A population-based twin study in women of smoking initiation and nicotine dependence


From the Departments of Psychiatry and Human Genetics, Virginia Institute for Psychiatric and Behavioral Genetics, Medical College of Virginia Commonwealth University, Richmond, VA, USA

ABSTRACT

Background. The development of drug dependence requires prior initiation. What is the relationship between the risk factors for initiation and dependence?

Methods. Using smoking as a model addiction, we assessed smoking initiation (SI) and nicotine dependence (ND) by personal interview in 1898 female twins from the population-based Virginia Twin Registry. We developed a twin structural equation model that estimates the correlation between the liability to SI and the liability to ND, given SI.

Results. The liabilities to SI and ND were substantially correlated but not identical. Heritable factors played an important aetiological role in SI and in ND. While the majority of genetic risk factors for ND were shared with SI, a distinct set of familial factors, which were probably partly genetic, solely influenced the risk for ND. SI was associated with low levels of education and religiosity, high levels of neuroticism and extroversion and a history of a wide range of psychiatric disorders. ND was associated with low levels of education, extroversion, mastery, and self-esteem, high levels of neuroticism and dependency and a history of mood and alcohol use disorders.

Conclusions. The aetiological factors that influence SI and ND, while overlapping, are not perfectly correlated. One set of genetic factors plays a significant aetiological role in both SI and ND, while another set of familial factors, probably in part genetic, solely influences ND. Some risk factors for SI and ND impact similarly on both stages, some act at only one stage and others impact differently and even in opposite directions at the two stages. The pathway to substance dependence is complex and involves multiple genetic and environmental risk factors.

INTRODUCTION

The development of psychoactive substance dependence requires prior drug exposure (Kozlowski & Harford, 1976). What is the relationship between the risk factors for initiation and dependence? Are the risk factors for these two stages similar or different? In this report, we take a genetic–epidemiological approach to this problem, exemplified by smoking initiation (SI) and nicotine dependence (ND).

A large proportion of cigarette smokers develop ND as nicotine is highly addictive (Hughes et al. 1987; US Department of Health and Human Services, 1988, 1994; Schelling, 1992; Hale et al. 1993; Breslau et al. 1994; American Psychiatric Association, 1996). Studies in rodents suggest genetic influences on the number of brain nicotine receptors (Marks et al. 1989) and the development of nicotine tolerance (Collins & Marks, 1989). In man, one adoption study (Eysenck, 1965) and many twin studies (Fisher, 1958; Friberg et al. 1959; Todd & Mason, 1959; Raaschou-Nielsen, 1960; Conterio & Chiarelli, 1962; Cederlof et al. 1977; Pedersen, 1981; Kaprio et al. 1982, 1984; Hannah et al. 1985; Carmelli et al. 1990; Heath...
et al. 1993; Kendler et al. 1993a; Boomsma et al. 1994; Edwards et al. 1995) (including small samples of twins reared apart (Fisher, 1958; Raaschou-Nielsen, 1960; Shields, 1962; Kaprio et al. 1984)) have examined smoking behaviour and consistently shown evidence for genetic influences. While most studies utilized only simple measures of cigarette use, such as ever v. never smoker, a handful have examined characteristics that may indirectly reflect ND, such as smoking persistence (Heath & Madden, 1995), heavy smoking (Kaprio et al. 1982) or cigarette consumption (Carmelli et al. 1990). We are unaware of studies in genetically informative populations that directly examined ND.

We seek here to answer three questions regarding SI and ND in a sample of 1898 female twins from a birth-certificate based twin registry: (1) What is the relationship between the genetic and environmental risk factors for SI v. ND?; (2) What is the magnitude of the aetiological role for familial and genetic factors in SI and ND?; and (3) Examining a wide range of potential correlates of smoking behaviour, can we find variables that differentially predict SI v. ND?

**METHOD**

**Sample and interview methods**

The Caucasian female same-sex twins studied in this report are part of a longitudinal study of genetic and environmental risk factors for common psychiatric and substance use disorders. The twins, ascertained from the population-based Virginia Twin Registry, were eligible to participate in wave 1 of this study if both members of the pair had previously responded to a mailed questionnaire, to which the individual response rate was 64%. In our wave 1 personal interview, we succeeded in interviewing 92% of the eligible individuals (N = 2163). Ninety per cent of the interviews were face-to-face, while the rest were completed by phone. Written informed consent was obtained prior to all face-to-face interviews while verbal assent was obtained for all phone interviews. The mean age of the participating twins was 30.1 ± 7.6 years. Zygosity was determined blindly by standard questions (Eaves et al. 1989), photographs, and when necessary, DNA (Spence et al. 1988).

We have completed two subsequent waves of telephone interviews, which succeeded in interviewing 2001 (92.5%) and 1898 (87.7%) of the original sample, respectively. The mean ± s.d. of months between the wave 1 and wave 3 interviews was 61.3 ± 5.1.

**Measures**

Our wave 3 interview section on smoking behaviour began with the question ‘Have you ever smoked regularly for at least a month?’ We defined ‘regular smoking’ as: ‘a regular pattern of use in which you would average at least 7 cigarettes per week for a month or more’. Those who responded positively were considered to have initiated smoking. We then asked each twin to define the period in her life when she smoked most heavily. We assessed ND for that period using the 8-item Fagerstrom Tolerance Questionnaire (Fagerstrom & Schneider, 1989), adapted for retrospective reporting, and four additional items, two of which were drug dependence criteria in DSM-IV (American Psychiatric Association, 1994): ‘During this time when you smoked most heavily, how often would you check to make sure you had cigarettes around to smoke?’ and ‘Have you ever seriously attempted to stop smoking?’. Two were included for their face validity: ‘Do you currently smoke regularly?’ and ‘During this time when you smoked most heavily, if you didn’t smoke for a period of time, how strong would your craving get for another cigarette?’.

We assessed a range of putative risk factors for smoking, the specific formats of which we have outlined elsewhere (Heath & Madden, 1995): three dimensions of religiosity (Kendler et al. 1997); eight personality-like traits assessed at wave 1 – neuroticism (Eysenck et al. 1985), extroversion (Eysenck et al. 1985), altruism (defined as the 7 items from The Interpersonal Reactivity Index (Davis, 1980) assessing sensitivity to the needs and feelings of others), interpersonal dependency (Hirschfeld et al. 1977), mastery (defined as a reversed coding of the Maddi et al. Powerlessness subscale of the Alienation test (Maddi et al. 1979), dispositional optimism (Scheier & Carver, 1985), self-esteem (Rosenberg, 1965) and locus of control (based on the learned resourcefulness subscale of the Attributional Style Questionnaire (Peterson et al. 1982)); and the lifetime history of seven
DSM-III-R-based syndromes (American Psychiatric Association, 1987) as assessed in the three waves of interviews — major depression (MD), generalized anxiety disorder (GAD — where we used a 1-month rather than a 6-month minimum duration of illness), panic disorder, phobias (diagnosed when an irrational fear was accompanied by significant objective behaviour change or impairment), bulimia nervosa and alcohol dependence. Details of these diagnostic algorithms are available elsewhere (Kendler et al. 1991, 1992a, b, c; 1993b).

Statistics
The structural equation model used in this report, a causal, contingent, common pathway model, is shown in Fig. 1. The model is: causal because it assumes a direct path (b) from the liability to initiation to the liability to dependence; contingent because dependence can be assessed only in those who have initiated smoking; and common pathway (Kendler et al. 1987), in that genetic and environmental effects on initiation can only affect dependence by flowing through the observed phenotype of initiation. This constraint is necessary to identify the model as we lack assessment of ND in twins who never initiated smoking.

The risk factors for initiation are divided into additive genetic (A), common environmental (C) and individual specific environmental components (E), where the subscript, indicates that they are specific for initiation (Neale & Cardon, 1992; Kendler, 1993). Individuals with high liability on this dimension initiate substance use and then become susceptible to develop dependence.

The liability to dependence derives from two sources: risk factors shared with initiation (reflected in path b) and risk factors independent of initiation (subdivided into additive genetic (A), common environmental (C) and individual specific environmental components (E), where the subscript, indicates specificity for dependence.

If all risk factors for SI similarly influenced risk for ND, then b would approach 1-0, and a, c, and e would approach zero. If risk factors for SI and ND were unrelated, b would approach zero and dependence would be entirely due to the effects of a, c, and e. If risk factors for SI and ND are related, but not identical, then b would be between zero and unity, and ND would be caused both by genetic and environmental risk factors that were in common with SI and unique to ND.

Our model is based on a multiple threshold model, widely used in genetic studies (Reich et al. 1972). Given that SI is binary, and we trichotomized ND scores for these analyses, there are six possible individual outcomes. However, only four of these can be observed because ND cannot be assessed in the absence of

Fig. 1. A bivariate twin model for substance initiation and subsequent substance dependence. The model begins with the risk factors for initiation (indicated by the subscript), which are divided into additive genetic (A), common environmental (C) and individual specific environmental components (E). For individuals above the threshold on this dimension, initiation occurs and they become susceptible to develop dependence. The vulnerability to dependence derives from two sources: (i) risk factors shared between initiation and dependence, reflected in path (b), and (ii) risk factors that influence dependence are those which are independent of the liability to initiation. These are also subdivided into additive genetic (A), common environmental (C) and individual specific environmental components (E), where the subscript indicates that they are specific for dependence. Path coefficients, indicated by lower case letters (a, c and e), reflect standardized partial regression coefficients. Therefore, the proportion of variance in the dependent variables accounted for by the independent variable is equal to the square of the connecting path. Heritability of initiation, for example, equals a². Observed variables are depicted in boxes and latent variables in circles and ellipses. The model is constrained so that a²+c²+e² = 1.0 and b²+a²+c²+e² = 1.0. Furthermore, the total heritable influences on drug dependence can be subdivided into those that are shared with initiation, which equal a²×b² and those that are specific to dependence a².
Possible observed outcomes for a pair of twins assessed on a dichotomous smoking initiation variable (SI− v. SI+), and a nicotine dependence (ND) variable trichotomized into low (L), medium (M) and high (H). ND is assessed only in those who initiate, SI+.

Table 1.

<table>
<thead>
<tr>
<th>Twin 1</th>
<th>Twin 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>SI−</td>
<td>SI−</td>
</tr>
<tr>
<td>LND</td>
<td>LND</td>
</tr>
<tr>
<td>MND</td>
<td>MND</td>
</tr>
<tr>
<td>HND</td>
<td>HND</td>
</tr>
<tr>
<td>SI+</td>
<td>SI+</td>
</tr>
<tr>
<td>LND</td>
<td>LND</td>
</tr>
<tr>
<td>MND</td>
<td>MND</td>
</tr>
<tr>
<td>HND</td>
<td>HND</td>
</tr>
</tbody>
</table>

As shown in Table 1, we can cross-tabulate these four outcomes in members of a twin pair. The probability of observing the 16 possible outcomes depends on: (i) the threshold for SI; (ii) the two thresholds that separate low from medium from high ND; and (iii) the 4 × 4 covariance matrix between liabilities to SI and ND in twin 1 and twin 2. The covariance matrix is predicted from the parameters \( a_i, c_i, e_i, b, a_d, c_d \) and \( e_d \) of the model in Fig. 1, according to the rules of path analysis (Bollen, 1989; Neale & Cardon, 1992). The cell proportions are computed by integrating the multivariate normal distribution over 2, 3 or 4 dimensions, in cell 1, cells 2, 3, 4, 8, 9, 10 and cells 5, 6, 7, 11, 12, 13, 14, 15, 16 or Table 1, respectively. These integrals were computed with the function mnor in Mx (Neale, 1997), which uses the numerical integration software of Genz (1992).

Predicted cell frequencies are obtained by multiplying the predicted cell proportions by the observed sample size for that group. The model is fitted by minimizing the \( \chi^2 \) computed for the observed and expected frequencies. Minimum \( \chi^2 \) has been shown to have superior properties to maximum likelihood estimates when cell frequencies are low (Agresti, 1990). We utilize Akaike's Information Criteria (AIC) as a guide to selecting the model with the optimal combination of goodness-of-fit and parsimony (Akaike, 1987).

These models assume that MZ and DZ twins are equally correlated in their exposure to environmental risk factors for the trait under investigation. We tested this assumption using logistic regression where, controlling for zygosity, we predict twin concordance from measures of environmental similarity in childhood (Loehlin & Nichols, 1976) and frequency of contact as adults.

Factor analysis was carried out by the method of principal components using PROFACTOR in SAS (SAS Institute, 1990). The relationships between the putative risk factors and SI and ND were assessed by logistic regression, which also estimates odds ratios (ORs). For comparability, ND was dichotomized by a median split. All the independent variables, except lifetime psychiatric diagnoses, were standardized. Therefore, the OR indicates the changes in risk for the dependent measure for every s.d. change in the independent variable.

We tested whether putative predictor variables for smoking were equally related to the first stage of smoking behaviour (abstinence v. SI) and to the second stage (low v. high ND) by means of the continuation ratio (CR). When the CR is positive, the variable has a greater impact on the second than the first stage of smoking. When negative, the variable has a smaller impact on the second than the first stage of smoking. The significance of the CR is tested by logistic regression (MacLean, 1988).

RESULTS

Assessment of nicotine dependence

Of the 1898 interviewed twins, 759 (40.0%) reported SI. A scree plot derived from the 12 interview items that assessed ND at the time of the respondent's heaviest smoking revealed the presence of one large factor explaining 28.5% of the total variance. Table 2 depicts the loadings of these 12 items on this first factor. For further
Twin study of smoking initiation and nicotine dependence

Table 2. Items used to assess nicotine dependence in lifetime regular smokers and factor loadings on first principal component

<table>
<thead>
<tr>
<th>Item summary</th>
<th>Origin</th>
<th>Response format</th>
<th>Factor loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 First cigarette how soon after waking</td>
<td>FTQ</td>
<td>Minutes</td>
<td>+049</td>
</tr>
<tr>
<td>2 Difficulty refraining where forbidden</td>
<td>FTQ</td>
<td>Very, somewhat, a little, none</td>
<td>+072</td>
</tr>
<tr>
<td>3 First cigarette most satisfying</td>
<td>FTQ*</td>
<td>Yes/no</td>
<td>+036</td>
</tr>
<tr>
<td>4 Cigarettes/day when smoking most heavily</td>
<td>FTQ</td>
<td>Number of cigarettes</td>
<td>+070</td>
</tr>
<tr>
<td>5 Smoked most in AM</td>
<td>FTQ</td>
<td>Yes/no</td>
<td>+034</td>
</tr>
<tr>
<td>6 Smoked when ill in bed</td>
<td>FTQ</td>
<td>Yes/no</td>
<td>+054</td>
</tr>
<tr>
<td>7 Nicotine level of cigarette brand</td>
<td>FTQ</td>
<td>Low, medium, high</td>
<td>+018</td>
</tr>
<tr>
<td>8 Inhale</td>
<td>FTQ</td>
<td>Yes/no</td>
<td>+034</td>
</tr>
<tr>
<td>9 Check to make sure cigarettes available</td>
<td>DSM-IV</td>
<td>Often, sometimes, rarely, never</td>
<td>+061</td>
</tr>
<tr>
<td>10 Ever seriously tried to quit</td>
<td>DSM-IV</td>
<td>Yes/no</td>
<td>+015</td>
</tr>
<tr>
<td>11 Strength craving when denied</td>
<td>—</td>
<td>Very, somewhat, moderate, hardly any</td>
<td>+079</td>
</tr>
<tr>
<td>12 Current smoker</td>
<td>—</td>
<td>Yes/no</td>
<td>+047</td>
</tr>
</tbody>
</table>

FTQ, Fagerstrom Tolerance Questionnaire (Fagerstrom & Schneider, 1989).


*Modified from original wording of ‘Which cigarette would you hate most to give up?’

analyses, we defined ND as the score of the first factor from these 12 items.

Equal environment assumption

Childhood environment similarity was not associated with twin resemblance for SI ($\chi^2 = 0.86$, all df = 1, NS) or ND ($\chi^2 = 0.39$, NS). Frequency of current contact did not predict SI ($\chi^2 = 0.22$, NS), but was significantly associated with twin resemblance for ND ($\chi^2 = 6.24$, $P = 0.01$).

Twin model fitting

Our sample contained 851 pairs (497 MZ and 354 DZ) with complete information about SI and ND. Concordance rates and tetrachoric correlations for SI were 76.0% and +0.84 in MZ and 61.1% and +0.44 in DZ twins. The sample contained 234 twin pairs (136 MZ and 98 DZ) concordant for SI. In these pairs, substantial twin resemblance was seen for ND which was greater in MZ (Pearson’s $r = +0.47$) than in DZ pairs ($r = +0.30$).

The full model (Fig. 1) fitted well ($\chi^2 = 25.6$, df = 19, $P = 0.14$, AIC = −12.4) and produced the following parameter estimates: $a_i = 0.78$, $c_i = 0.07$, $c_i^2 = 0.15$, $b = +0.77$, $a_i = 0.13$, $c_i = 0.08$ and $c_i^2 = 0.20$. We next constrained $b$ — the path from initiation to dependence — to zero and unity. Both of these models fit substantially worse and were rejected: $\chi^2 = 34.0$, df = 20, AIC = −6.0 and $\chi^2 = 59.3$, df = 22, AIC = +15.3, respectively.

We next attempted to simplify the causes of SI. When we tested a model with no familial–environmental risk factors, we obtained an improvement in the AIC (i.e. $c_i = 0$; $\chi^2 = 25.8$, df = 20, AIC = −14.2). However, a model containing no genetic risk factors for SI fit quite poorly and could be rejected (i.e. $a_i = 0$; $\chi^2 = 57.4$, df = 20, AIC = +17.4).

Assuming an AE model for SI, we then tested for additional familial aggregation for ND by assuming no genetic or familial–environmental risk factors for ND (i.e. $a_i = 0$ and $c_i = 0$). This model failed badly ($\chi^2 = 46.0$, df = 22, AIC = +2.0). Lastly, we compared the fit of models which assumed no familial–environmental risk factors for ND (i.e. $a_i = 0$, $\chi^2 = 26.2$, df = 21, AIC = −15.8) and those which assumed no genetic risk factors for ND (i.e. $a_i = 0$, $\chi^2 = 26.6$, df = 21, AIC = −15.4). Although the model with only genetic risk factors for ND fit better, the difference was quite modest. We cannot, with confidence, determine the extent to which the specific familial aggregation for ND, excluding genetic risk factors for SI, is due to genetic or environmental factors.

The parameter estimates for the best-fit $A_i E_i b A_i E_i$ model were: $a_i = 0.85$, $c_i = 0.15$, $b = +0.77$, $a_i = 0.22$, and $c_i = 0.19$. For both the full and best fit models, the estimates of $b$ are high, indicating a strong relationship between the risk factors for SI and ND. Around 60% of the variance in liability to ND ($0.77^2 = 0.59$) is in common with the liability to SI.
The full model suggests that heritable factors responsible for ND is due to factors specific to dependence. A negative ratio indicates a greater impact on initiation than on dependence, while a positive ratio indicates a greater impact on dependence than on initiation, while OR indicates odds ratio.

### Correlates of smoking initiation and persistence

Table 3 presents results from logistic regressions which examine the relationship between four domains – demographic, religious, personality and lifetime psychopathology – and (i) the probability of SI and (ii) the probability of ND given SI.

SI was significantly predicted by four of nine demographic/religious variables: fewer years of education, low levels of personal religious devotion, low levels of religious institutional

<table>
<thead>
<tr>
<th>Domain</th>
<th>Putative risk factor</th>
<th>Initiation (OR)</th>
<th>Dependence (OR)</th>
<th>Continuation ratio†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic</td>
<td>Age</td>
<td>1.07</td>
<td>1.06</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Years of education</td>
<td>0.72****</td>
<td>0.68****</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>Family income</td>
<td>0.94</td>
<td>0.88</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>Individual income</td>
<td>1.06</td>
<td>0.84*</td>
<td>-0.21**</td>
</tr>
<tr>
<td></td>
<td>Ever divorced</td>
<td>2.51***</td>
<td>1.20</td>
<td>-0.55**</td>
</tr>
<tr>
<td></td>
<td>Currently cohabiting</td>
<td>0.93</td>
<td>1.10</td>
<td>0.15</td>
</tr>
<tr>
<td>Religious</td>
<td>Personal devotion</td>
<td>0.67****</td>
<td>0.94</td>
<td>+0.26**</td>
</tr>
<tr>
<td></td>
<td>Personal conservatism</td>
<td>0.99</td>
<td>1.11</td>
<td>+0.12</td>
</tr>
<tr>
<td></td>
<td>Institutional conservatism</td>
<td>0.84****</td>
<td>1.04</td>
<td>+0.18*</td>
</tr>
<tr>
<td>Personality</td>
<td>Neuroticism</td>
<td>1.25****</td>
<td>1.31***</td>
<td>+0.09</td>
</tr>
<tr>
<td></td>
<td>Extraversion</td>
<td>1.25****</td>
<td>0.84*</td>
<td>-0.35****</td>
</tr>
<tr>
<td></td>
<td>Altruism</td>
<td>0.87**</td>
<td>1.09</td>
<td>+0.20</td>
</tr>
<tr>
<td></td>
<td>Interpersonal dependency</td>
<td>1.00</td>
<td>1.18*</td>
<td>+0.16</td>
</tr>
<tr>
<td></td>
<td>Mastery</td>
<td>0.99</td>
<td>0.76*</td>
<td>-0.26**</td>
</tr>
<tr>
<td></td>
<td>Dispositional optimism</td>
<td>0.91</td>
<td>0.76**</td>
<td>-0.19</td>
</tr>
<tr>
<td></td>
<td>Self-esteem</td>
<td>0.95</td>
<td>0.80**</td>
<td>-0.18</td>
</tr>
<tr>
<td></td>
<td>Locus of control</td>
<td>1.01</td>
<td>0.80**</td>
<td>-0.22</td>
</tr>
<tr>
<td>Lifetime psychopathology</td>
<td>Major depression</td>
<td>1.89****</td>
<td>2.11***</td>
<td>+0.24</td>
</tr>
<tr>
<td></td>
<td>Generalized anxiety disorder</td>
<td>1.97****</td>
<td>2.28***</td>
<td>+0.28</td>
</tr>
<tr>
<td></td>
<td>Panic disorder</td>
<td>3.51****</td>
<td>1.64</td>
<td>-0.51</td>
</tr>
<tr>
<td></td>
<td>Phobia</td>
<td>1.54****</td>
<td>1.33</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>Bulimia</td>
<td>1.58</td>
<td>1.26</td>
<td>-0.14</td>
</tr>
<tr>
<td></td>
<td>Problem drinking</td>
<td>4.11****</td>
<td>1.86***</td>
<td>-0.51**</td>
</tr>
<tr>
<td></td>
<td>Alcohol dependence</td>
<td>5.63****</td>
<td>2.47***</td>
<td>-0.48**</td>
</tr>
</tbody>
</table>

OR indicates odds ratio.

† The continuation ratio tests whether the impact of the risk factor on the initial stage of initiation v. abstinence differs from the impact on the second stage – here, low v. high nicotine dependence. A positive ratio indicates a greater impact on dependence than on initiation, while a negative ratio indicates a greater impact on initiation than on dependence.

‡ P < 0.10 (for continuation ratio only); * P < 0.05; ** P < 0.01; *** P < 0.001; **** P < 0.0001.

Approximately 40% of the variance in liability to ND is due to factors specific to dependence. The full model suggests that heritable factors unrelated to SI account for ~13% of the total variance in liability to ND, or ~30% of the variance in risk factors unique to ND. An additional 8% of the total and 20% of the unique variance to ND is ascribed, in the full model, to the effect of familial–environmental factors that act specifically on ND.

The best-fit model, on the basis of modest statistical power, sets the unique familial–environmental contribution to ND to zero. In this model, genetic factors specific to ND are responsible for ~22% of the total liability to ND and ~55% of the variance in liability that is unique to ND. These results suggest that familial factors, which are at least in part genetic, exist that are specific to ND.

The clauses of variation in liability to ND in the best-fit model can be viewed from another perspective. Risk factors can be divided into the traditional categories of A, C and E and within each of these categories into those that are shared with SI and those that are unique to ND. In this model, genetic factors contribute a total of 72% of the variance in liability to ND, of which 69% also influence SI while 31% are unique to ND. Unique environmental factors contribute 28% of the variance in liability to ND, of which 32% is shared with SI and 68% unique to ND.
conservatism and divorce. By contrast, among smokers, ND was weakly associated with these characteristics, significant relationships being seen only with years of education and individual income. Four of these variables were significantly different in their association with SI and ND: individual income, personal devotion, institutional conservatism and history of divorce. While a history of divorce and low levels of personal devotion and institutional conservatism were strongly associated with an increased risk for SI, they were not predictive of risk for ND in regular smokers.

SI was significantly associated with three of eight personality traits: higher levels of neuroticism (N) and extroversion (E) and lower levels of altruism. By contrast, seven of the eight personality traits were significantly associated with ND—high levels of N and interpersonal dependency and low levels of E, mastery, dispositional optimism, self-esteem and external locus of control. Every personality measure except neuroticism differed significantly in its association with SI and with ND. Of particular interest, while high levels of E predicted SI, among smokers, low levels of E predicted ND.

A lifetime history of six of the seven disorders examined was significantly associated with SI. The association was particularly strong (OR > 3) for panic disorder, problem drinking and alcohol dependence. By contrast, among smokers ND was significantly associated only with major depression, generalized anxiety disorder and alcohol problems. Three disorders differed significantly in their association with SI and ND, all because they were more strongly related to SI than to ND: panic disorder, problem drinking and alcohol dependence.

DISCUSSION

Heritability of smoking initiation and nicotine dependence

The concordance rates for SI in this sample (MZ twins 76% and DZ twins 61%) are within the range previously reported (Fisher, 1958; Friberg et al. 1959; Todd & Mason, 1959; Raaschou-Nielsen, 1960; Conterio & Chiarelli, 1962; Cederlof et al. 1977; Pedersen, 1981; Kaprio et al. 1982, 1984; Hannah et al. 1985; Carmelli et al. 1990; Heath et al. 1993; Kendler et al. 1993a; Boomsma et al. 1994; Edwards et al. 1995). The estimated heritability for SI in this sample (a² = 0.78–0.85) was toward the upper range of that reported in other population or volunteer twin samples from Sweden (Pedersen, 1981) (a² = 0.84), Australia (Heath et al. 1993) (a² = 0.77), Holland (Boomsma et al. 1994) (a² = 0.55), and the United States (Health et al. 1993; Kendler et al. 1993a) (a² = 0.48–0.54).

We are unaware of previous twin studies of ND. Our results suggest that in smokers the liability to ND is substantially influenced by genetic factors, with an estimated total heritability of ~ 72%. These findings are consistent with prior studies showing substantial heritability for variables that may indirectly reflect ND such as smoking persistence (Heath & Madden, 1995), heavy smoking (Kaprio et al. 1982) and number of cigarettes consumed (Carmelli et al. 1990).

The relationship between smoking initiation and nicotine dependence

A goal of this study was to clarify the aetiological relationship between SI and ND. By model fitting, we rejected the hypotheses that the risk factors for SI and ND were completely correlated or entirely independent. Several previous studies have examined issues that bear resemblance to our analyses with varying results ( Heath, 1990; Meyer et al. 1992; Heath & Martin, 1993). To our knowledge, our study is the first to address the relationship between SI and measured ND and suggests some risk factors were shared by SI and ND, while others impacted uniquely on ND.

To clarify this issue further, we examined the relationship between a range of potential risk factors and SI and ND. The resulting pattern was complex. While demographic and religious factors were considerably more strongly associated with SI, personality traits more strongly predicted ND. Neuroticism was an exception to this pattern, correlating strongly with both SI and ND. By contrast, while high levels of extroversion predicted SI, low levels among smokers predicted ND. While mastery, dispositional optimism and self-esteem were unrelated to SI, these traits were strongly and inversely related to ND. Our analysis of putative risk factors replicated our modelling results. Some risk factors (e.g. educational level, neuroticism, depression, alcoholism) were shared
both for SI and ND. Other risk factors, by contrast, substantially influenced only SI (e.g. religiosity) or only ND (e.g. low self-esteem).

Our results are generally consistent with prior evidence for a relationship between personality and smoking (Breslau et al. 1993; Gilbert & Gilbert, 1995; Heath et al. 1995). Consistent with prior work (Glassman et al. 1990; Breslau et al. 1991; Breslau, 1995), affective, anxiety and alcohol use disorders were all associated with a significantly elevated risk for SI. The association of psychiatric disorders with ND was more specific than that seen for SI – reaching statistical significance only for MD, GAD and alcohol use disorders.

Conclusion
Using cigarette smoking as a paradigmatic substance-use problem, our results suggest that the pathway to dependence is complex. Both genetic and sociocultural factors play a significant aetiological role at the stages of initiation and dependence. Moreover, the genetic and environmental risk factors that act at these discrete stages, while correlated, are not identical. While some familial risk factors influence dependence through their impact on initiation, other familial factors (perhaps reflecting genetic variation in drug metabolism and interaction with end-organ receptors) play a particularly important role only in the vulnerability to substance dependence.

Limitations
This study should be interpreted in the context of five potential methodological limitations. First, our sample was entirely female and we do not know whether a similar pattern would be found in males (Jarvis, 1994).

Secondly, despite the large initial sample, we only had 234 twin pairs concordant for SI. In a conventional twin study of ND, a sample size of ~400 pairs would be needed to detect genetic effects at the 5% level with 80% power given a true heritability of ~45% (Neale et al. 1994). However, this problem is exacerbated by the high correlation between the liabilities to SI and ND, resulting in quite limited power to estimate the unique sources of liability to ND. While we are confident of the existence of familial factors that are unique to ND, we have quite limited power to determine if those factors are genetic or familial–environmental. However, both our model-fitting to a modest degree and prior studies on proxy variables for ND suggest that the most plausible model is one which includes genetic risk factors unique to ND.

Thirdly, frequency of adult contact predicted twin resemblance for ND, a possible violation of the equal environment assumption. Our sample of twin pairs concordant for SI was too small to evaluate by model-fitting methods (Hettema et al. 1995). A cruder method – discarding high contact MZ and low contact DZ pairs until the zygosity groups were matched for contact frequency – produced correlations for ND that remained higher in MZ than DZ twins, although the differences were diminished. Such corrections, however, are almost undoubtedly conservative, as they assume that frequent contact ‘causes’ resemblance in ND. It may be more plausible that twins with similar smoking habits (or similar ‘correlated’ traits such as personality, education and income) choose to be in closer contact.

Fourthly, while most of the associations observed between putative demographic, religious, personality, and psychopathologic risk factors and SI, SP and ND may be causal, this is unlikely to be uniformly true. While smoking is unlikely to influence education or religiosity, attempts to quit smoking might cause major depression (Glassman et al. 1988). To address further this critical question, we utilized the twin structure of our data. Table 3 contains 20 statistically significant relationships between putative personality and psychopathology risk factors in a twin and her own SI or ND status. We next examined whether these risk factors in one twin predicted the smoking status of her cotwin. Of these 20 analyses, 12 were statistically significant at the 5% and three at the 10% level. Such a pattern is inconsistent with the hypothesis that these associations result largely from smoking causing the putative risk factors.

Finally, the correlation observed between the liabilities to SI and ND may have been so high because our definition of SI – regular smoking for a month – was so stringent. Of those who first start to smoke, a high vulnerability to ND may be required to progress to 1 month of regular smoking. Therefore, we re-fitted our model illustrated in Fig. 1 with a substantially lower threshold for SI: the consumption, within
Twin study of smoking initiation and nicotine dependence

...a single week, of two or more cigarettes. The path from SI to ND (b) was estimated at +0.81, slightly higher than the +0.77 obtained with our original definition of SI. It is unlikely that the high correlation observed between the liabilities to SI and ND is an artefact resulting from our stringent definition of SI.

This work was supported by grants MH-40828, MH-49492, MH-01277, AA-09095 and DA-10228 from the United States National Institutes of Health. The Virginia Twin Registry, established and maintained by Dr W. Nance MD, and Dr L. Corey, is supported by the United States National Institutes of Health grants HD-26746 and NS-31564. Dr Charles MacLean provided advice on the use of the continuation ratio.

REFERENCES


