, mood, and core cognitive symptoms. These results, which reflect a pathoplastic effect of the environment on depressive symptom patterns,³⁴ are consistent with more fine-grained prior work in this cohort demonstrating different depressive symptomatic profiles associated with different classes of stressful life events.³⁵ Environmental stressors for MD do not simply produce aggregate increase in risk for the full depressive syndrome but rather impact, at least in part, selected groups of depressive symptoms.³⁶

A rich clinical literature has long debated the unity vs heterogeneity of the depressive syndrome,^{37,38} several themes of which are echoed in our findings. First, Beck and Alford³⁹ suggested etiological primacy for the cognitive changes in depression. Of their cognitive triad, only 1—worthlessness—is in the *DSM-IV A* criteria. However, this symptom is closely associated with suicidality,^{40,41} so these 2 criteria might roughly index cognitive depressive symptoms. This is consistent with a recent factor analytic study of patients with MD using the Zung scale, which identified a “cognitive” factor that includes “suicidal rumination.”⁴² These 2 criteria—worthlessness and suicidal ideation—form 1 distinct environmental common factor and a key part of 1 genetic common factor.

Second, neurovegetative symptoms often cluster together in factor analyses of depressive symptoms.^{43–46} Trainees are typically taught that they reflect the “biological” or “hypothalamic” symptoms of depression. In our analyses, the 3 *DSMA* criteria assessing neurovegetative symptoms—weight/appetite and sleep changes and fatigue—had high loadings on the same genetic factor. Environmental influences on these symptoms were shared with other criteria in the first general symptom environmental common factor or were criteria specific in their effects.

Third, it has been debated whether sad mood or anhedonia better reflects the core symptoms of MD.⁴⁷ In our analyses, these 2 symptoms closely tracked each other in both the genetic and environmental factor structures. These criteria appear to reflect quite closely related genetic and environmental causes that were, in turn, partially distinct from those that impacted the more specific depressive symptoms.

In an earlier clinical examination of this cohort,⁴⁸ we examined individuals meeting *DSM-IV* criteria for MD in the last year and found that the symptomatic criteria for MD differed substantially in their predictive relationship with a range of clinical validators. Part of this heterogeneity was captured by the distinction between cognitive and neurovegetative symptoms, a distinction reflected in the current analysis in both our genetic and environmental factors. Our current results provide further support for our conclusion of this early report: “These results challenge the equivalence assumption for the symptomatic criteria for MD and suggest a more than expected degree of ‘covert’ heterogeneity among these criteria.”⁴⁸(p1679)

Results from these current analyses should be viewed in light of findings that *DSM* criteria for alcohol dependence,⁴⁹ antisocial personality disorder,⁵⁰ and conduct disorder⁵¹ reflect multiple diverse genetic factors. Many of the historical-clinical syndromes that populate our nosology may reflect multiple dimensions of genetic risk.

Compared with results obtained for schizophrenia and bipolar disorder, efforts to identify, by genome-wide association, molecular genetic variants impacting risk for MD have, to date, been less successful.^{52–54} There are many reasons why this may be so, including the higher population prevalence and lower heritability of MD.¹⁰ However, one possible contributing factor may be the genetic complexity of MD. If the disorder results from the joint effects of multiple independent dimensions of genetic risk, case-control designs may not be an optimal strategy. Greater power may arise from assessing and analyzing these dimensions independently rather than as aggregated together in the clinical syndrome of MD.

LIMITATIONS

These results need to be considered in the context of 6 potential methodological limitations. First, these results could be biased by methodological limitations of the twin method, particularly failure of the equal environment assumption or high levels of assortative mating. Prior analyses of these potential biases in this sample have suggested that their impact is likely to be modest at most.^{12,55}

Second, this is a general population sample so a considerable amount of information is coming from individuals with subsyndromal depressive symptoms. Furthermore, on average, those who meet *DSM-IV* criteria for MD are only moderately ill. It is not clear how our results would compare with a more severely ill, hospitalized sample of individuals with MD.

Third, for the data analyzed herein, reports of lifetime episodes of MD have been shown in this sample to be of only moderate reliability.^{56,57} While our results could have arisen from recall bias, it is difficult to construct a plausible pattern of biases that would produce the results observed.

Fourth, the model-fitting process was challenging. Although we used a simple algorithm for model selection and a commonly used and well-tested fit index, how confident can we be in the models election process?⁵⁸ The best response to this legitimate concern is to review the model-fitting results. Our conclusion that more than a single genetic factor is needed to account for the patterning of risk liability underlying the 9 MD criteria has robust statistical support: the large improvement in fit (99 AIC units) going from 1 to 2 genetic factors (models 6–11). We are also reasonably sure that 3 genetic factors were present because of the improvement increment in fit in going from models 16 to 17 (from 2 to 3 genetic factors; 12 AIC units). By similar logic, it is likely that MD has multiple individual-specific environmental factors (eg, the gain of 17 AIC units from model 16 vs 11). However, we are

much less sure whether the correct number of environmental factors is 2 or 3 because of the nearly identical fit of models 22 and 17. As to the specific parameter estimates, some skepticism is appropriate. The patterns of loadings in our best-fit model, especially for the multiple genetic and individual-specific environmental factors, are clinically sensible and consistent with prior literature. They were also relatively stable across different models. Only the results of replication will address these appropriate questions more definitively.

Fifth, it would have been ideal if we had information on the presence of each depressive criterion in all twins. However, in the absence of depressive episodes, it is difficult to meaningfully assess, over a lifetime, the occurrence of many of these symptoms such as problems with sleep or weight. Because we had responses to the 2 “stem” items (sad mood and loss of interest) on our entire sample and especially in twins discordant for endorsing these stem items, the conditional missing data patterns generated by the skip-out structure could be directly modeled.⁵⁹ Furthermore, we assessed in these twins the presence of all *DSM-IV* symptoms of depression experienced in the last year. We found, as predicted by our missingness model, that the intercorrelations of nonstem MD criteria were similar in subjects who did vs did not endorse one of the stems. While our model of missing data is unlikely to be precisely true, it is improbable that major biases were thereby introduced. Furthermore, this is a far preferable approach to restricting analysis to only subjects who met full MD criteria.³³

Finally, our analyses were restricted to the 9 symptomatic *DSM-IV* criteria for MD. Our results might have differed considerably if we had analyzed a larger and more diverse set of depressive symptoms.

CONCLUSIONS

These results have potentially important implications for the investigation of genetic risk factors for MD. Prior twin studies that have examined the magnitude of aggregate genetic effects for MD, their development over time, or the resulting patterns of comorbidity have assumed that MD reflected a unitary dimension of genetic liability. These results will need to be reconsidered in light of evidence for multiple genetic factors underlying MD. Molecular genetic studies, particularly candidate gene and genome-wide association studies, have also focused almost exclusively on the comparison of subjects meeting criteria for MD with matched controls. If accurate, our results suggest that this approach would likely be inefficient. While these results should be replicated before any widespread changes in analytic strategy are appropriate, they raise our awareness of the widely accepted but rarely tested assumption that the criteria for *DSM-IV* psychiatric and substance use disorders reflect single dimensions of genetic risk. This assumption is probably unwarranted and should be empirically evaluated rather than assumed.

Acknowledgments

Funding/Support: This research was supported in part by grants MH-40828 and MH/AA/DA-49492 from the National Institutes of Health. The Virginia Twin Registry is supported by grant UL 1RR031990.

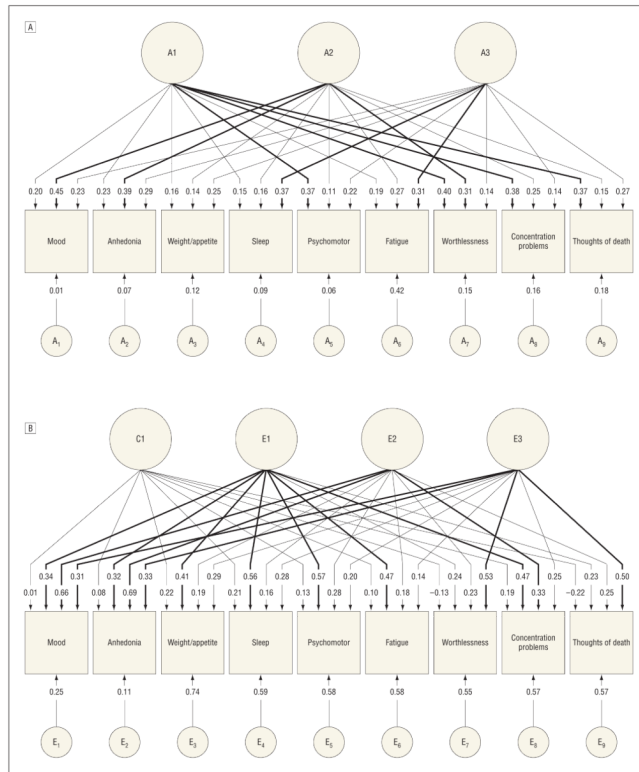
References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4. Washington, DC: American Psychiatric Association; 1994.
2. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 3. Washington, DC: American Psychiatric Association; 1980.
3. Spitzer, R.L.; Endicott, J.; Robins, E. Research Diagnostic Criteria for a Selected Group of Functional Disorders. 2. New York: New York Psychiatric Institute; 1975.

4. Feighner JP, Robins E, Guze SB, Woodruff RA Jr, Winokur G, Munoz R. Diagnostic criteria for use in psychiatric research. *Arch Gen Psychiatry*. 1972; 26 (1):57–63. [PubMed: 5009428]
5. Stone TT, Burris BC. Melancholia: clinical study of 50 selected cases. *J Am Med Assoc*. 1950; 142(3):165–168. [PubMed: 15399769]
6. Kendler KS, Muñoz RA, Murphy G. The development of the Feighner criteria: a historical perspective. *Am J Psychiatry*. 2010; 167(2):134–142. [PubMed: 20008944]
7. Muthen B. Dichotomous factor analysis of symptom data. *Sociol Methods Res*. 1989; 18(1):19–65.
8. Aggen SH, Neale MC, Kendler KS. *DSM* criteria for major depression: evaluating symptom patterns using latent-trait item response models. *Psychol Med*. 2005; 35(4):475–487. [PubMed: 15856718]
9. Simon GE, Von Korff M. Medical co-morbidity and validity of *DSM-IV* depression criteria. *Psychol Med*. 2006; 36(1):27–36. [PubMed: 16202189]
10. Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. *Am J Psychiatry*. 2000; 157(10):1552–1562. [PubMed: 11007705]
11. Kendler KS, Gatz M, Gardner CO, Pedersen NL. A Swedish national twin study of lifetime major depression. *Am J Psychiatry*. 2006; 163(1):109–114. [PubMed: 16390897]
12. Kendler, KS.; Prescott, CA. *Genes, Environment, and Psychopathology: Understanding the Causes of Psychiatric and Substance Use Disorders*. New York, NY: Guilford Press; 2006.
13. Kendler KS, Prescott CA. A population-based twin study of lifetime major depression in men and women. *Arch Gen Psychiatry*. 1999; 56(1):39–44. [PubMed: 9892254]
14. Spitzer, RL.; Williams, JB.; Gibbon, J. *Structured Clinical Interview for DSM-III-R—Patient Version (SCID-P, 4/1/87)*. New York: New York State Psychiatric Institute; 1987.
15. Neale, MC.; Boker, SM.; Xie, G.; Maes, HH. *Mx: Statistical Modeling*. 6. Richmond: Dept of Psychiatry, Virginia Commonwealth University Medical School; 2003.
16. Little, RJA.; Rubin, DB. *Statistical Analysis With Missing Data*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. New York, NY: John Wiley & Sons, Inc; 1987.
17. Akaike H. Factor analysis and AIC. *Psychometrika*. 1987; 52(3):317–332. doi:10.1007/BF02294359
18. Williams L, Holahan P. Parsimony-based fit indices for multiple-indicator models: do they work? *Struct Equ Modeling*. 1994; 1(2):161–189. doi:10.1080/10705519409539970.
19. Markon KE, Krueger RF. An empirical comparison of information-theoretic selection criteria for multivariate behavior genetic models. *Behav Genet*. 2004; 34(6):593–610. [PubMed: 15520516]
20. SAS [computer program]. Version 9.1.3. Cary, NC: SAS Institute Inc; 2005.
21. Hirschfeld RMA, Klerman GL, Lavori PW, Keller MB, Griffith P, Coryell W. Pre-morbid personality assessments of first onset of major depression. *Arch Gen Psychiatry*. 1989; 46(4):345–350. [PubMed: 2649038]
22. Klein, MH.; Kupfer, DJ.; Shea, MT. *Personality and Depression: A Current View*. New York, NY: The Guilford Press; 1993.
23. Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ. A longitudinal twin study of personality and major depression in women. *Arch Gen Psychiatry*. 1993; 50(11):853–862. [PubMed: 8215811]
24. Boyd JH, Burke JD Jr, Gruenberg E, et al. Exclusion criteria of *DSM-III*: a study of co-occurrence of hierarchy-free syndromes. *Arch Gen Psychiatry*. 1984; 41(10):983–989. [PubMed: 6477056]
25. Weissman MM, Bland R, Joyce PR, Newman S, Wells JE, Wittchen HU. Sex differences in rates of depression: cross-national perspectives. *J Affect Disord*. 1993; 29(2–3):77–84. [PubMed: 8300980]
26. Ritsher JE, Warner V, Johnson JG, Dohrenwend BP. Inter-generational longitudinal study of social class and depression: a test of social causation and social selection models. *Br J Psychiatry Suppl*. 2001; 40:s84–s90. [PubMed: 11315232]
27. Muthen, LK.; Muthen, BO. *Mplus User's Guide*: 1998-2010. 6. Los Angeles, CA: Muthen & Muthen; 2010.

28. Kendler KS, Prescott CA, Myers J, Neale MC. The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Arch Gen Psychiatry*. 2003; 60(9):929–937. [PubMed: 12963675]
29. Kendler KS, Aggen SH, Knudsen GP, Røysamb E, Neale MC, Reichborn-Kjennerud T. The structure of genetic and environmental risk factors for syndromal and subsyndromal common *DSM-IV* axis I and all axis II disorders. *Am J Psychiatry*. 2011; 168(1):29–39. [PubMed: 20952461]
30. Kendler KS, Gardner CO, Neale MC, Prescott CA. Genetic risk factors for major depression in men and women: similar or different heritabilities and same or partly distinct genes? *Psychol Med*. 2001; 31(4):605–616. [PubMed: 11352363]
31. Sullivan PF, Kovalenko P, York TP, Prescott CA, Kendler KS. Fatigue in a community sample of twins. *Psychol Med*. 2003; 33(2):263–281. [PubMed: 12622305]
32. Sullivan PF, Evengård B, Jacks A, Pedersen NL. Twin analyses of chronic fatigue in a Swedish national sample. *Psychol Med*. 2005; 35(9):1327–1336. [PubMed: 16168155]
33. Korszun A, Moskvina V, Brewster S, et al. Familiality of symptom dimensions in depression. *Arch Gen Psychiatry*. 2004; 61(5):468–474. [PubMed: 15123491]
34. Birnbaum, K. The making of a psychosis: the principles of structural analysis in psychiatry. In: Hirsch, SR.; Shepherd, M., editors. *Themes and Variations in European Psychiatry: An Anthology*. Charlottesville: University Press of Virginia; 1974. p. 199-238.
35. Keller MC, Neale MC, Kendler KS. Association of different adverse life events with distinct patterns of depressive symptoms. *Am J Psychiatry*. 2007; 164(10):1521–1529. quiz 1622. [PubMed: 17898343]
36. Cramer AO, Borsboom D, Aggen SH, Kendler KS. The pathoplasticity of dysphoric episodes: differential impact of stressful life events on the pattern of depressive symptom inter-correlations. *Psychol Med*. 2012; 42(5):957–965. [PubMed: 22093641]
37. Lewis A. States of depression. *Br Med J*. 1938; 2(4060):875–878. [PubMed: 20781840]
38. MacFadyen HW. The classification of depressive disorders, part I: I, a review of statistically based classification studies. *J Clin Psychol*. 1975; 31(3):380–394. [PubMed: 1100676]
39. Beck, AT.; Alford, BA. *Depression: Causes and Treatment*. 2. Philadelphia: University of Pennsylvania Press; 2008.
40. Williams JMG, Van der Does AJW, Barnhofer T, Crane C, Segal ZS. Cognitive reactivity, suicidal ideation and future fluency: preliminary investigation of a differential activation theory of hopelessness/suicidality. *Cognit Ther Res*. 2008; 32(1):83–104. 10.1007/s10608-006-9105-y
41. Lynch TR, Johnson CS, Mendelson T, et al. Correlates of suicidal ideation among an elderly depressed sample. *J Affect Disord*. 1999; 56(1):9–15. [PubMed: 10626775]
42. Romera I, Delgado-Cohen H, Perez T, Caballero L, Gilaberte I. Factor analysis of the Zung self-rating depression scale in a large sample of patients with major depressive disorder in primary care. *BMC Psychiatry*. 2008; 8:4. [PubMed: 18194524]
43. Hamilton MA. Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol*. 1967; 6(4):278–296. [PubMed: 6080235]
44. Bech P, Fava M, Trivedi MH, Wisniewski SR, Rush AJ. Factor structure and dimensionality of the two depression scales in STAR*D using level 1 datasets. *J Affect Disord*. 2011; 132(3):396–400. [PubMed: 21440308]
45. Uher R, Farmer A, Maier W, et al. Measuring depression: comparison and integration of three scales in the GENDEP study. *Psychol Med*. 2008; 38(2):289–300. [PubMed: 17922940]
46. Shafer AB. Meta-analysis of the factor structures of four depression questionnaires: Beck, CES-D, Hamilton, and Zung. *J Clin Psychol*. 2006; 62(1):123–146. [PubMed: 16287149]
47. Klein DF. Endogenomorphic depression: a conceptual and terminological revision. *Arch Gen Psychiatry*. 1974; 31(4):447–454. [PubMed: 4420562]
48. Lux V, Kendler KS. Deconstructing major depression: a validation study of the *DSM-IV* symptomatic criteria. *Psychol Med*. 2010; 40(10):1679–1690. [PubMed: 20059797]
49. Kendler KS, Aggen SH, Prescott CA, Crabbe J, Neale MC. Evidence for multiple genetic factors underlying the *DSM-IV* criteria for alcohol dependence. *Mol Psychiatry*. 2012; 17(12):1306–1315. [PubMed: 22105626]

50. Kendler KS, Aggen SH, Patrick CJ. A multivariate twin study of the *DSM-IV* criteria for antisocial personality disorder. *Biol Psychiatry*. 2012; 71(3):247–253. [PubMed: 21762879]
51. Kendler KS, Aggen SH, Patrick CJ. Familial influences on conduct disorder reflect 2 genetic factors and 1 shared environmental factor. *JAMA Psychiatry*. 2013; 70(1):78–86. [PubMed: 23117573]
52. Sullivan PF, de Geus EJ, Willemsen G, et al. Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. *Mol Psychiatry*. 2009; 14(4):359–375. [PubMed: 19065144]
53. Bosker FJ, Hartman CA, Nolte IM, et al. Poor replication of candidate genes for major depressive disorder using genome-wide association data. *Mol Psychiatry*. 2011; 16(5):516–532. [PubMed: 20351714]
54. Wray NR, Pergadia ML, Blackwood DH, et al. Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. *Mol Psychiatry*. 2012; 17(1): 36–48. [PubMed: 21042317]
55. Maes HH, Neale MC, Kendler KS, et al. Assortative mating for major psychiatric diagnoses in two population-based samples. *Psychol Med*. 1998; 28(6):1389–1401. [PubMed: 9854280]
56. Foley DL, Neale MC, Kendler KS. Reliability of a lifetime history of major depression: implications for heritability and comorbidity. *Psychol Med*. 1998; 28(4):857–870. [PubMed: 9723141]
57. 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 Psychiatry*. 1993; 50(11):863–870. [PubMed: 8215812]
58. Preacher KJ, Merkle EC. The problem of model selection uncertainty in structural equation modeling. *Psychol Methods*. 2012; 17(1):1–14. [PubMed: 22268762]
59. Neale MC, Aggen SH, Maes HH, Kubarych TS, Schmitt JE. Methodological issues in the assessment of substance use phenotypes. *Addict Behav*. 2006; 31(6):1010–1034. [PubMed: 16723188]

**Figure.**

Parameter estimates for genetic and environmental factors from the best-fit multivariate twin model applied to the 9 *DSM-IV* criteria for major depression. A, Parameter estimates for genetic factors from the best-fit multivariate twin model applied to the 9 *DSM-IV* criteria for major depression. Three genetic common factors were identified as well as genetic factors unique to each criterion. Factor loadings from the common factors of 0.31 or more are bold. A1 through A3 refer to the first through third common genetic factors. A1 through A9 refer to the criteria-specific genetic influences for criteria 1 through 9, respectively. Squares are observed variables and circles/ovals indicate latent variables. B, Parameter estimates for environmental factors from the best-fit multivariate twin model applied to the 9 *DSM-IV* criteria for major depression. One common and 3 individual-specific environmental common factors were identified as well as individual-specific environmental factors unique to each criterion. Factor loadings from the common factors of 0.31 or more are bold. C1 and E1 through E3 refer, respectively, to the single common shared environmental and the first through third common individual-specific environmental factors. E1 through E9 refer to the criteria-specific individual-specific environmental influences for criteria 1 through 9, respectively. Squares are observed variables and circles/ovals indicate latent variables.

Table 1Endorsement Frequencies for the Individual *DSM-IV* Criteria for Major Depression

<i>DSM-IV</i> A Criterion for Major Depression	Description	Endorsement, %	
		Entire Sample ^a	Sample Meeting Screening Criteria ^b
1	Depressed mood	47.2	96.4
2	Loss of interest/pleasure	39.1	79.8
3	Weight or appetite change	30.4	62.1
4	Insomnia or hypersomnia	35.1	71.8
5	Psychomotor changes	28.4	58.1
6	Fatigue	28.8	58.8
7	Worthlessness/guilt	31.2	63.7
8	Problems of concentration	31.7	64.8
9	Suicidal ideation	10.7	21.8

^aPercentages reflect the proportion of the entire sample who endorsed this criterion.

^bPercentages reflect the proportion of individuals who endorsed this criterion among those who endorsed depressed mood and/or loss of interest and hence were asked about the presence of all 9 criteria.

Table 2
Model Fitting Results for Independent Pathway Models Fitted to the 9 DSM-IV Major Depression Symptomatic (or A) Criteria

Model No.	Twin Model	A _c C _c E _c A _s C _s E _s ^a	Sex Effects, Quantitative/Qualitative ^b	2	df	AIC
1 ^c	111_111		-/-			
2	111_111		+/-	350.0	354	58.1
3	111_111		-/+	311.5	31	39.5
4	111_111		+/+	371.3	355	38.8
5	111_011		-/+	33.8	8	319.8
6	111_101		-/+	36.1	8	322.1
7	111_001		-/+	15.8	17	318.1
8	011_101		-/+	87.3	17	53.3
9	101_101		-/+	29.5	17	34.5
10	110_101		-/+	1672.2	17	1638.2
11	211_101		-/+	3121.5	0	3121.5
12	121_101		-/+	378.0	0	377.9
13	112_101		-/+	3107.3	0	3107.2
14	311_101		-/+	3133.9	37	3119.8
15	221_101		-/+	3144.9	38	3128.8
16	212_101		-/+	3154.2	38	3138.2
17	312_101		-/+	3180.2	315	3150.1
18	222_101		-/+	3165.1	316	3133.0
19	213_101		-/+	3171.0	315	3140.9
20	412_101 ^d		-/+	3172.7	315	3142.6
21	322_101		-/+	393.9	323	347.8
22 ^e	313_101		-/+	3194.7	322	3150.6
23	413_101 ^d		-/+	3168.4	322	3124.3
24	323_101		-/+	3183.7	324	3135.6
25	314_101 ^d		-/+	3177.2	322	3133.1

Abbreviations: A, additive genetic effects; AIC, Akaike information criterion; c, common, so AC reflects the number of genetic common factors; C, shared family environment; E, individual-specific environment; s, criterion specific, so Cs reflects whether criterion-specific shared environmental factors are included in the model.

^aFactors included in the model are indicated by a 1. Factors absent in the model are indicated by a 0. For example, a 212_101 model includes 2 common genetic, 1 common shared environmental, and 2 individual-specific common factors and criterion-specific genetic and individual-specific environmental factors.

^bTwo types of sex effects are modeled. In this column, $-/-$ refers to the absence in the model of both quantitative and qualitative effects, while $-/+$ means that the model contains qualitative but not quantitative effects.

^cModel 1 had a -2 log likelihood = 44 336.6, df = 40 443, and 72 estimated parameters.

^dModel fitted with a 4-4-4 common factor background structure. All other models were fitted with a 3-3-3 background structure. We did follow-up analyses to confirm that this change did not impact the fit statistics or the model parameters.

^eBest-fit model.

Table 3

Proportion of Variation of Each *DSM-IV* Criterion for Major Depression Due to Genetic, Shared Environmental, and Individual-Specific Environmental Factors in Our Best-Fit Multivariate Model and the Proportion of the Heritability Due to the 3 Common Genetic Factors

Criteria	%					
	Total a ²	F1	F2	F3	sp	Total e ²
Sad mood	29	14	69	17	0	0
Anhedonia	29	17	55	28	0	1
Weight/appetite	12	25	17	50	8	5
Sleep change	19	11	16	68	5	4
Psychomotor	20	70	5	25	0	2
Fatigue	36	8	19	25	47	1
Worthlessness	29	55	31	7	7	2
Concentration	26	57	23	8	12	4
Suicidal ideation	27	52	7	30	11	5

Abbreviations: a², additive genetic variance; e², shared environmental variance; c², unshared environmental variance; F1, psychomotor/cognitive genetic risk factor; F2, mood genetic risk factor; F3, neurovegetative genetic risk factor; sp, criterion-specific genetic.

Table 4
 Prediction of Patterns of Comorbidity and Features of Major Depression by the 3 Additive Genetic Factors Scores With Age and Sex as Covariates

Validator Variable	Effect Size of the Relationship Between the Validator and the Factors			Constrained Model Setting Effect Size of All Factors Equal ^e	
	Factor 1 Cognitive/Motor	Factor 2 Core Mood	Factor 3 Neurovegetative	χ^2	P Value
Neuroticism	0.54 ^b	0.27 ^b	0.30 ^b	17.8	<.001
Generalized anxiety disorder	0.77 ^b	0.49 ^b	0.47 ^b	9.3	.001
Panic disorder	0.67 ^b	0.21 ^b	0.48 ^b	5.7	.06
Alcohol dependence	0.62 ^b	0.29 ^b	0.32 ^b	11.7	.003
Major depression ^c					
Age at onset	-3.00 ^b	-1.15 ^d	0.37	33.4	<.001
No. of episodes	0.55 ^b	0.16 ^d	0.23 ^b	21.9	<.001
Treatment seeking ^c	0.63 ^b	0.30 ^e	0.48 ^b	9.1	.01
Melancholia subtype ^c	1.05 ^b	-0.43 ^b	1.34 ^b	173.8	<.001
Unreactive mood ^c	0.82 ^b	0.10	0.79 ^b	38.8	<.001
Early-morning awakening ^c	0.36 ^b	-0.54 ^b	1.39 ^b	241.7	<.001
Sex ^f	0.07	-0.23 ^b	-0.10	10.3	.006
Years of education	-0.49 ^b	1.17 ^b	-0.38 ^c	76.2	<.001

^aThe constrained model tests for the equality of the effect sizes across the 3 factors.

^b $P < .001$.

^cData available on close to 100% of cases ($n = 3761$) who endorsed at least 1 of the major depression screening criteria except for variables. These variables, for which information was available on about 45% to 48% of these cases, were assessed in the final part of the depression section of the interview, which required the endorsement of more than 2 depressive symptoms.

^d $P < .05$.

^e $P < .01$.

^f0, male, 1, female.