Psychological Medicine, 2007, **37**, 1163–1172. © 2007 Cambridge University Press
doi:10.1017/S0033291707000244 First published online 4 April 2007 Printed in the United Kingdom

A longitudinal study of personality and major depression in a population-based sample of male twins

AYMAN H. FANOUS[1,2,3]*, MICHAEL C. NEALE[3], STEVEN H. AGGEN[3] AND KENNETH S. KENDLER[3]

[1] *Washington VA Medical Center, Washington, DC, USA*; [2] *Georgetown University Medical Center, Washington, DC, USA*; [3] *Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA*

ABSTRACT

Background. The relationship between personality and psychiatric illness is complex. It is not clear whether one directly causes the other.

Method. In a population-based sample of male twins ($n=3030$), we attempted to predict major depression (MD) from neuroticism (N) and extraversion (E) and vice versa, to evaluate the causal, scar, state, and prodromal hypotheses. In a longitudinal, structural equation twin model, we decomposed the covariation between N and MD into (*a*) genetic and environmental factors that are common to both traits, as well as specific to each one and (*b*) direct causal effects of N at time 1 on subsequent MD, as well as between MD and subsequent N.

Results. E was negatively correlated with lifetime and one-year prevalence of MD. N predicted the new onset of MD, and was predicted by both current and past MD. It did not predict the time to onset of MD. All of the covariation between N and MD was due to additive genetic and individual-specific environmental factors shared by both traits and a direct causal path between MD and N assessed later. No genetic factors were unique to either trait.

Conclusions. In men, N may be a vulnerability factor for MD but does not cause it directly. However, MD may have a direct causal effect on N. The genetic overlap between N and MD in men may be greater than in women.

INTRODUCTION

The notion that certain personality configurations are associated with more severe psychopathology dates back to Hippocrates (Jackson, 1986). In the twentieth century, Kraepelin (1921), Schneider (1958), and Kretschmer (1936) described personality types that were predisposed to depression, mania, and psychosis. More recently, the etiological relationship between personality and psychiatric illness has been of nosological as well as genetic significance. This may be due in part to the ascendancy of a spectrum concept in the affective (Akiskal & Pinto, 1999) and psychotic disorders (Siever *et al.* 1993; Kendler *et al.* 1995), according to which some Axis I disorders have personality-based substrates. In addition, the use of endophenotypes in genetic studies of psychiatric illness may provide greater power to detect susceptibility genes (Gottesman & Gould, 2003). Personality traits as endophenotypes improved the power to detect linkage to alcoholism in one study (Czerwinski *et al.* 1999), while the inclusion of schizophrenia-related personality disorders in the definition of affection

* Address for correspondence: Ayman H. Fanous, M.D., Washington VA Medical Center, 50 Irving St NW, Washington, DC 20422, USA.
(Email: ayman.fanous@med.va.gov)

in some linkage studies of schizophrenia have resulted in higher LOD scores (Riley & McGuffin, 2000).

Eysenck proffered a parsimonious theory of personality comprising the factors neuroticism (N), extraversion (E), and psychoticism (Eysenck & Eysenck, 1985). Rating scales of N- and E-like traits have subsequently been developed within the frameworks of other theories of personality. N, thought to represent emotionality or the predisposition to experience negative affect, has been linked to MD in clinical, epidemiological, family, and twin studies. It may predict the later onset of MD in never-ill individuals (Nystrom & Lindegard, 1975; Boyce *et al.* 1991; Kendler *et al.* 1993), suggesting a causal relationship. Additionally, N may be increased in individuals with MD even after recovery, indicating a 'scar' effect (Kendler *et al.* 1993), although two clinical studies could not confirm this (Zeiss & Lewinsohn, 1988; Duggan *et al.* 1991).

The state of being depressed, however, is itself associated with higher levels of reported N (Coppen, 1965; Kerr *et al.* 1970; Hirschfeld *et al.* 1983*a*; Farmer *et al.* 2002) (henceforth called the 'state' effect), which may be confounded with a true trait effect of N on MD in cross-sectional studies. Furthermore, personality pathology may be associated with a worse prognosis (Weissman *et al.* 1978; Boyce & Parker, 1985; Duggan *et al.* 1990). Therefore, clinical samples may be enriched with patients having high N, introducing bias. These considerations underscore the importance of epidemiological samples and longitudinal designs.

E measures sociability, liveliness, and the level of ease and pleasure felt in the company of others (Eysenck & Eysenck, 1985). Its relationship to MD may be in the opposite direction to that seen with N, but is probably less robust (Akiskal *et al.* 1983; Hirschfeld & Cross, 1983). Several studies support a state effect (Coppen, 1965; Kerr *et al.* 1970; Hirschfeld *et al.* 1983*a*; Farmer *et al.* 2002), whereby subjects in a depressive episode experience lower levels of E, but two others fail to support it (Bartholomew, 1959; Mezey *et al.* 1963). In two studies, low E did not predict future MD (Hirschfeld *et al.* 1989; Boyce *et al.* 1991). The confounding of state and trait effects and the selection bias of clinical samples mentioned above for N also apply to E.

The foregoing studies were designed to test hypotheses about the prediction of personality from MD and *vice versa*. However, since association alone does not imply causation, these studies cannot address hypotheses about the causes of covariation. Genetic-epidemiologic designs can do this, and are not susceptible to confounds and biases inherent in within-person designs (Neale & Cardon, 1992; Neale & Eaves, 1993). Five studies do (Wetzel *et al.* 1980; Krieg *et al.* 1990; Maier *et al.* 1992, 1995; Lauer *et al.* 1997), while four do not indicate that N and MD share common familial factors (Hirschfeld *et al.* 1983*b*; Katz & McGuffin, 1987; Ouimette *et al.* 1992; Farmer *et al.* 2002). Two studies support the same hypothesis for low E (Wetzel *et al.* 1980; Hirschfeld *et al.* 1983*b*), while six do not (Katz & McGuffin, 1987; Krieg *et al.* 1990; Maier *et al.* 1992; Ouimette *et al.* 1992; Kendler *et al.* 1993; Farmer *et al.* 2002).

In a previously published longitudinal twin study in women of personality and MD (Kendler *et al.* 1993), we reported evidence supporting the causal, scar, and state effects of N using regression analyses. In a longitudinal, structural equation twin model, most of the covariation between N and MD was due to common genetic factors and a small scar effect. We found no evidence of a significant relationship between E and MD.

There are several reasons to hypothesize that the causal relationship between personality and MD may be different in men than in women. First, the prevalence of MD (Regier *et al.* 1988; Weissman *et al.* 1993; Kessler *et al.* 1994) and the mean levels of N are higher, while mean E is lower (Floderus-Myrhed *et al.* 1980; Tambs *et al.* 1991; Macaskill *et al.* 1994; Jang *et al.* 1996), in women than in men. Second, the rates of substance abuse and antisocial personality disorders, both of which have been linked to dimensions of personality (Cox, 1985; Sher & Trull, 1994), are higher in men (Regier *et al.* 1988; Kessler *et al.* 1994). Third, in this sample, there is evidence of sex-specific genetic factors for both N and MD (Kendler & Prescott, 1999; Kendler *et al.* 2001; Fanous *et al.* 2002), as well as sex differences in the extent to which the correlation between N and MD arises from genetic factors common to both traits (Fanous

et al. 2002). However, when we previously examined sex differences in the relationship between N and MD, we used lifetime measures of MD and did not model the direct causal effect of these traits on each other. In this report, we combine longitudinal and genetic-epidemiologic designs to clarify the relationship between N and MD in men, similar to those previously used in a female sample (Kendler *et al*. 1993), with the goal of comparing the genetic *and* causal relationships between personality and MD across the sexes.

METHODS

Subjects

This report is based on data from the first and second waves of interviews in our study of adult male twins from the Virginia Twin Registry (now part of the Mid-Atlantic Twin Registry), details of which have been outlined previously (Kendler *et al*. 2000). Briefly, twins were eligible for participation if one or both members were successfully matched to state Division of Motor Vehicles records, and if they were white, a member of a multiple birth that included at least one male, and born between 1940 and 1974. After a full explanation of the research protocol, signed informed consent was obtained before all face-to-face interviews and verbal assent before all telephone interviews. Subjects were aged 20 to 58 years (mean 36·8, S.D. 9·1 years).

Wave One Assessment

Of 9417 eligible individuals for the first wave, 6812 (72·4%) completed initial interviews. At these interviews, which were conducted by telephone in over 95% of subjects, MD, N, and E were assessed for the first time. To assess test-retest reliability, 150 members of male-male twins were re-interviewed a mean of 4·4 (S.D. 1·1) weeks after their initial interview.

Wave Two Assessment

Those who completed the Wave One Assessment were recontacted by telephone or mail to schedule a face-to-face interview at least one year later. This second-wave interview (1994 through 1998), in which MD was assessed again, was performed face to face in 79·8% of the sample, and by telephone in the remainder. N and E were assessed for the second time by a self-report questionnaire mailed out after the Wave Two interview.

Our analyses are based on 1517 male-male pairs [867 monozygotic (MZ) and 650 dizygotic (DZ)]. Interviewers had a master's degree in a mental-health-related field or a bachelor's degree in this area plus two years of clinical experience. Members of a twin pair were interviewed by different interviewers who were unaware of clinical information about the co-twin. Zygosity determination was made using a discriminant function analysis based on six standard zygosity questions. The algorithm was developed on 227 twin pairs who underwent genotyping using eight or more highly polymorphic DNA markers (Kendler *et al*. 2000).

Measures

N and E were assessed using the 12 and 8 items, respectively, of the short form of the Eysenck Personality Questionnaire (Eysenck & Eysenck, 1975). The stabilities of N and E over 19·3 months in this sample were +0·69 and +0·73, respectively. We will refer to N and E assessed at time 1 and time 2 as N_1, E_1, N_2, and E_2, respectively. Lifetime and last-year prevalence of MD was assessed by a structured psychiatric interview based on the SCID (Spitzer *et al*. 1987). Our assessment differed from the standard SCID in two major ways. First, we assessed last-year history of MD and a lifetime history of MD prior to the last year in two separate sections. An individual who was positive for one or more episodes of MD in either of these sections is here called positive for lifetime MD. Second, SCID questions for MD were modified so that, from the 'A criteria' for depression, we independently noted the presence of the 14 disaggregated symptoms (i.e. separately assessing weight loss, weight gain, decreased appetite, and increased appetite). Diagnoses were generated from the interview response data using a computer algorithm. For the last year, onsets and offsets of episodes of MD were estimated in units of a month. Therefore, current MD means that the individual reported symptoms meeting criteria for MD in the month of the interview.

https://doi.org/10.1017/S0033291707000244 Published online by Cambridge University Press

Statistical analysis

Regression analyses

Regression analyses were used to describe the association between personality at two time-points and the prevalence of MD between these time-points. Logistic regression was used when MD was the dependent variable, and linear regression when personality was the dependent variable. Generalized estimating equations (GEE) corrected for spuriously low standard errors caused by the non-independence of two members of a twin pair. All analyses were implemented in SAS using the procedure GENMOD (SAS Institute, Cary, NC, USA).

We performed five kinds of regression analyses, which we termed descriptive, causal, prodromal, state and scar. In the descriptive analyses, we sought to describe the overall association between N, E and MD and therefore included all individuals, using their personality scores to predict lifetime or last-year MD.

In the causal analyses, we eliminated twins with one or more episodes of MD prior to or at time 1 and examined whether N_1 or E_1 could prospectively predict a first onset of MD in the year prior to the time 2 interview.

In the prodromal analyses, our goal was to determine whether any of the putative causal relationships between N, E and MD could result from these personality measures reflecting prodromal depressive symptoms. Therefore, we tested, only in individuals who went on to develop a new onset of MD, whether N_1 and E_1 predicted the time to onset of their episode.

In the state analyses, we sought to determine whether being in an episode of MD influenced personality scores. We eliminated twins who were in an episode of MD at their time 1 assessment (so that this assessment was not influenced by state effects of MD), and then, controlling for their time 1 personality score, determined whether being in an episode of MD at their time 2 assessment influenced time 2 personality scores.

In the scar analyses, our goal was to determine whether personality was influenced by prior episodes of MD. Therefore, we eliminated twins who had had an episode of MD prior to or at their time 1 assessment (so that their personality was unaffected by possible earlier scar effects) and twins who were in an episode of MD at their time 2 assessment (the presence of which would contaminate the scar and state effects).

In the remaining sample, we determined whether having had an episode of MD in the year between the time 1 and time 2 assessments influenced the time 2 personality scores, controlling for the time 1 personality scores.

Twin modeling

We constructed a longitudinal, structural equation twin model for the relationship between personality, as assessed at two times of measurement, and the one-year prevalence of MD that occurs between these times. We assume a multifactorial-threshold model for MD, the strengths and limitations of which have been discussed elsewhere (Neale, 1992). Owing to the substantially lower prevalence of last-year than lifetime MD, there was substantially reduced power to detect a genetic effect in this sample. We therefore used a more inclusive definition of an episode of MD, which differed from standard DSM-III-R criteria in the following ways only: subjects were positive if they endorsed three or more symptoms for 1 week or more ($n = 286$, prevalence = 9·43 %), rather than at least five symptoms for 2 weeks or more ($n = 106$, prevalence = 3·49 %). Criteria were otherwise identical to those for standard DSM-III-R criteria.

The full model (model 1) incorporated two sets of additive genetic (A), common environmental (C), and individual specific environmental factors (E): those common to all three observed variables (level 1), and those specific to each one (level 2). In addition, it included two causal paths (level 3): one from N_1 to MD, (α) and another from MD to N_2 (β). This causal effect reflects the combined impact of both 'scar' effects of episodes of MD in the year between time 1 and time 2 and 'state' effects of individuals who are in an episode of MD at the time 2 assessment. In addition, in order to assure identification of the two causal paths, we assume that the genetic and environmental factors produce an equal influence on time 1 and time 2 N. This is a reasonable assumption because this is a sample in early to mid-adult life and the time period between the two points of measurement is relatively short. In one study, the heritability of N declined with age (Viken *et al.* 1994). Furthermore, we added MD prior

https://doi.org/10.1017/S0033291707000244 Published online by Cambridge University Press

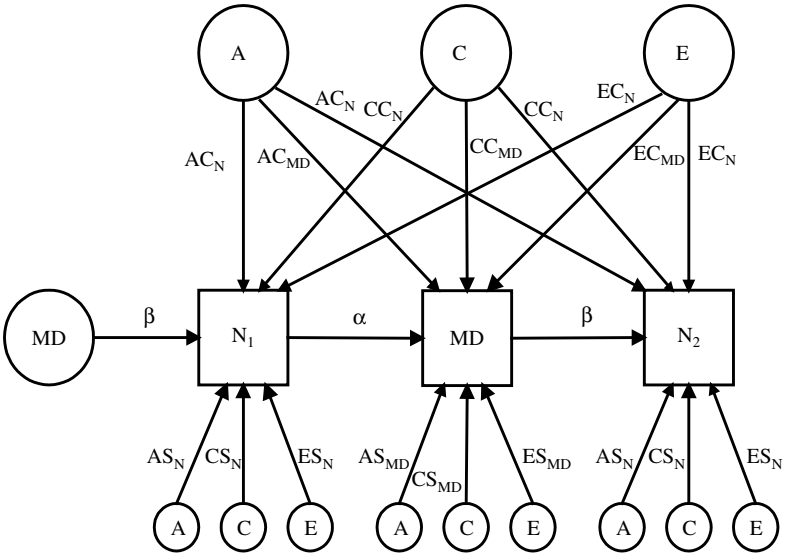

FIG. 1. A longitudinal, structural-equation twin model for the interrelationship between neuroticism as assessed at two time-points (designated herein as N_1 and N_2) and the 1-year prevalence of major depression (MD) occurring between these two times of measurement. The model contains three 'levels' of paths, represented as arrows: (1) those from the additive genetic, common environmental, and individual-specific environmental factors at the top of the figure that *commonly* influence N as assessed at both times (AC_N, CC_N, and EC_N, respectively) *as well as* 1-year prevalence of MD (AC_{MD}, CC_{MD}, and EC_{MD}, respectively); (2) those from the additive genetic, common environmental and individual-specific environmental factors at the bottom of the figure that *specifically* influence N at one of the two time-points (AS_N, CS_N, and ES_N, respectively) *or* MD (AS_{MD}, CS_{MD}, and ES_{MD}, respectively); and lastly, (3) direct causal paths connecting N_1 to MD (α) and MD to N_2 (β). Latent variables are depicted in circles, while observed variables are depicted in squares.

to time 1 as an additional latent variable in order to equalize as much as possible the inputs to N_1 and N_2. This model is depicted graphically in Fig. 1. To choose submodels to test, we systematically eliminated paths at each of the three levels sequentially, and used the 95% confidence intervals (CI) of the parameter estimates of each of these paths in the full model to guide us.

Models were fitted directly to raw data matrices by maximum likelihood using the computer program MX (Neale *et al.* 1999). Twelve thresholds were specified for the N variable, corresponding to the 12 levels of the scale. Alternative models were evaluated by comparing the difference in their χ^2 values relative to the difference in their degrees of freedom, according to the principle of parsimony – models with fewer parameters are preferable if they do not provide significantly worse fit. We operationalize parsimony by using the Akaike Information Criterion (AIC) (Akaike, 1987), calculated as χ^2 minus twice the degrees of freedom. The goal was to produce the model with the lowest (i.e. largest negative) value for the AIC.

RESULTS

Regression analyses

In the epidemiologic, within-person analyses, the following relationships were observed: (*a*) N1 predicted the onset of MD between time 1 and time 2, (*b*) the presence of MD predicted N at time 2, while adjusting for N at time 1, and (*c*) the occurrence of an episode of MD between time 1 and time 2 predicted subsequent N at time 2, again, while adjusting for N at time 1. These findings supported the causal, state, and scar hypotheses, respectively. There was no evidence of the prodromal hypothesis, however, as N at time 1 did not predict the time to onset of MD. There was an overall negative correlation between E and MD although there was no evidence of causal, scar, state, or prodromal effects. Regression coefficients, relative risks, and significance levels are presented in Table 1.

https://doi.org/10.1017/S0033291707000244 Published online by Cambridge University Press

Table 1. *Epidemiological analyses of neuroticism, extraversion, and major depression*

Analysis	Exclusionary criteria	Key variables: Independent	Key variables: Dependent	Neuroticism: N	Neuroticism: β	Neuroticism: P	Neuroticism: RR	Extraversion: β	Extraversion: P	Extraversion: RR
Descriptive	None	P_1	LTMD	3030	**0·59**	**< 0·0001**	**1·8**	**−0·19**	**< 0·0001**	**0·83**
	None	P_2	LTMD	2478	**0·58**	**< 0·0001**	**1·78**	0·05	0·28	
	None	P_1	1-y MD	3030	**1·10**	**< 0·0001**	**3·01**	**−0·23**	**0·0001**	**0·80**
	None	P_2	1-y MD	2478	**1·05**	**< 0·0001**	**2·87**	**−0·15**	**0·04**	**0·86**
Causal	Any previous MD at time 1	P_1	1-y MD	1862	**0·61**	**< 0·0001**	**1·85**	−0·11	0·35	
State	1-y MD at time 1	MD_2	P_1	2248	**0·92**	**< 0·0001**	**2·52**	0·11	0·26	
Scar	1-y MD at time 1 *or* MD at time 2	1-y MD	P_2	2193	**0·63**	**< 0·0001**	**1·88**	0	0·98	
Prodromal	No 1-y MD	P_1	TTO MD	177	0·03	0·66		−0·11	0·14	

Descriptive and causal analyses used logistic regression; state, scar and prodromal analyses used linear regression.
LT, Lifetime; 1-y: one-year; P_1 and P_2, personality at times 1 and 2 respectively; MD_2, MD at time 2; TTO, time to onset.
Bold values denote $p < 0·05$.

With zygosity, age and the twin's current depressive status as control variables, co-twin's lifetime history of MD significantly predicted the level of N ($\beta = 0·304$, $p = 0·02$) but not E ($\beta = -0·147$, $p = 0·18$). When lifetime history of MD was used as a control variable instead of current status, the results were similar, but more significant: co-twin's lifetime history of MD predicted the relative's level of N ($\beta = 0·397$, $p = 0·003$) but not E ($\beta = -0·152$, $p = 0·17$).

Twin modeling

As the parameter estimates in model 1 for the common C paths were close to zero and their 95% CIs included zero, these paths were eliminated in model 2, which had a minimally worse fit than model 1 ($\chi^2 = +0·008$, df = 2), but was preferable (AIC = −3·992). All subsequent models were submodels of model 2. Working at level 2, we eliminated all possible combinations of the following paths: specific A to N, specific A to MD, specific C to N, and specific C to MD, yielding models 3–15 (details available on request). Among these, the best-fitting model was one in which we eliminated the paths of specific A to N, specific A to MD, and specific C to N. This model (model 9) fit slightly worse than model 1 ($\chi^2 = +0·603$), but was preferable because of its parsimony (AIC = −9·397).

Working then from model 9, we sequentially eliminated paths at level 3. First, we eliminated α, yielding model 16, which fit slightly worse than model 9, but was preferable owing to its parsimony (df = 6, AIC = −10·785). We could not improve on this model, as eliminating β to yield model 17 caused substantial worsening of fit ($\chi^2 = 24·384$, AIC = +10·383). Finally, we eliminated the specific C path to MD from model 9 to yield model 18. This model fit slightly worse ($\chi^2 = +3·780$), and although it was more parsimonious (df = 7), it was not significantly different than model 16, differing by less than one unit of AIC (AIC = −10·220). Therefore, model 16 was the best-fitting model. However, the difference in fit between it and the fuller model, model 9, was only 1·4 units. Following the guidelines of Burnham & Anderson (1998), we are therefore unable to reject model 9 as a difference in AIC of less than 2 suggests that the fuller model is preferable. Nevertheless, in model 9, the path that was removed, α, had a lower 95% confidence interval of 0. We were therefore able to reject this model on the grounds that it *both* had a less negative AIC *and* a non-significant path, compared with model 16. Parameter estimates and 95% confidence intervals for model 9 and the full model, model 1, are presented in Table 2.

DISCUSSION

The results we obtained in the epidemiological analyses of N and MD were broadly similar to those from the study of female twins. Consistent with three prior studies (Nystrom & Lindegard, 1975), N predicted the new onset of MD (Tambs *et al.* 1991; Kendler *et al.* 1993). MD predicted N_2 controlling for N_1, consistent with a scar effect, as did our previous study (Kendler *et al.* 1993), while two clinical studies failed to provide evidence for it (Duggan *et al.* 1991; Zeiss &

https://doi.org/10.1017/S0033291707000244 Published online by Cambridge University Press

Table 2. *Parameter estimates in longitudinal structural equation twin model of neuroticism (N) and major depression (MD) in males*

Parameter	Full model	Best-fitting model
A. Paths commonly influencing N and MD		
Additive genetic path to N (AC_N)	0·55 (0·53–0·60)	0·55 (0·50–0·60)
Additive genetic path to MD (AC_{MD})	0·26 (0·07–0·57)	0·31 (0·29–0·43)
Common environmental path to N (CC_N)	0·03 (−0·40–0·40)	—
Common environmental path to MD (CC_{MD})	−0·04 (−0·53–0·53)	—
Individual-specific environmental path to N (EC_N)	0·54 (0·47–0·60)	0·53 (0·48–0·58)
Individual-specific environmental path to MD (EC_{MD})	0·18 (−0·10–0·41)	0·29 (0·15–0·40)
B. Paths specific to N *or* MD		
Additive genetic path to N (AS_N)	0·09 (0·0–0·21)	—
Additive genetic path to MD (AS_{MD})	0·24 (0·0–0·53)	—
Common environmental path to N (CS_N)	0·02 (0·0–0·16)	—
Common environmental path to MD (CS_{MD})	0·24 (0·24–0·51)	0·35 (0·0–0·52)
Individual-specific environmental path to N (ES_N)	0·52 (0·48–0·55)	0·53 (0·52–0·55)
Individual-specific environmental path to MD (ES_{MD})	0·84 (0·83–0·94)	0·83 (0·73–0·92)
C. Direct causal paths		
From N to MD (α)	0·10 (0·0–0·31)	—
From MD to N (β)	0·19 (0·10–0·35)	0·16 (0·10–0·22)

See Fig. 1 for graphic representation of model. Path estimates are standardized regression coefficients, so they must be squared to equal the proportion of variance in the dependent variable accounted for by the dependent variable. Ninety-five per cent confidence intervals are in parentheses. Dashes indicate that a path was eliminated in its respective model.

Lewinsohn, 1988). While controlling for N_1, we were able to isolate the effect of the state of being depressed on concurrent N at time 2. An ongoing episode of MD at time 2 strongly and significantly predicted N_2. The effect size in this analysis (RR 2·52) was greater than that for both the causal (RR 1·85) and scar effects (RR 1·88) and confirms the long-held notion that personality is significantly altered by clinical mood states, as previously reported (Coppen, 1965; Kerr *et al.* 1970; Hirschfeld *et al.* 1983*a*; Farmer *et al.* 2002), validating the DSM-IV's caution against diagnosing Axis II disorders during acute episodes of Axis I disorders. Although we found no evidence supporting the prodromal hypothesis, in which individuals with high N would have a shorter time to onset of MD, our sample size was reduced to only 5·8% of the original sample when we admitted only individuals with MD. Because of the consequent loss of power, we cannot confidently rule out the possibility that N does in fact represent prodromal symptoms of MD.

These epidemiologic analyses clearly demonstrate significant associations between N and MD consistent with previous studies. However, they cannot determine whether one has a direct causal effect on the other, or whether they merely share a common familial liability. The genetic-epidemiologic regression analyses, in which either trait predicted the other *across* twins, suggest that at least a portion of the association between N and MD is a result of shared familial factors. They cannot, however, differentiate genetic from common environmental sources of covariation.

To more definitively resolve these issues, our longitudinal twin model incorporated genetic and environmental factors that are shared by N and MD, those that are specific to each trait, and direct causal paths between N at time 1 and MD, as well as between MD and N at time 2. In the best-fitting model, all of the association between N and MD was the result of additive genetic and individual specific environmental factors shared by both traits, as well as a scar path between MD and N_2. Although in our regression analyses, N_1 predicted the one-year prevalence of MD at time 2, this association was explained in the twin model by common factors influencing both traits. This tells against a direct causal effect of N on MD, but supports the notion that N is an index of the vulnerability to MD resulting from genetic and environmental

factors and suggests that N may be useful as a potential endophenotype in genetic studies of MD. In a recent linkage study of N, regions on chromosomes 1q and 12q were implicated. These two regions overlapped with those implicated, respectively, in either alcoholism or MD (Nurnberger *et al.* 2001), and MD alone (Abkevich *et al.* 2003). However, rigorous methods for evaluating the overlap of specific molecular genetic factors for N and MD have yet to be employed.

There were similarities and differences between the current analyses and the study of female twins previously reported (Kendler *et al.* 1993) that may provide clues about the nature of the sex difference in MD. In both genders, there was evidence of a scar effect, but no evidence that N directly causes MD, suggesting that the causal relationship between the two traits does not differ across the sexes (Kendler & Prescott, 1999). The sex difference seen in MD, therefore, is not likely due to the higher mean N found in women (Floderus-Myrhed *et al.* 1980; Tambs *et al.* 1991; Macaskill *et al.* 1994; Jang *et al.* 1996), as we find no evidence of a causal effect of N on MD in either sex. It is more likely that higher N in women is a consequence of the higher prevalence of MD (Regier *et al.* 1988; Weissman *et al.* 1993; Kessler *et al.* 1994), mediated by a scar effect operant in both sexes.

One major difference was that there was evidence of specific genetic factors influencing *either* trait in women only. In men, N and MD completely shared genetic factors. We previously showed that a model in which the genetic overlap between N and *lifetime* MD was higher in men could not be rejected (Fanous *et al.* 2002). Taken together, these results suggest that there may be greater genetic heterogeneity in women, which could possibly be related to their greater susceptibility to MD. This also suggests that N may be more useful as a proxy endophenotype of MD in men than in women. The heritability of last-year MD was substantially lower in men than in women (12% *v.* 46%). This is consistent with an analysis of the same sample, in which the diagnosis of MD was determined on two occasions to correct for the unreliability of assessment. In that analysis, the heritability of lifetime MD was 30% greater in women than in men (Kendler *et al.* 2001).

The male and female samples also differed in the relationship between E and MD. In the current study of males, there was no evidence supporting previous reports of state effects (Coppen, 1965; Kerr *et al.* 1970; Hirschfeld *et al.* 1983*a*; Farmer *et al.* 2002), or evidence of causal, scar, or prodromal effects of E on MD. However, there was an overall inverse relationship between E and MD in the descriptive analyses, which was not found in the female sample. This difference could be a result of the greater power to detect an effect in the male sample, as it was substantially larger. As the descriptive analyses did not exclude any subjects, they may also have had more power to detect a true effect than did the other analyses, which excluded different portions of the sample. Of note, the causal, state, and scar analyses of E and MD had regression slopes in the expected direction, but they did not reach statistical significance, although in the causal analysis, there was a trend towards significance ($p=0{\cdot}07$). A sex difference in the relationship between E and MD may be a result of greater emotional expression in women (Kring & Gordon, 1998), leading to a greater degree of social interaction in women with MD than in their male counterparts.

The results of this report should be interpreted in the context of five potentially important methodological limitations. First, as the criteria for MD used in the twin models presented here were less stringent than those used in female twins, i.e. three *versus* five symptoms (see Method section), differences between these models and those presented for female twins should be interpreted with caution. However, in the female sample, the number of depressive symptoms in an episode was related to the genetic risk of MD in a monotonic fashion, telling against the validity of the current requirement in DSM-IV that five symptoms be present to meet criteria for an episode of MD (Kendler & Gardner, 1998). Second, our elimination of different segments of the sample in the regression analyses to isolate the causal, state, scar, and prodromal effects is susceptible to selection effects. Furthermore, as the prevalence of MD differs in men and women, the selection effects may also be different, further complicating the comparison between the male and female samples. Third, our use of the word 'causal'

https://doi.org/10.1017/S0033291707000244 Published online by Cambridge University Press

should be understood with the proviso that we cannot directly assess causal effects of one trait on another using the current design, or any other that we know of. Causal effects in the twin model are inferred by the covariance of N and MD *not* accounted for by genetic and environmental factors common to both traits, and by the 'causative' variable temporally preceding the 'caused' variable. Fourth, the diagnoses of MD were derived from face-to-face interviews in 79·8% and by telephone interviews in the remainder. However, there was little difference in the assessment of psychiatric disorders or in the heritability of MD by telephone as opposed to face-to-face interviews in a previous study of female twins (Kendler *et al.* 1992). Fifth, inclusion of twins with a wide range of ages makes it difficult to detect potential difference in the relationship between personality and MD in different age groups. One study reported a decline in the genetic influences on N with age (Viken *et al.* 1994), although it has not been replicated. Lastly, the one-year interval between interviews, as well as the fact that the sample is of adults, makes it difficult to test more developmental hypotheses about the relationship between personality and MD, for example, whether N in childhood or adolescence predicts MD in adult life.

ACKNOWLEDGEMENTS

This work was supported by NIH grants MH-40828, MH/AA/DA-49492, IT-32 MH-20030, DA-11287 and AA-09095. We acknowledge the contribution of the Virginia Twin Registry, now part of the Mid-Atlantic Twin Registry (MATR), to ascertainment of subjects for this study. The MATR, directed by Dr Judy L. Silberg, has received support from the National Institutes of Health, the Carman Trust and the W. M. Keck, John Templeton and Robert Wood Johnson Foundations. Sarah Burns, M.A., Frank Butera, M.S., and Patsy Waring directed data collection. Indrani Ray, B.S., provided database support. Dr Carol Prescott played a major role in the collection of this data.

DECLARATION OF INTEREST

There is no conflict of interest with regard to this work.

REFERENCES

Abkevich, V., Camp, N. J., Hensel, C. H., Neff, C. D., Russell, D. L., Hughes, D. C., Plenk, A. M., Lowry, M. R., Richards, R. L., Carter, C., Frech, G. C., Stone, S., Rowe, K., Chau, C. A., Cortado, K., Hunt, A., Luce, K., O'Neil, G., Poarch, J., Potter, J., Poulsen, G. H., Saxton, H., Bernat-Sestak, M., Thompson, V., Gutin, A., Skolnick, M. H., Shattuck, D. & Cannon-Albright, L. (2003). Predisposition locus for major depression at chromosome 12q22-12q23.2. *American Journal of Human Genetics* **73**, 1271–1281.

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika* **52**, 317–332.

Akiskal, H. S. & Pinto, O. (1999). The evolving bipolar spectrum. Prototypes I, II, III, and IV. *Psychiatric Clinics of North America* **22**, 517–534, vii.

Akiskal, H. S., Hirschfeld, R. M. & Yerevanian, B. I. (1983). The relationship of personality to affective disorders. *Archives of General Psychiatry* **40**, 801–810.

Bartholomew, A. A. (1959). The temporal reliability of the Maudsley Personality Inventory. *Journal of Mental Science* **105**, 238–240.

Boyce, P. & Parker, G. (1985). Neuroticism as a predictor of outcome in depression. *Journal of Nervous and Mental Disease* **173**, 685–688.

Boyce, P., Parker, G., Barnett, B., Cooney, M. & Smith, F. (1991). Personality as a vulnerability factor to depression. *British Journal of Psychiatry* **159**, 106–114.

Burnham, K. P. & Anderson, D. R. (1998). *Model Selection and Inference: a Practical Information – Theoretic Approach*. Springer-Verlag: New York.

Coppen, A. (1965). Effects of a depressive illness on MPI scores. *British Journal of Psychiatry* **111**, 236–239.

Cox, W. M. (1985). Personality correlates of substance abuse. In *Determinants of Substance Abuse: Biological, Psychological, and Environmental Factors* (ed. M. Galizio and S. A. Maisto), pp. 209–246. Plenum Press: New York.

Czerwinski, S. A., Mahaney, M. C., Williams, J. T., Almasy, L. & Blangero, J. (1999). Genetic analysis of personality traits and alcoholism using a mixed discrete continuous trait variance component model. *Genetic Epidemiology* **17** (suppl. 1), S121–S126.

Duggan, C. F., Lee, A. S. & Murray, R. M. (1990). Does personality predict long-term outcome in depression? *British Journal of Psychiatry* **157**, 19–24.

Duggan, C. F., Sham, P., Lee, A. S. & Murray, R. M. (1991). Does recurrent depression lead to a change in neuroticism? *Psychological Medicine* **21**, 985–990.

Eysenck, H. J. & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire*. Hodder & Stoughton: London.

Eysenck, H. J. & Eysenck, M. W. (1985). *Personality and Individual Differences. A Natural Science Approach*. Plenum Press: New York.

Fanous, A., Gardner, C. O., Prescott, C. A., Cancro, R. & Kendler, K. S. (2002). Neuroticism, major depression and gender: a population-based twin study. *Psychological Medicine* **32**, 719–728.

Farmer, A., Redman, K., Harris, T., Mahmood, A., Sadler, S., Pickering, A. & McGuffin, P. (2002). Neuroticism, extraversion, life events and depression. The Cardiff Depression Study. *British Journal of Psychiatry* **181**, 118–122.

Floderus-Myrhed, B., Pedersen, N. & Rasmuson, I. (1980). Assessment of heritability for personality, based on a short-form of the Eysenck Personality Inventory: a study of 12898 twin pairs. *Behavior Genetics* **10**, 153–162.

Gottesman, I. I. & Gould, T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *American Journal of Psychiatry* **160**, 636–645.

Hirschfeld, R. M. & Cross, C. K. (1983). Personality, life events, and social factors in depression. In *Psychiatry Update: The American Psychiatric Association Annual Review* (ed. L. Grinspoon), pp. 382–405. American Psychiatric Press: Washington, DC.

Hirschfeld, R. M., Klerman, G. L., Clayton, P. J. & Keller, M. B. (1983*b*). Personality and depression. Empirical findings. *Archives of General Psychiatry* **40**, 993–998.

Hirschfeld, R. M., Klerman, G. L., Clayton, P. J., Keller, M. B., McDonald-Scott, P. & Larkin, B. H. (1983*a*). Assessing personality: effects of the depressive state on trait measurement. *American Journal of Psychiatry* **140**, 695–699.

Hirschfeld, R. M., Klerman, G. L., Lavori, P., Keller, M. B., Griffith, P. & Coryell, W. (1989). Premorbid personality assessments of first onset of major depression. *Archives of General Psychiatry* **46**, 345–350.

Jackson, S. W. (1986). *Melancholia and Depression: From Hippocratic Times to Modern Times*. Yale University Press: New Haven, CT.

Jang, K. L., Livesley, W. J. & Vernon, P. A. (1996). Heritability of the big five personality dimensions and their facets: a twin study. *Journal of Personality* **64**, 577–591.

Katz, R. & McGuffin, P. (1987). Neuroticism in familial depression. *Psychological Medicine* **17**, 155–161.

Kendler, K. S. & Gardner Jr., C. O. (1998). Boundaries of major depression: an evaluation of DSM-IV criteria. *American Journal of Psychiatry* **155**, 172–177.

Kendler, K. S. & Prescott, C. A. (1999). A population-based twin study of lifetime major depression in men and women. *Archives of General Psychiatry* **56**, 39–44.

Kendler, K. S., Gardner, C. O., Neale, M. C. & Prescott, C. A. (2001). Genetic risk factors for major depression in men and women: similar or different heritabilities and same or partly distinct genes? *Psychological Medicine* **31**, 605–616.

Kendler, K. S., Karkowski, L. M., Neale, M. C. & Prescott, C. A. (2000). Illicit psychoactive substance use, heavy use, abuse, and dependence in a US population-based sample of male twins. *Archives of General Psychiatry* **57**, 261–269.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C. & Eaves, L. J. (1992). A population-based twin study of major depression in women. The impact of varying definitions of illness. *Archives of General Psychiatry* **49**, 257–266.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C. & Eaves, L. J. (1993). A longitudinal twin study of personality and major depression in women. *Archives of General Psychiatry* **50**, 853–862.

Kendler, K. S., Neale, M. C. & Walsh, D. (1995). Evaluating the spectrum concept of schizophrenia in the Roscommon Family Study. *American Journal of Psychiatry* **152**, 749–754.

Kerr, T. A., Schapira, K., Roth, M. & Garside, R. F. (1970). The relationship between the Maudsley Personality Inventory and the course of affective disorders. *British Journal of Psychiatry* **116**, 11–19.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H. U. & Kendler, K. S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Archives of General Psychiatry* **51**, 8–19.

Kraepelin, E. (1921). *Manic-Depressive Illness and Paranoia*. E & S Livingstone: Edinburgh, UK.

Kretschmer, E. (1936). *Physique and Character*. Kegan Paul Trench Trubner & Co.: London.

Krieg, J. C., Lauer, C. J., Hermle, L., von Bardeleben, U., Pollmacher, T. & Holsboer, F. (1990). Psychometric, polysomnographic, and neuroendocrine measures in subjects at high risk for psychiatric disorders: preliminary results. *Neuropsychobiology* **23**, 57–67.

Kring, A. M. & Gordon, A. H. (1998). Sex differences in emotion: expression, experience, and physiology. *Journal of Personality and Social Psychology* **74**, 686–703.

Lauer, C. J., Bronisch, T., Kainz, M., Schreiber, W., Holsboer, F. & Krieg, J. C. (1997). Pre-morbid psychometric profile of subjects at high familial risk for affective disorder. *Psychological Medicine* **27**, 355–362.

Macaskill, G. T., Hopper, J. L., White, V. & Hill, D. J. (1994). Genetic and environmental variation in Eysenck Personality Questionnaire scales measured on Australian adolescent twins. *Behavior Genetics* **24**, 481–491.

Maier, W., Lichtermann, D., Minges, J. & Heun, R. (1992). Personality traits in subjects at risk for unipolar major depression: a family study perspective. *Journal of Affective Disorders* **24**, 153–163.

Maier, W., Minges, J., Lichtermann, D., Franke, P. & Gansicke, M. (1995). Personality patterns in subjects at risk for affective disorders. *Psychopathology* **28** (suppl. 1), 59–72.

Mezey, A. G., Cohen, S. I. & Knight, E. J. (1963). Personality assessment under varying physiological and psychological conditions. *Journal of Psychosomatic Research* **7**, 237–240.

Neale M. C. & Cardon, L. R. (1992). *Methodology for Genetic Studies of Twins and Families*. Kluwer Academic: Dordrecht, The Netherlands.

Neale, M. C. & Eaves, L. J. (1993). Estimating and controlling for the effects of volunteer bias with pairs of relatives. *Behavior Genetics* **23**, 271–277.

Neale M. C., Boker, S. M., Xie, G. & Maes, H. H. (1999). *Mx: Statistical Modeling*, 5th edn. Medical College of Virginia of Virginia Commonwealth University, Department of Psychiatry: Richmond, VA.

Nurnberger Jr., J. I., Foroud, T., Flury, L., Su, J., Meyer, E. T., Hu, K., Crowe, R., Edenberg, H., Goate, A., Bierut, L., Reich, T., Schuckit, M. & Reich, W. (2001). Evidence for a locus on chromosome 1 that influences vulnerability to alcoholism and affective disorder. *American Journal of Psychiatry* **158**, 718–724.

Nystrom, S. & Lindegard, B. (1975). Predisposition for mental syndromes: a study comparing predisposition for depression, neurasthenia and anxiety state. *Acta Psychiatrica Scandanavica* **51**, 69–76.

Ouimette, P. C., Klein, D. N. & Pepper, C. N. (1992). Personality traits in offspring of parents with unipolar affective disorder: an exploratory study. *Journal of Affective Disorders* **1**, 91–98.

Regier, D. A., Boyd, J. H., Burke Jr., J. D., Rae, D. S., Myers, J. K., Kramer, M., Robins, L. N., George, L. K., Karno, M. & Locke, B. Z. (1988). One-month prevalence of mental disorders in the United States. Based on five Epidemiologic Catchment Area sites. *Archives of General Psychiatry* **45**, 977–986.

Riley, B. P. & McGuffin, P. (2000). Linkage and associated studies of schizophrenia. *American Journal of Medical Genetics* **97**, 23–44.

Schneider, K. (1958). *Psychopathic Personalities*. Cassell: London.

Sher, K. J. & Trull, T. J. (1994). Personality and disinhibitory psychopathology: alcoholism and antisocial personality disorder. *Journal of Abnormal Psychology* **103**, 92–102.

Siever, L. J., Kalus, O. F. & Keefe, R. S. (1993). The boundaries of schizophrenia. *Psychiatric Clinics of North America* **16**, 217–244.

Spitzer, R. L., Williams, J. B. & Gibbon, J. (1987). *Structured Clinical Interview for DSM-III-R Patient Version*. New York Psychiatric Institute: New York.

Tambs, K., Sundet, J. M., Eaves, L., Solaas, M. H. & Berg, K. (1991). Pedigree analysis of Eysenck Personality Questionnaire (EPQ) scores in monozygotic (MZ) twin families. *Behavior Genetics* **21**, 369–382.

Viken, R. J., Rose, R. J., Kaprio, J. & Koskenvuo, M. (1994). A developmental genetic analysis of adult personality: extraversion and neuroticism from 18 to 59 years of age. *Journal of Personality and Social Psychology* **66**, 722–730.

Weissman, M. M., Bland, R., Joyce, P. R., Newman, S., Wells, J. E. & Wittchen, H. U. (1993). Sex differences in rates of depression: cross-national perspectives. *Journal of Affective Disorders* **29**, 77–84.

Weissman, M. M., Prusoff, B. A. & Klerman, G. L. (1978). Personality and the prediction of long-term outcome of depression. *American Journal of Psychiatry* **135**, 797–800.

Wetzel, R. D., Cloninger, C. R., Hong, B. & Reich, T. (1980). Personality as a subclinical expression of the affective disorders. *Comprehensive Psychiatry* **21**, 197–205.

Zeiss, A. M. & Lewinsohn, P. M. (1988). Enduring deficits after remissions of depression: a test of the scar hypothesis. *Behavioral Research Therapy* **26**, 151–158.