A Registry-Based Twin Study of Depression in Men

Michael J. Lyons, PhD; Seth A. Eisen, MD, MSC; Jack Goldberg, PhD; William True, PhD, MPH; Nong Lin, PhD; Joanne M. Meyer, PhD; Rosemary Toomey, PhD; Stephen V. Faraone, PhD; Mary Merla-Ramos, MA; Ming T. Tsuang, MD, PhD, DSc, FRCPsych

Background: The only large, registry-based twin study of depression using diagnostic criteria assessed by structured interview included only women. We present results from a comparable study of men.

Methods: Data were collected using a standardized telephone interview of men from the Vietnam Era Twin Registry. Both twins from 3372 pairs participated. Probands were randomly selected to analyze the data.

Results: Both twins from 3372 pairs participated. Probands and co-twins were unaware of the diagnostic status of the co-twin and the zygosity of the pair (MZ vs DZ). Monozygotic twins did not differ from DZ twins in their concordance rates for dysthymia. Other studies do not distinguish bipolar disorder from unipolar illness. Others contain methodological limitations relating to sample ascertainment. The optimal design of a twin study uses criteria-based diagnoses derived from structured diagnostic interviews in a sample ascertained to be representative of a population. To our knowledge, only the study by Kendler et al1 of women included an adequately sized sample and fulfilled these criteria. The one study4 that included men and used a population-based sample and structured interview was based on a small sample.

We report results from a study of the Vietnam Era Twin Registry. The advantages of the current study are that (1) assessment and diagnoses were carried out by researchers who were unaware of the diagnostic status of the co-twin and the zygosity of the pair and (2) twins were taken from a population-based registry and their conditions were diagnosed using accepted diagnostic criteria assessed through structured interview.

RESULTS

About 9.2% of twins met the DSM-III-R criteria for MD and 2.4% met the criteria for dysthymia at some time during their lives. Monozygotic twins were significantly more likely than DZ twins to be concordant for DSM-III-R MD, indicating a genetic effect (Table 1). Monozygotic twins did not differ from DZ twins in their concordance rates for dysthymia, which is not consistent with a genetic effect on dysthymia. When individuals with MD were subdivided into the DSM-III-R severity categories of mild, moderate, and severe psychotic, there was not a significant difference in MZ vs DZ concordance for mild and moderate severity levels. However, MZ twins had a significantly higher rate of concordance than DZ twins for severe/psychotic MD. There was a significantly greater MZ than DZ concordance rate for early, but not late, onset. Table 2 presents parameter estimates from the full model in the first row and then results from reduced models. The best-fitting

Depression is one of the most common serious psychiatric disorders.1 Numerous approaches have been brought to bear to investigate its etiology. Molecular genetic techniques have been applied,2 but sophisticated approaches have become available only recently, and results have been varied. Also, the probable multifactorial etiology of depression may impede rapid progress in determining the molecular genetic basis for risk. Therefore, traditional genetic epidemiological approaches remain important.

Although results of previous twin studies generally support the importance of genetic factors in the etiology of depression, some older studies do not distinguish bipolar disorder from unipolar illness. Others contain methodological limitations relating to sample ascertainment. The optimal design of a twin study uses criteria-based diagnoses derived from structured diagnostic interviews in a sample ascertained to be representative of a population. To our knowledge, only the study by Kendler et al1 of women included an adequately sized sample and fulfilled these criteria. The one study4 that included men and used a population-based sample and structured interview was based on a small sample.

We report results from a study of the Vietnam Era Twin Registry. The advantages of the current study are that (1) assessment and diagnoses were carried out by researchers who were unaware of the diagnostic status of the co-twin and the zygosity of the pair and (2) twins were taken from a population-based registry and their conditions were diagnosed using accepted diagnostic criteria assessed through structured interview.

©1998 American Medical Association. All rights reserved.
SAMPLE AND METHODS

SAMPLE

Participants were members of the Vietnam Era Twin Registry, which comprises male-male twin pairs born between 1939 and 1957 in which both members served in the military during the Vietnam War era (1965-1975). Zygosity was determined using a questionnaire and blood group typing methods that achieved 95% accuracy. Of 10,300 eligible individuals (5,150 pairs), 47 were deceased or incapacitated. Of those remaining, 8,169 (79.6%) were successfully interviewed by telephone. The 1874 monozygotic (MZ) and 1498 dizygotic (DZ) pairs in which both members responded to the depression items are the subject of this report (pairwise response rate, 66.1%). The mean age of respondents was 44.6 years (SD, ±2.8 years; range, 36-55 years); 90.4% were non-Hispanic white, 4.9% were African American, 2.7% were Hispanic, 1.3% were Native American, and 0.7% were other; 33.3% were high school graduates and 38.6% were college graduates; and 92.6% were employed full-time and 1.8% were employed part-time.

MEASURES

Subjects were interviewed using the Diagnostic Interview Schedule Version III Revised. Interviews were performed by telephone by the Institute for Survey Research at Temple University, Philadelphia, Pa, using experienced interviewers who recorded responses directly into a computer database. Interviewers were trained by one of the investigators (M.J.L.) and their supervisor, who had attended a training course conducted by the developers of the interview. Telephone interviews were used instead of face-to-face interviews because of the geographical diversity of the sample.

Diagnoses of depressive disorders were based on meeting the DSM-III-R diagnostic criteria for major depression (MD) or dysthymia. Based on DSM-III-R, individuals who have experienced a manic episode are excluded from these diagnoses. Men with MD were further classified using the DSM-III-R categories of mild (few, if any, symptoms more than the minimum to make the diagnosis and only minor impairment in usual social activities or occupational functioning or social activities or relationships), moderate (impairment and symptoms intermediate between mild and severe), and severe/psychotic (several symptoms more than the minimum required for diagnosis, marked interference with occupational or social functioning, and/or delusions or hallucinations). For the purposes of some analyses, men were divided into groups based on early and late onset of their first episode of MD. Age 30 years was selected as the cutoff point because this divided the men into groups of approximately equal size.

STATISTICAL ANALYSIS

Similarity within twin pairs is quantified through the calculation of probandwise concordance rates separately for MZ and DZ pairs. Concordance rates were also calculated for MD subdivided into the severity categories of mild, moderate, and severe/psychotic and for early vs late onset. Significance tests were carried out using the test of the difference between proportions. Significantly greater MZ than DZ concordance indicates a genetic effect.

In the analyses of the subcategories of depression (ie, severity level and age of onset), pairs were classified as concordant if both members of the pair had the same severity level or age of onset. Pairs were classified as discordant if one twin had the relevant severity level or age of onset and his co-twin had no lifetime history of MD or a different level of severity or age of onset.

Structural equation models that provide estimates of the proportion of phenotypic variance due to additive genetic effects (heritability or A), common or shared environmental effects (C), and unique or nonshared environmental effects (E) are fitted to the twin correlations. A computer analysis software package, LISREL, was used to estimate parameters by the method of weighted least squares, using the inverse of the asymptotic variances of the correlations as the weights. The first model is a full model that includes A, C, and E. The term E reflects effects that are specific to individuals rather than the pair and random error. The full model is compared with reduced models. Models without A test whether the twin correlation is due solely to common environmental effects and models without C test whether familial aggregation is due only to additive genetic effects. Models that include only E were not tested because all analyses demonstrated familial resemblance. To assess whether a reduced model fits the data worse than the full model, a χ² difference test is used. If this χ² is not significant, the reduced model is accepted as the more parsimonious explanation of the observed results. When the superior model cannot be identified by this series of likelihood tests, the Akaike information criterion is used. The Akaike information criterion equals χ² − 2 df and provides a reasonable balance between parsimony and goodness of fit.

“Nonadditive” genetic effects are tested statistically by fitting a model in which C is replaced with a nonadditive genetic parameter (D). In no case did the “best-fitting” model include a D term.

model for MD included a moderate effect from genetic factors and no significant effect from the family environment. The lack of an effect from the family environment is also suggested by the MZ tetrachoric correlation in Table 1, which is more than double the DZ correlation. The selection of the AE model as the best fit is supported by the significant χ² value for the CE model, which indicates that deleting additive genetic effects causes a statistically significant decrement in the fit of the model. The superiority of the AE model is also supported by the observation that this model has the lowest value for the Akaike information criterion.

The best-fitting model for dysthymia by the Akaike information criterion (AIC) was a model that included C but not A. Similar tetrachoric correlations for MZ and DZ twins for dysthymia also suggest that the family environment, rather than genetic factors, is responsible for the observed family resemblance. However, the CE model could not be distinguished from the AE model on the basis of the χ² goodness-of-fit test. Using the χ² goodness-of-fit test it was not possible to distinguish between shared genes vs shared environment as the source of the family resemblance for dysthymia.

Mild and moderate MD followed patterns that were similar to those for dysthymia. In both cases, the best-fitting model using the AIC was a model that included a significant effect from the family environment but not from additive genetic effects. Also in both cases, the χ² goodness-of-fit test could not distinguish between genetic and family environmental effects as the source of family resemblance.
The best-fitting model for early-onset MD was a model that included additive genetic effects but not the family environment. The superiority of this model was supported by the AIC and a $\chi^2$ goodness-of-fit test that rejected a model without A. The large difference between the MZ vs DZ tetrachoric correlations further supports the role of genes in early-onset depression. The best-fitting model for late-onset depression as indicated by the AIC was a model that included A but not C. The $\chi^2$ goodness-of-fit test could not distinguish between genes and the family environment as the source of family resemblance for late-onset depression.

Both early- and late-onset MD are affected by additive genetic factors and not the family environment, and the heritability of early-onset MD (0.47) is considerably greater than the heritability of late-onset MD (0.10). However, the 2 separate models that were fitted are not a direct test of whether the determinants of early- and late-onset MD are significantly different. To test this possibility, a heterogeneity test that constrained the 2 heritabilities to be equal was compared with the model that allowed the 2 heritability estimates to be different and was found to provide a significantly poorer fit ($\chi^2=4.57, df=1, P=.034$). Therefore, we conclude that early-onset MD is significantly more heritable than late-onset MD.
Our results demonstrate that familial effects, which may include inherited factors as well as aspects of the environment shared by the members of a twin pair, have a substantial effect on DSM-III-R–defined mood disorders. The comparison of MZ concordance rates to those for DZs demonstrated a genetic effect on overall MD, on severe/psychotic MD, and on early-onset MD. The probability of MD is affected significantly by genetic factors and experiences that are unique to each sibling but not by environmental experiences shared by siblings. The nonshared environment includes aspects of parental behavior that are not administered equally to twins, peers that are not shared, and many life experiences outside the family of origin.

Dysthymia, in contrast to MD, does not seem to be significantly affected by genetic factors. This is consistent with 212,13 of the 312-14 previous studies of less severe depressive disorders. Moreover, our results suggest that for dysthymia, unlike MD, there may be a meaningful contribution to etiology from the shared or family environment.

Our results suggest that the most severe level of depression may be more strongly affected by genetic factors. However, Kendler et al13 described a plausible alternative interpretation. Their work followed from previous observations that reports of psychiatric disorders on a lifetime basis are not highly reliable.16-18 Rice et al19 also suggested that severity may be associated with reliability of recall. More severe depression is likely to be retrospectively reported more reliably. Kendler et al13 determined that MD diagnosed using DSM-III-R criteria was moderately heritable. However, without an explicit inclusion of unreliability in the analyses, genetic effects may be underestimated and unique environmental effects may be overestimated because dissimilarities within pairs are attributed to unique environmental effects. When Kendler et al13 corrected for unreliability of lifetime reports of MD, heritability estimates increased from 0.42 to 0.70, a level comparable to schizophrenia and bipolar disorder. They concluded that MD is a highly heritable disorder diagnosed with moderate reliability rather than a moderately heritable disorder diagnosed with high reliability.

In their registry-based study of depression in women, Kendler et al3 reported that additive genetic effects contributed 42% of the variance in DSM-III-R MD and that the common environment did not contribute a detectable effect, which is similar to our values of 36% and 0%, respectively. Given the substantial sex difference in prevalence, it is interesting that genetic and environmental factors are important to the same degree in men and women. Our research does not address the question of the degree to which the same genetic and environmental factors affect depression in men and women. In a clinical sample of twins with MD, Torgersen20 reported a “maximum estimate” of heritability of 0.54 and a contribution of 0.03 of the variance from the shared environment. McGuffin et al21 reported a heritability of 0.48 and no significant shared environmental effect on MD in a clinical sample. Thus, there does not seem to be a difference between heritabilities derived from clinical samples vs those that we observed in our population-based sample.

Our results suggest that more severe subtypes of depression may be more strongly affected by genetic factors and that lower levels of depression may be more responsive to environmental factors—those within the family of origin and those outside it. This is consistent with results of previously published research. A family study of MD22 found that relatives of probands with more severe subtypes of depression had higher rates of depression than relatives of probands with less severe forms. A study of the Maudsley Hospital Twin Register21 found a significantly higher concordance rate for MZ pairs vs DZ pairs among twins with endogenous depression but not the International Classification of Diseases, Ninth Revision neurotic subtype. In the Maudsley twins, there was no evidence for earlier age of onset being associated with greater familiarity of affective disorder, but Weissman et al23 and Kuper et al24 found earlier age of onset in probands to be associated with greater morbidity among relatives, with greatest risk among relatives of probands with onset before age 20 years. There is other evidence,25 however, that prepubertal and adolescent onset may not be associated with greater risk of depression to relatives than the risk to relatives of probands with adult onset.

Our findings indicate that genetic factors play a greater role in MD with onset before age 30 years than in cases with later onset. We believe that the difference in heritability between early- and late-onset MD cannot be attributed solely to unreliability of reports related to the recency of the episode. Kendler et al13 did not find an association between age of onset and reliability. If anything, one might expect earlier, more temporally remote episodes to be reported less reliably, which would produce a pattern opposite the one we observed.

Face-to-face interviewing, the most common technique, was not feasible in the present study because of the large and geographically diverse sample. However, several studies3,16-27 support the comparability of telephone and face-to-face interviews. The classic twin study assumes that MZ and DZ twins do not differ with regard to the similarity of their exposure to relevant environmental effects (the “equal environments assumption”). Although our analyses did not include an assessment of the equal environments assumption, in a previous twin study of depression³ neither the similarity of childhood environment nor the frequency of contact during adulthood was related to twin similarity for depression. A study using hospital- and population-based samples in Sweden28 found no association between measures of childhood and adult environmental similarity and twin similarity for affective illness. Another approach for testing of the equal environments assumption is to compare twin resemblance based on the twins’ belief about their zygodity in contrast to their actual zygodity. In an application of this approach to the study of depression,29 there was no evidence for an effect of perceived zygodity on twin resemblance for MD, further supporting the validity of the equal environments assumption.

The computation of tetrachoric correlations requires the assumption that the dichotomous phenotype reflects a threshold imposed on an underlying, unobserved, normally distributed dimension of liability. Although we cannot directly observe this hypothesized continuum of vulnerability to depression, we believe that such a model is likely to be correct, which supports the use of the tetrachoric correlation. Moreover, tests that use concordance rates (Table 1) are nonparametric and do not require any distributional assumptions. Our agnostic approach to data analysis based on concordance rates supports the biometric modeling results.
There is a question about the generalizability of our findings to nonveterans. The lifetime prevalence of DSM-III-R MD in our sample was 9.2% and the lifetime prevalence of dysthymia was 2.4%. These figures are comparable to the prevalence of 12.7% and 4.8%, respectively, reported in the National Comorbidity Survey. By contrast, the Epidemiologic Catchment Area Study reported a lifetime risk for depression for men of 2.6% and for dysthymia of 2.2%. The National Vietnam Veterans Readjustment Study found a lifetime prevalence of depressive episodes of 5.1% for men who had served in the Vietnam theater, 4.0% for Vietnam-era veterans who did not, and 1.5% for matched civilian control subjects. The differences observed among participants in these studies are not readily explained. Vietnam-era veterans are drawn from a specific generation (those born between 1939 and 1957), which may be especially relevant given the evidence for a cohort effect in depression.

Accepted for publication July 31, 1997.

From the Harvard Institute of Psychiatric Epidemiology and Genetics (Drs Lyons, Toomey, Faraone, and Tsuang); Harvard Medical School Department of Psychiatry at the Massachusetts Mental Health Center (Drs Lyons, Toomey, Faraone, and Tsuang); the Department of Psychology, Boston University (Drs Lyons and Toomey); Research and Medical Services, St Louis Veterans Affairs Medical Center, St Louis, Mo (Drs Eisen, True, and Lin); the Department of Internal Medicine, Division of General Medical Sciences, Washington University (Dr Eisen); the School of Public Health, University of Illinois–Chicago (Dr Goldberg); the Veterans Administration Cooperative Studies Program Coordinating Center, Hines Veterans Administration Hospital, Hines, Ill (Drs Goldberg and Meyer); the School of Public Health, St Louis University Medical Center (Dr True); the Department of Human Genetics, Medical College of Virginia, Richmond (Dr Meyer); the Department of Psychology, Long Island University, Long Island, NY (Ms Merla-Ramos); and the Department of Epidemiology, Harvard School of Public Health (Dr Tsuang).

Supported by grants AA10586 (Dr Lyons) and AA10339 (Dr True) from the National Institute on Alcohol Abuse and Alcoholism, grant DAO4604 from the National Institute on Drug Abuse (Dr Tsuang); and the Department of Veterans Affairs Health Services Research and Development Service (study 992), Washington, DC.

Important contributions to the success of this study were made by the following: Department of Veterans Affairs Health Services and Development Service: Chief Research and Development Officer, John R. Feussner, MD; Former Director, Daniel Deykin, MD; Administrative Officer, Janet Gold; Cooperative Studies in Health Services: Program Manager, Charles Welch III, PhD; Health Services Research & Development: Deputy Director, Shirley Meehan, MBA, PhD; Hines Veterans Administration Cooperative Studies Program Coordinating Center, Vietnam Era Twin Registry: Director, William G. Henderson, PhD; Registry Coordinator, Mary Ellen Vitek; Programmer, Kenneth Bukowski; Statistical Assistant, Mary Biondich; Vietnam Era Twin Registry Advisory Committee: Theodore Colton, ScD, Ralph Paffenbarger, MD, Walter Nance, MD, Myrna Weissman, PhD, and Roger Williams, MD. Irving Gottesman, PhD, and Jag Khalsa, PhD, also made important contributions. Most important, we thank the members of the Vietnam Era Twin Registry for their participation and cooperation. They willingly provided sensitive information and considerable time in responding to the survey. Without their contribution this research project would not have been possible.

Reprints: Michael J. Lyons, PhD, Psychology Department, Boston University, 64 Cummings St, Boston, MA 02115.

REFERENCES


