Clinical Characteristics of Major Depression That Predict Risk of Depression in Relatives

Kenneth S. Kendler, MD; Charles O. Gardner, PhD; Carol A. Prescott, PhD

Background: Major depression (MD) is both clinically and etiologically heterogeneous. We attempt to relate clinical and etiologic heterogeneity by determining those features of MD that reflect a high familial liability to depressive illness.

Methods: Our sample, 3786 personally interviewed twin pairs from a population-based registry, contained 1765 people with a lifetime history of MD by DSM-III-R criteria, of whom 639 (36.2%) had affected co-twins. We examine, using Cox proportional hazard models, the clinical features of MD in affected twins that predicted the risk for MD in the co-twin. Control variables were zygosity, age at interview, and sex of the twin and co-twin.

Results: The best-fitting model contained 4 significant predictors: number of episodes, duration of longest episode, recurrent thoughts of death or suicide, and level of distress or impairment. These 4 clinical features were similarly predictive of the risk for MD in the co-twins of male and female twins and predicted risk of illness more strongly in monozygotic than in dizygotic twins. Variables that did not uniquely predict risk of MD in the co-twin included age at onset and number of depressive symptoms. For number of episodes, the best-fitting model indicated an inverted U-shaped function with greatest co-twin risk for MD with 7 to 9 lifetime episodes.

Conclusions: The clinical features of MD in epidemiologic samples can be meaningfully related to the familial vulnerability to illness. Familial MD is best characterized by intermediate levels of recurrence, long duration of episodes, high levels of impairment, and recurrent thoughts of death or suicide. These clinical features probably reflect a high genetic liability to depressive illness.

Arch Gen Psychiatry. 1999;56:322-327

Major depression (MD) is both etiologically heterogeneous, with a wide range of risk factors, and clinically heterogeneous, presenting with varying patterns of symptoms, ages at onset, severity, duration, and recurrence. This article attempts to clarify further the relationship between etiologic and clinical heterogeneity in MD.

Because familial factors consistently influence risk for MD, investigators have identified a range of clinical features of MD related to the risk of illness in relatives, including age at onset, recurrence, impairment, and number or kind of depressive symptoms.

These studies have several potential limitations. First, all but 2 ascertained depressed patients in clinical settings, possibly confounding family history of MD and help seeking. Second, with few exceptions, previous studies examined only 1 or a few clinical features of MD, which is not the best way to detect the individual clinical variables that uniquely predict familial liability to depression. Third, many studies relied on family history assessments or only compared depressive probands with or without a positive family history for MD. Fourth, while some of the sample sizes of depressed probands were reasonably large (eg, 75 probands, 133 probands, or 177 probands), larger samples may be needed to resolve the effect of more modest predictors of familial liability. Fifth, all but 3 of the prior studies examined nuclear families and, therefore, could not disentangle the effect of genetic vs familial-environmental effects. Sixth, since the risk for MD differs substantially in men and women, clinical features of depressive illness that reflect high familial vulnerability may differ by sex. To our knowledge, this has not been examined previously.

In this article, we examine the relationship between a range of clinical fea-
SUBJECTS AND METHODS

SUBJECTS

The twins examined in this article are from 2 interrelated projects that studied white twins ascertained through the population-based Virginia Twin Registry, which was formed from a systematic review of all birth certificates in the Commonwealth of Virginia. The female-female twin pairs used in this study are from 2 related samples, together covering birth years 1934 to 1974. Twin pairs became eligible to participate if both members had previously responded to a mailed questionnaire, the response rate to which was approximately 64%. Eighty-eight percent of our sample was first interviewed face-to-face between 1987 and 1989 and has subsequently been the subject of 3 additional telephone interview waves. We examine results from the last of these waves, completed in 1997, a mean (±SD) of 92 (±7) months after their first assessment. The remainder of our female-female twin sample was first interviewed face-to-face between 1992 and 1994 and then interviewed by telephone 35 (±9) months later, between 1996 and 1997.

The male-male and male-female twin pairs, covering the birth years 1940 to 1974, were ascertained in a separate study beginning in 1993 in which we succeeded in interviewing 73% of those eligible.

From these 2 samples, we formed a total of 3786 twin pairs in which both members were personally interviewed. Zygosity was determined by standard questions, photographs, and when necessary in the female-female twins, by DNA testing. We validated our zygosity algorithm by analyzing 15 highly informative DNA polymorphisms in a random sample of 184 male-male twin pairs. The algorithm classified 177 pairs correctly (κ [±SD], 0.91 [±0.03]). All interviewers were blind to information about the co-twin.

MEASURES

Lifetime MD was assessed, with high interrater reliability (κ [±SD], 0.96 [±0.04]), by structured psychiatric interview based on the Structured Clinical Interview for DSM-III-R (SCID) with 2 modifications. First, we assessed last-year and lifetime history for MD prior to the last year in 2 separate sections. Second, SCID questions for MD were modified so that, from the “A criteria” for depression, we independently assessed the presence of the 14 disaggregated symptoms (eg, we separately recorded weight loss and weight gain). Twins with a history of depression were also asked about their age at onset of depression, the number of episodes of depression, length of the longest episode, degree of impairment during the worst episode, level of distress during the worst episode, and depression-related help seeking. Interviewers had at least a master’s degree in social work, psychology, or another mental health–related field, or a bachelor’s degree in 1 of these areas plus 2 years’ clinical experience.

STATISTICAL ANALYSIS

We focused on pairs: a proband twin with MD and the co-twin. Since the onset of MD is age dependent and our sample was variable in age, we used the Cox proportional hazard model, as operationalized in the PHREG procedure in SAS, to predict the hazard rate for MD in the co-twin as a function of clinical characteristics in the proband twin. The magnitude of impact of independent variables in the Cox model is assessed by a hazard ratio (HR) (or ratio of hazards) presented with the 95% confidence interval (CI), which reflects the increase in hazard rate associated with a unit change in the explanatory variable, adjusting for all covariates. To avoid an undue influence from the few twins who reported very large numbers of episodes or long durations, we truncated depressive episodes at 100 and duration at 3 years.

We tested the ability of the 14 disaggregated DSM-III-R symptoms to predict co-twin risk for MD. To avoid spurious positive results, we decided, a priori, to include in the main regression analyses only symptoms that predicted risk in the co-twin at P ≤ .01. Relevant values are expressed as mean ± SD.

RESULTS

SAMPLE

The sample comprised 1791 twins with a lifetime history of MD chosen from pairs where both members were interviewed. In 26 (1.5%) of these twins, 1 or more of the clinical variables of interest were missing and were therefore excluded, leaving a total of 1765 pairs of proband twins and their co-twins. Of the 1765 co-twins, 639 (36.2%) reported a lifetime history of MD. The proband sample had a mean age at onset and interview of 25.9 (±10.2) and 36.1 (±8.5), respectively. They endorsed 6.9 (±1.3) A criteria and the mean and median number of episodes they experienced were 6.1 (±13.7) and 3. The mean and median duration, in weeks, of the longest episode were 28.2 (±51.5) and 9. Because of the rightward skew in number of episodes and duration of longest episode, we log-transformed these variables for further analyses. Impairment and level of distress during the worst episode were assessed on 3- and 4-point scales, respectively (none, moderate, and severe; and not at all, somewhat, moderately, and very) and had mean scores of 2.0 (±0.7) and 2.9 (±0.7), respectively. Severe impairment was defined as “marked impairment in main life task so that respondent was almost nonfunctional, eg, could not go to job, do any housework.” Fifty-two percent of the affected twins reported seeking some kind of professional help for their illness.

CONTROL VARIABLES

We included 4 covariates in the Cox regression analyses, which all significantly predicted the co-twin hazard.
rate for MD: age at interview ($\chi^2_1 = 66.8; P < .001$; HR, 0.62; 95% CI, 0.55-0.70), sex of co-twin ($\chi^2_1 = 49.6; P < .001$; HR, 1.89; 95% CI, 1.58-2.26), zygosity ($\chi^2_1 = 29.2; P < .001$; HR, 0.64; 95% CI, 0.55-0.76), and sex of twin ($\chi^2_1 = 4.7; P = .03$; HR, 1.22; 95% CI, 1.02-1.45). The co-twin’s hazard rate for MD was substantially higher when the co-twin was female, younger, and belonged to a monozygotic (MZ) pair, and was slightly higher when the proband twin was female.

**SYMPTOMS**

We tested 14 disaggregated DSM-IV A criteria for MD separately, using an a priori type 1 error rate of 1%. Only 1 symptom met this criterion: recurrent thoughts of death or suicide ($\chi^2_1 = 14.5; P < .001$; HR, 1.38; 95% CI, 1.17-1.62). One other symptom, feelings of guilt, predicted risk in the co-twin at the 5% level ($\chi^2_1 = 4.5; P = .03$; HR, 1.21; 95% CI, 1.01-1.44), but was excluded from subsequent analyses.

**MODELS**

Since prior evidence pointed most strongly to age at onset and recurrence as predictors of risk of illness in relatives, we began with these variables. Model 1 included only age at onset, which significantly predicted the hazard rate in co-twins ($\chi^2_1 = 13.8; P = .002$; HR, 0.85; 95% CI, 0.78-0.93) (Table). Model 2 added number of episodes that also significantly predicted risk for MD in the co-twin. In model 3, we added (number of episodes)$^2$, which was significant and negative, indicating a lower risk for MD with very high and very low numbers of episodes. Adding (number of episodes)$^2$ to the model also strengthened the linear effect of number of episodes (HR from 1.18 to 1.98). In model 3, both number of episodes ($\chi^2_1 = 17.5; P < .001$; HR, 1.98; 95% CI, 1.43-2.73) and (number of episodes)$^2$ ($\chi^2_1 = 11.4; P = .007$; HR, 0.78, 95% CI, 0.67-0.90) were substantially stronger predictors of risk of MD in the co-twin than was age at onset ($\chi^2_1 = 6.3; P = .01$; HR, 0.89; 95% CI, 0.81-0.98). Model 4 added duration of longest episode, which also strongly predicted the hazard rate for MD in the co-twin ($\chi^2_1 = 13.5; P = .002$; HR, 1.23; 95% CI, 1.10-1.38). Adding (length of episode)$^2$, however, produced no improvement in model fit. Model 5 added recurrent thoughts of death/suicide, which was a unique predictor ($\chi^2_1 = 6.0; P = .01$; HR, 1.23; 95% CI, 1.04-1.47). In this model, however, age at onset no longer significantly predicted risk for MD even at a trend level ($\chi^2_1 = 2.2; P = .14$; HR, 0.93; 95% CI, 0.85-1.02) and was dropped from the model. The final significant improvement in the model came with the addition of degree of impairment during the worst episode ($\chi^2_1 = 3.9; P = .05$; HR, 1.13; 95% CI, 1.00-1.28). We attempted to then add to this model our remaining variables, but none were significant: age at onset ($\chi^2_1 = -2.0; P = .16$), level of distress ($\chi^2_1 = 3.1; P = .08$), number of A criteria ($\chi^2_1 = 0.1; P = .81$), and help seeking ($\chi^2_1 = 0.4; P = .54$).

We examined the HR for MD in the co-twin as a function of an increasing number of lifetime episodes (Figure). Although the HR was significantly greater in those with 2 or more vs only 1 lifetime episode of MD ($\chi^2_1 = 15.0; P < .001$; HR, 1.44; 95% CI, 1.20-1.73), the HR maximized in co-twins of twins with 7 to 9 lifetime episodes, being lower in those with fewer and more episodes.
the 10% level. However, of the 5 predictor variables, 1 demonstrated no interaction with zygosity (length of episode, $\chi^2 = 0.0, P = .89$), 1 demonstrated a nonsignificantly stronger effect in MZ vs dizygotic (DZ) co-twins (all 1-tailed $P$ values) (impairment, $\chi^2 = 1.7, P = .10$) and 3 were significantly stronger predictors of risk for MD in MZ than in DZ co-twins: number of episodes ($\chi^2 = 6.7, P = .005$), (number of episodes)$^2$ ($\chi^2 = 3.0, P = .04$), and recurrent thoughts of death ($\chi^2 = 11.0, P = .009$).

Although human recall is of limited reliability in general, and for depressive episodes in particular, we found 4 retrospectively reported clinical features of MD that significantly discriminated, among people with a lifetime history of depression, those who have a high vs low familial vulnerability to illness. We review these results herein.

**RECURRENT**

Five previous studies have suggested that recurrent episodes of MD are associated with higher familial vulnerability to illness. However, aside from our previous preliminary report in female-female pairs, these studies were restricted to clinical samples and divided patients into a few groups such as 1 vs 2 or more, 1, 2, or more than 3, or less than 2 vs more than 2 episodes. More information may be obtained by examining the entire range of numbers of episodes. While the risk of illness in the co-twin increased steadily as the number of reported lifetime episodes increased from 1 to 9, further increases were associated with a decline in risk to the co-twin. Given the importance of the number of episodes in these analyses, we examined its reliability in 192 randomly selected female twins who were interviewed twice, by different interviewers, with a mean interval of 4.3 ± 1.5 weeks. In the 62 twins who reported a lifetime history of MD on both occasions, the Spearman rank correlation for reported number of lifetime episodes was $+0.48 (P<.001)$, similar to or higher than that found previously.

With respect to the number of lifetime episodes, our results suggest that persons with MD can be divided into 3 groups: (1) those with few lifetime episodes where non-familial risk factors (eg, stressful life events) probably play a major role and familial factors play a minor causative role; (2) those with intermediate levels of recurrence with the highest level of familial risk; and (3) those who report very large numbers of episodes who have a low familial risk for depressive illness and may suffer from difficulties with affective modulation that are partially distinct from those that underlie MD.

**DURATION**

In the Maudsley twin sample, the MZ/DZ concordance ratio was substantially higher in persons with MD of less than the median episode length of 13 months. By contrast, we find increasing evidence for familial loading with increasing length of duration of the worst episode. Compared with people who experienced a 2-week episode, the HR for MD in co-twins was similarly increased in twins with a maximum duration of 15 to 30 days (HR = 1.48), 31 to 60 days (HR = 1.52), 61 to 90 days (HR = 1.51), 91 to 180 days (HR = 1.46), and 181 to 365 days (HR = 1.45), but more increased in those with episodes of 1 year (HR = 1.96). The Maudsley sample contained twins hospitalized for MD where the mean duration of illness was twice that found in our epidemiologic sample. Consistent with our findings, we observed in our female-female twins that genetic risk for MD was 1 independent predictor of duration of depressive episode.

**IMPAIRMENT**

The degree of impairment during the worst episode also significantly predicted risk of depression in the co-twin. We obtained a similar finding in an earlier interview wave with our female-female twins. The only other study of this issue that we are aware of found that a significantly higher proportion of depressed relatives of depressed probands than depressed relatives of control probands were incapacitated during their depression. Controlling for other factors, level of impairment in a depressive episode reflects the familial liability to illness.

**SYMPTOMS OF DEPRESSION**

We could not replicate previous results that risk of MD in relatives is correlated with the number of DSM-III-R criteria in depressed probands. Examining individual symptoms, 2 (recurrent thoughts of death or suicide and feelings of guilt) were significant predictors and only the former remained significant after correcting for multiple testing. We are unaware of a previous precedent for this finding. Both of these symptoms reflect derogatory self-evaluation, a core feature of the cognitive changes that may predispose to and accompany MD.

**AGE AT ONSET**

We could not replicate previous findings that early age at onset of MD uniquely predicts increased risk of MD.
in relatives.\textsuperscript{13-21,23} Two prior twin studies\textsuperscript{10,23} have examined this question and our results are consistent with one\textsuperscript{10} of them.

Four reasons for this discrepancy are possible. First, in addition to age at onset, our analyses included other variables such as recurrence and severity. In our own analyses, age at onset significantly predicted risk of illness in the co-twin prior to the inclusion of other predictor variables. Second, unlike most (but not all)\textsuperscript{35,18} studies, we included age at interview as a control variable. Numerous studies have shown higher hazard rates for MD in younger people, due to true cohort differences and/or differential recall.\textsuperscript{35,44} In any sample with a variable age structure, younger twins with MD will have earlier ages of onset. If not controlled for, a noncausal association between age at onset and risk in relatives could be induced. Third, only our study examined a non-treated sample. If both early onset and positive family history commonly produces help seeking (eg, a parent with prior depression insisting that a depressed teenager get treatment), the association in treated samples might be artifactual. Fourth, our sample is young, few persons having an onset of MD after age 40 years, a group that may have a particularly low familial loading.\textsuperscript{39} Our restricted age range may have limited our power to detect an association between age at onset and risk in relatives.

**SEX AND GENETICS**

Although differences in the lifetime prevalence of MD in men and women is consistently shown in epidemiologic studies of this condition,\textsuperscript{29,30} we found no significant differences in the clinical features of MD between the sexes that predicted risk of depression in the co-twin. High familial liability to illness seems to effect the clinical features of MD similarly in men and women.

Because our study contained MZ and DZ twins, who differ in their degree of genetic relatedness, we could assess whether the clinical features of MD that predicted rates of illness in the co-twin were likely to reflect genetic vs familial environmental risk factors. Of the 5 significant variables, 1 was equally predictive of risk in MZ and DZ co-twins, while the remaining 4 were more predictive in MZ co-twins, 1 nonsignificantly and 3 significantly. These results suggest that the identified significant predictors are largely indices of a high genetic liability to MD. This result is consistent with prior evidence in this and other twin samples that the familial aggregation of MD is largely or entirely due to genetic factors.\textsuperscript{10,31,33}

**RELIABILITY OF REPORTING**

Could the variables found to predict co-twin risk for MD reflect reliability of reporting rather than familial risk? In results from an earlier wave of interviews with the female-female twin pairs in this study, 3 clinical variables uniquely predicted stability of the diagnosis of MD: number of symptoms, treatment-seeking, and number of episodes.\textsuperscript{39} Other studies have found number of symptoms and treatment seeking to predict diagnostic stability in depression.\textsuperscript{38,60} Since some clinical variables predict only reliability of reporting but not risk of illness in the co-twin (eg, treatment-seeking and number of depressive symptoms) or risk of illness in the co-twin but not reliability of reporting (eg, duration and impairment), it is unlikely that predictors of familial risk are simply indices of reliability of reported depressive episodes.

**IMPLICATIONS**

In a large epidemiologic sample of twins suffering from lifetime MD, we identified several clinical indices of a high familial/genetic vulnerability to illness. Given the recall and measurement problems inherent in the assessment of lifetime psychopathologic disorders, it is likely that we have underestimated the relationship between these measures and the genetic risk of illness. Most importantly, these results suggest that MD is indeed an etiologically heterogeneous syndrome and that this heterogeneity can be reflected in clinical heterogeneity.

With rapidly increasing interest in linkage studies of MD, the appropriate definition of the highly “familial” or “genetic” forms takes on new urgency. The recent report from the National Institute of Mental Health Genetics Workgroup\textsuperscript{37} has declared that “early-onset depression” is the form of depressive illness that should be considered for “large-scale molecular approaches.” This conclusion is not consistent with the results of this investigation.

**LIMITATIONS**

We have obtained the lifetime history of mania only in the female-female pairs. When we removed twins with a prior history of mania or hypomania from the sample, virtually no change was seen in the results of the full model (model 6 in the Table). It is unlikely that our findings are substantially skewed by the small subset of affected people with bipolar illness.

The Cox models we employed assume the independence of the observations in those twin pairs concordant for MD, which constitute 22% of the analyzed pairs. In these pairs, we use clinical features of twin A to predict risk in twin B and clinical features of twin B to predict risk in twin A. For our final model we performed 1000 replicates of nonparametric bootstrap simulations and thereby obtained empirical SEs. These were 2% to 4% greater than those obtained by the standard Cox analyses, suggesting that the bias was introduced into our analyses because the twin structure of our data was slight.

Accepted for publication January 14, 1999.

This work was supported by National Institutes of Health grants MH/AA-49492 and AA-09095 and a Research Scientist Award 6(MH-01277) to Dr Kendler. The Virginia Twin Registry, established by Walter Nance, MD, PhD, and maintained by Linda Corey, PhD, is supported by the US National Institutes of Health grants HD-26746 and NS-31564. These data were collected under the direction of Sarah Burns, MA, Frank Butera, MS, Lisa Halberstadt, MS, and Barbara Brooke, MSW.

Reprints: Kenneth S. Kendler, MD, Box 980126, Richmond, VA 23298-0126.
REFERENCES