Behav Genet (2006) 36:795–808
DOI 10.1007/s10519-006-9085-4

ORIGINAL PAPER

Genetic and Cultural Transmission of Smoking Initiation: An Extended Twin Kinship Model

Hermine H. Maes · Michael C. Neale · Kenneth S. Kendler · Nicholas G. Martin · Andrew C. Heath · Lindon J. Eaves

Received: 13 September 2005 / Accepted: 18 May 2006 / Published online: 30 June 2006

Abstract
Background Considerable evidence from twin and adoption studies indicates that genetic and shared environmental factors play a significant role in the initiation of smoking behavior. Although twin and adoption designs are powerful to detect genetic and environmental influences, they do not provide information on the processes of assortative mating and parent–offspring transmission and their contribution to the variability explained by genetic and/or environmental factors.

Methods We examined the role of genetic and environmental factors for smoking initiation using an extended kinship design. This design allows the simultaneous testing of additive and non-additive genetic, shared and individual-specific environmental factors, as well as sex differences in the expression of genes and environment in the presence of assortative mating and combined genetic and cultural transmission. A dichotomous lifetime smoking measure was obtained from twins and relatives in the Virginia 30,000 sample.

Results Results demonstrate that both genetic and environmental factors play a significant role in the liability to smoking initiation. Major influences on individual differences appeared to be additive genetic and unique environmental effects, with smaller contributions from assortative mating, shared sibling environment, twin environment, cultural transmission and resulting genotype–environment covariance. The finding of negative cultural transmission without dominance led us to investigate more closely two possible mechanisms for the lower parent–offspring correlations compared to the sibling and DZ twin correlations in subsets of the data: (i) age × gene interaction, and (ii) social homogamy. Neither mechanism provided a significantly better explanation of the data, although age regression was significant.

Conclusions This study showed significant heritability, partly due to assortment, and significant effects of primarily non-parental shared environment on smoking initiation.

Keywords Smoking initiation · Extended twin kinship design · Genetics · Assortment · Cultural transmission

Edited by Stacey Cherny

H. H. Maes (✉) · M. C. Neale · K. S. Kendler · L. J. Eaves
Department of Human Genetics, Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, 980003, 23298-0003 Richmond, VA, USA
e-mail: hmaes@vcu.edu

H. H. Maes
Massey Cancer Center, Virginia Commonwealth University, Richmond, VA, USA

H. H. Maes
Faculty of Kinesiology and Rehabilitation Sciences, Katholieke Universiteit Leuven, Leuven, Belgium

M. C. Neale · K. S. Kendler · L. J. Eaves
Department of Psychiatry, Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA

N. G. Martin
Queensland Institute of Medical Research, Brisbane, Australia

A. C. Heath
Department of Psychiatry, Washington University School of Medicine, St. Louis, MI, USA

Introduction

Smoking is a serious public health problem. Briefly, tobacco is the second major cause of death in the world. Half of the 1.3 million people that smoke today will eventually be killed by tobacco (World Health Organization 2005). In the US, cigarettes are estimated to be responsible for 30% of all cancer deaths (87% of lung cancer deaths) and 21% of deaths from cardiovascular disease (Ries et al. 2005). Smoking harms nearly every organ in the body, causing many diseases and reducing health in general (CDC 2004). The economic costs of tobacco use are equally devastating. A 1994 report estimated that the use of tobacco resulted in an annual global net loss of US$ 200,000 million, a third of this loss being in developing countries (World Health Organization 2005).

Considerable evidence exists that genetic and environmental factors play a significant role in the initiation of smoking behavior. This evidence primarily stems from twin and adoption studies. Here, we summarize twin studies that have examined lifetime or current use of tobacco products (which we will term smoking initiation or SI). There are more than 15 published twin studies of SI, originating from seven different countries, with the majority reporting on adult samples. Detailed reviews of this literature have been published by Sullivan and Kendler (1997), Heath et al. (1998), and Li (2003). Estimates of the heritability (h^2) of SI were generally high, with most values falling between 40% and 70% (median = 57%). The unweighted mean (±SD) estimate of h^2 for the 37 total samples was 0.53±.15 and did not differ appreciably in the adult and adolescent subsets. Estimates of the proportion of variance in liability due to shared environmental effects (c^2) were more variable, with most ranging from 0% to 50%. The unweighted mean (±SD) estimate of c^2 for the 31 adult samples (0.22±0.18) was appreciably lower than that estimated from the 6 adolescent samples (0.41±0.20) (for all samples 0.27±.20). Shared environmental factors, particularly peer group effects, may be more important for SI in adolescence than in adulthood. The unweighted mean estimate for individual-specific environmental effects (e^2) was 0.18 (±.12). One of the most striking aspects of this literature is the consistent evidence for moderate to high heritability estimates across a wide range of ages, countries and gender (Madden et al. 2004).

Five studies have examined SI in twins reared apart, a combination of the twin and adoption study designs. Concordance for smoking in the 147 MZ pairs reported in these studies was high (75%) and only slightly lower than the concordance rate reported from studies of MZ twins reared together (84%) (see Sullivan and Kendler 1997; Kendler et al. 2000). The only available standard adoption study (Eaves and Eysenck 1980) reports substantial correlations for smoking status between biological parents and offspring, but not between adoptive parents and their adoptive offspring. Similarly, correlations between biological siblings (either nontwin siblings or DZ twins) were significantly different from zero, while those between adoptive siblings were not. In aggregate, the reared-apart twin studies and adoption studies support the hypothesis of strong genetic influences on SI.

Although the twin and adoption designs are powerful to detect genetic and environmental influences, they do not generally provide information on the processes of assortative mating and parent–offspring transmission and their contribution to the variability explained by genetic and/or environmental factors. The addition of parents to the twin design may inform about the mechanism of assortment (through the marital correlation) and about inter-generational transmission (through the parent–offspring correlation). To our knowledge, only one paper used a twin-parent model for smoking to estimate the degree of assortment and cultural transmission (Boomsma et al. 1994). They found that the association between spouses for 'ever smoked' was rather low (.18) and that the parent–offspring resemblance could be accounted for completely by their genetic relatedness. Adding more first and second degree relationships to the twin-parent design further allows us to evaluate the consistency of these parameters over a large range of relationships.

In this paper, we will examine the role of genetic and environmental factors for smoking initiation using an extended kinship design. This design allows the simultaneous testing of additive and non-additive genetic, shared and individual-specific environmental factors, as well as sex differences in the expression of genes and environment in the presence of assortative mating and combined genetic and cultural transmission. First, we estimated the correlations between relatives and consider their overall pattern across the different types of relative. Second, we fit a model to the data for the purpose of formal hypothesis testing. Third, we test several assumptions of the model, such as the consistency of estimates across age, and the mechanism of assortment and cultural transmission with subsets of the data. Finally, we refit the full model, allowing for a change in prevalence with age.

Materials and Methods

The Virginia 30,000

The Virginia 30,000 sample contains data from 14,763 twins, ascertained from two sources (Eaves et al. 1999; Truett et al. 1994). Public birth records and other public records in the Commonwealth of Virginia were used to obtain current address information for twins born in Virginia between 1915 and 1971, with questionnaires mailed to twins who had returned at least one questionnaire in previous surveys. A second group of twins was identified through their response to a letter published in the newsletter of the American Association of Retired Persons (AARP, 9476 individuals). Twins participating in the study were mailed a 16 page ''Health and Lifestyles'' questionnaire, and were asked to supply the names and addresses of their spouses, siblings, parents and children for the follow-up study of relatives of twins. Completed questionnaires were obtained from 69.8% of twins invited to participate in the study, which was carried out between 1986 and 1989.

The original twin questionnaire was modified slightly to provide two additional forms, one appropriate for the parents of twins and another for the spouses, children and siblings of twins. Modifications affected only those aspects of the questionnaire related to twinning. The response rate from relatives (44.7%) was much lower than that from the twins. Of the complete sample of 28,492 individuals (from 8567 extended kinships) with valid smoking data, 58% were female, with 50% of respondents under 50 years of age. Table 1 breaks down the sample sizes by type of relative and sex, as well as by zygosity and ascertainment sample for the twins only.

Zygosity Determination

Zygosity of twins was determined on the basis of responses to standard questions about similarity and the degree to which others confused them. This method has been shown to give at least 95% agreement with diagnosis based on extensive blood typing (Eaves et al. 1989; Ooki et al. 1990).

Measures

In all questionnaires mailed to twins and their relatives, self-report data on smoking were obtained from three items. Respondents were asked to indicate the number corresponding to the frequency which best described their smoking habits during your lifetime. The four possible response values were: 'never smoked', 'used to smoke but gave it up', 'smoked on and off', 'smoked most of your life'. Smoking quantity was measured as the number which expressed their best estimate of the DAILY cigarette consumption with six response categories: 'never', '1–5 per day', '5–10 per day', '11–20 per day', '21–40 per day', '>40 per day'. Age of onset was recorded as the age at which they started smoking. Based on these three variables, we created a dichotomous variable reflecting whether they had ever smoked or not. If they responded ''never smoked'' to the smoking frequency question and ''never'' to the smoking quantity question and did not report an age of onset for smoking, they were coded zero on the dichotomous smoking variable. If on the other hand, they reported any of the other three response categories for smoking frequency OR any of the other five categories for smoking quantity OR an age of onset, they were coded one. Responses were consistent across the three variables for 86% of the sample. Eleven percentage of the sample was coded a smoker based on two out of three variables. In less than 1% of the sample was someone coded a smoker on the basis of only one of these three variables. 732 respondents (1%) were assigned missing values for the dichotomous smoking variable.

Table 1 Sample sizes by type of relative and sex (top panel), and by zygosity and ascertainment sample for the twins only (bottom panel)

	Twins	Parents	Siblings	Spouses	Children	Total
Male	5,322	914	1,260	2,515	1,890	11,901
Female	9,434	1,448	1,924	1,875	1,910	16,591
Total	14,756	2,362	3,184	4,390	3,800	28,492
	MZM	DZM	MZF	DZF	DZO	Total
VTR	610	718	960	810	1,294	4,392
AARP	1,014	500	2,986	1,726	1,498	7,724
Total VA30k	1,626	1,220	3,948	2,542	2,792	12,128[a]

VTR:Virginia Twin Registry subsample; AARP:American Association for Retired Persons subsample

[a]For 2628 twins, zygosity information was either not available or inconclusive

Statistical Methods

Structural modeling of the data was undertaken using methods described in Eaves et al. (1999) and Truett et al. (1994), which assess the contributions of additive and dominant genetic effects in the presence of effects such as vertical cultural inheritance, phenotypic assortative mating, shared twin and sibling environments and within-family environment. Phenotypic assortment occurs when mate selection is based at least partly on the trait being studied, and generates a correlation between the observed phenotypes of spouses. Vertical cultural inheritance is the transmission of traits from parent to child by non-genetic means, and reflects the impact of the parent's phenotype on the environmental of their children.[1] Between-family environmental effects make family members relatively more similar, which include sibling and twin effects. Sibling environments are those environmental factors shared between all types of offspring. A special twin environment is an additional correlation between the environment of twins (in addition to the sibling environment) which makes both MZ and DZ twins more alike than ordinary siblings even in the absence of genetic effects (Neale and Cardon 1992). While all these sources of common environment contribute to variation among individuals regardless of relationship, they differ in their effect on the covariation between types of relatives. The contribution of the genetic and environmental factors may be dependent upon sex, both in their magnitude and nature. Figure 1 presents a path diagram of the extended twin design model (ET model). Note that only two generations are shown as all the model parameters can be depicted with drawing just an opposite-sex pair of twins and their parents.

In its early applications, the ET model was fit to correlations between relatives after the data had been corrected for the linear and quadratic effects of age, sex, twin status (twin versus other relative), source of ascertainment (Virginian birth records versus AARP) and interactions between these effects (Eaves et al. 1999). After the implementation in the statistical modeling package Mx (Neale et al. 2003) the ET model was fit to the raw ordinal data to obtain maximum likelihood estimates of the model parameters. Given the three generational nature of the some of the data, we opted for a binary approach, as ordinal data with more categories would be much more computationally intensive. Also, this approach allowed us to add covariates such as age to the model, thereby addressing one of the criticisms of the analysis of kinship data. A detailed description of the specification of the ET model is given in Maes et al. (1999).

[1] This model, which assumes that assortment and cultural transmission are based on the measured phenotype, is only one of the possible mechanisms for family resemblance (Fulker 1988; Heath and Eaves 1985).

Results

Response Frequencies

Responses frequencies for the smoking questionnaire items are listed in Table 2. This cohort demonstrated a marked difference between smoking behaviors of men and women, with greater frequency and quantity among men. While a decrease in prevalence of smoking behaviors over three generations (fathers versus twins/brothers/husbands versus sons) was apparent for males, both the frequency and quantity of smoking are comparable across three generations in females. Regression analyses (not shown) indicated significant regressions for sample (Virginia Twin Registry versus AARP) and sex and for the interactions between sex × age, sex × age^2, age × sample and sex × sample. It should be noted 'sample' is highly correlated with age, because most of the AARP related individuals are older than 50, whilst the majority of the Virginia Twin Registry sample is under 50.

Maximum Likelihood Estimation of Correlations

Due to advances in computational speed and efficiency, it is now feasible to use maximum likelihood methods to estimate tetrachoric correlations and model genetic and environmental effects in pedigrees of this complexity, allowing us to obtain unbiased estimates of all parameters. Table 3 lists the maximum likelihood estimates of the tetrachoric correlations. When the correlations are subdivided by sex of the pairs, a total of 88 different estimates are obtained. The precision of these correlations is low when the sample size is small which is the case for the more distant relatives. However, when we constrain correlations to be the same across sex of the pairs within categories with the same genetic and environmental relatedness, the pattern of correlations becomes much more stable (see Fig. 2). Furthermore, none of the gender homogeneity tests for correlations was significant, except for DZ opposite sex correlation compared to the DZ same sex correlations.

The correlations for relatives of different degree of relatedness—both genetic and environmental—'tell

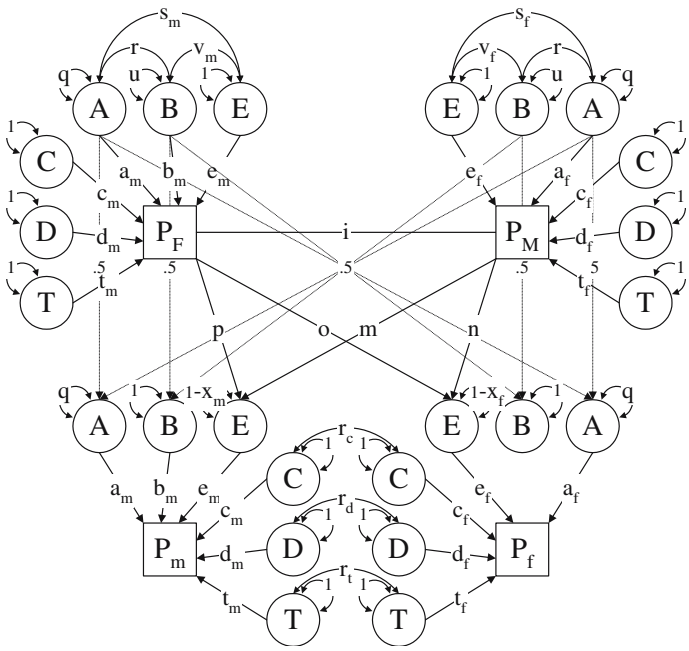

Fig. 1 Full extended family resemblance model for opposite-sex DZ twins and their parents. Path coefficients are the same in both generations, and gene–gene and gene–environment correlations occur in both generations a_f = gender-common additive genes–females; a_m = gender-common additive genes–males; b_m = male-specific additive genes–males; r = induced correlation between gender-common and male-specific additive genetic effects; d_f = non-additive genes–females; d_m = non-additive genes–males; r_d = correlation between male and female non-additive genetic effects; c_f = common environment–females; c_m = common environment–males; r_c = correlation between male and female common environment; t_f = special twin environment-females; t_m = special twin environment–males; r_t = correlation between male and female special twin environmental effects; n = maternal cultural transmission-females; m = maternal cultural transmission-males; o = paternal cultural transmission-females; p = paternal cultural transmission-males; e_f = specific environment parameter-females; e_m = specific environment parameter-males; i = assortative mating parameter; s_f = correlation between gender-common additive genetic effects and environment-females; s_m = correlation between gender-common additive genetic effects and environment-males; v_f = correlation between male-specific additive genetic effects and environment-females; v_m = correlation between male-specific additive genetic effects and environment-males

the story' about the genetic and environmental contributions to the phenotype. If *additive genetic factors* contribute to the variability, correlations are expected to decrease with a decreasing genetic relatedness, e.g., MZ twins should correlate higher than first-degree relatives (DZ twins, siblings, parent–child pairs) who share on average half of their genes, followed by second-degree relatives (grandparents, half-siblings, avuncular pairs) who share on average 1/4 of their genes, and more distant relatives, such as cousins, who share 1/8 of their genes. From a genetic perspective, avuncular MZ pairs (nieces and nephews with their aunts/uncles who are the identical co-twins of their mother/father) are similar to parent–child pairs. Cousins through MZ twins are genetically related as half-siblings. Such a pattern of decreasing correlations with decreasing genetic relatedness is observed for the tetrachoric correlations for SI in the genetically related pairs, except for somewhat lower parent–child and avuncular correlations. Significant *dominance* contributions would be reflected in DZ and sibling correlations being less than half the MZ correlations (expectation is $1/4\ rMZ \leq rDZ \leq 1/2\ rMZ$) with zero correlations for other pairs of relatives. There is no clear evidence for this pattern in the SI correlations.

The contribution of *shared environmental factors* should increase similarity between people living together or having grown up in the same home. This would affect all first-degree relatives and MZ twins. With different types of first-degree relatives, this source of variance can be further partitioned in factors shared with parents—typically referred to as *vertical cultural transmission*—and non-parental environmental factors. The latter are environmental factors that siblings may have in common, such as school environment, peers, friends and could explain greater similarity in siblings than parent–offspring pairs. Given the availability of twins, we can also assess the presence of a *special twin environment*

Table 2 Response frequencies (*N* and %) of self-reported smoking behavior in the Virginia 30,000 sample

		Fathers	Mothers	MaleTwins	FemaleTwins	Brothers	Sisters	Husbands	Wives	Sons	Daughters
Smoking frequency											
Never Smokers	*N*	202	778	1,991	5,319	460	1,076	772	1,054	960	1,593
	%	22	55	38	57	37	57	31	57	51	55
Past smokers	*N*	461	324	1,894	2,012	467	426	1,190	492	506	702
	%	51	23	36	22	37	22	48	27	27	24
Casual smokers	*N*	51	85	427	722	112	162	130	106	157	257
	%	6	6	8	8	9	9	5	6	8	9
Regular smokers	*N*	187	227	940	1,203	212	237	400	197	259	345
	%	21	16	18	13	17	12	16	11	14	12
Smoking quantity											
No cigarettes	*N*	230	751	2,097	5,232	459	1,058	833	1,046	969	1,596
	%	36	62	41	58	49	63	44	63	64	62
1–5/day	*N*	111	149	610	1,073	154	223	319	235	196	356
	%	18	12	12	12	16	13	17	14	13	14
5–10/day	*N*	83	97	388	667	99	126	194	125	89	196
	%	13	8	8	7	10	8	10	8	6	8
11–20/day	*N*	203	213	1,050	1,267	223	275	550	248	259	428
	%	32	18	21	14	24	16	29	15	17	17
21–40/day	*N*	8	6	757	663	5	5	7	2	1	1
	%	1		15	7	1					
>40/day	*N*	2	1	186	105	1					
	%			4	1						
Age of onset of smoking											
No onset	*N*	185	730	1,915	5,101	433	1,045	750	1,029	942	1,575
	%	22	53	37	57	36	56	31	57	51	55
Onset	*N*	660	635	3,220	3,899	777	810	1,683	789	909	1,301
	%	78	47	63	43	64	44	69	43	49	45
Smoking (dichotomous)											
Never smokers	*N*	185	730	1,915	5,101	433	1,045	750	1,029	942	1,575
	%	21	53	37	56	35	56	30	56	50	55
Smokers	*N*	703	649	3,292	4,001	799	832	1,736	803	926	1,314
	%	79	47	63	44	65	44	70	44	50	45

from the additional similarity in twins compared to siblings. Although different mechanisms of intergenerational transmission are possible, transmission based on the phenotype of the parent seems reasonable. If this type of cultural transmission exists for a trait which also has a genetic component, it will result in genotype–environment covariance. Positive cultural transmission (and resulting GE correlation) will increase the similarity of siblings and parent–child pairs relative to MZ twins. The results for SI suggest non-parental shared environmental factors and likely some contribution of special twin environment. There is no evidence for cultural transmission; on the contrary, the pattern of correlations might be more consistent with negative cultural transmission because parent–offspring correlations are less (rather than greater) than might be expected from genetic factors alone.

Another source of similarity which may have both genetic and environmental implications is non-random mating. Again, the effects of *assortative mating* depend on the mechanism by which people choose their spouses. One possibility is that selection is based on the phenotype of the partner (phenotypic assortment). This process generates correlations between the sources of variance in one spouse with those of the other spouse, and thus increases the genetic and/or environmental covariance between parent–offspring and sibling pairs. Phenotypic assortment increases correlations through marriage (e.g. parents with in-laws of their children) such that they will be a combination of the spousal correlation and the correlation between the spouse and the other relative (e.g. parent–child correlation * spousal correlation). If genetic factors contribute significantly to the trait, the correlations through marriage will be a function of the genetic relatedness of the pairings (e.g. the correlation between a spouse and the MZ sister of his wife will be expected to be greater than that between a spouse and his wife's DZ sister). The spousal correlation for SI is highly significant, suggesting some form of assortment. The pattern of correlations through marriage observed for smoking behavior is consistent with a genetic contribution to SI. While the spousal correlation for smoking could arise

Table 3 Maximum likelihood correlations for smoking behavior in the VA30,000 sample

	Male–Male[a]	Female–Female	Male–Female	Female–Male	Homogeneity
Nuclear families					
MZ twins	0.834	0.794	–	–	0.806
DZ twins	0.584	0.533	0.333	–	0.462[b]
Siblings	0.361	0.358	0.286	–	0.317
Parent–Child	0.185	0.212	0.212	0.159	0.189
Grandparents via					
Father	0.139	0.263	0.259	0.322	0.228
Mother	0.376	0.123	0.585	0.308	
Avuncular via					
Father's MZ co-twin	0.243	–	0.193	–	0.165
Mother's MZ co-twin	–	0.172	–	0.122	
Father's DZ co-twin	−0.182	0.097	0.070	−0.081	0.063
Mother's DZ co-twin	−0.010	0.174	−0.090	0.090	
Father's sibling	−0.048	0.189	−0.042	−0.114	0.048
Mother's sibling	0.055	0.012	0.114	0.067	
Cousins via					
MZ male twins	0.296	0.221	0.280	–	0.220
MZ female twins	0.212	0.213	0.173	–	
DZ male twins	0.013	0.020	−0.166	–	
DZ female twins	−0.019	0.308	0.055	–	
DZ male-female twins	0.536	0.072	0.108	0.184	0.120
Spouses	–	–	0.377	–	0.373
Spouse of twin with					
MZ co-twin	–	–	0.306	0.238	0.274
DZ co-twin	0.205	0.169	0.244	0.216	0.209
Sibling of twin	0.176	0.258	0.130	0.077	0.167
Parent of twin	0.053	0.172	0.067	−0.026	0.079
Spouse of MZ co-twin	0.286	0.514	–	–	0.397
Spouse of DZ co-twin	0.177	0.370	0.046	–	0.176
Affine avuncular via					
Father's MZ co-twin	–	0.065	–	0.049	0.100
Mother's MZ co-twin	0.063	–	0.204	–	
Father's DZ co-twin	0.190	0.219	0.208	−0.015	
Mother's DZ co-twin	0.198	0.218	0.222	−0.034	0.132

[a]First sex refers the older relative

[b]Significant sex difference

from spousal interaction rather than assortment, it was not significantly mediated by duration of marriage.

Maximum Likelihood Estimation of Genetic and Environmental Contributions

Maximum likelihood estimates of the genetic and environmental parameters under the ET model and the derived proportions of variance for the genetic and environmental effects from the analysis of individual observations are shown in Table 4. Additive genetic effects accounted for 60% of the variance in smoking in males and 64% in females. These proportions include the effects due to assortative mating (about 10%), consistent with the highly significant spousal correlation r=.38. The contribution of genetic dominance was negligible. The shared environmental effects on smoking arise from non-parental sources, special twin environment and cultural transmission. In males, these sources explained 23, 14 and 7% of the variance. The corresponding proportions for females were 14, 15 and 3%. Genotype-environment covariance was estimated to be negative for males and females, resulting in a negative contribution of this source of variance, respectively −21% for males and −16% for females. Individual specific environmental factors made up the remainder of the variance (17% in males, 20% in females). The estimates from the full model provide no evidence for sex-specific genetic effects. However, the correlation between the non-parental and twin shared environments in males and females are .01 and .49 respectively, suggesting that partly different shared environmental factors account for similarity in smoking behavior in males and females.

The 95% confidence intervals could be obtained from Mx using the method of Neale and Miller (1997). However, given the large number of estimated parameters and the ordinal data input, estimating

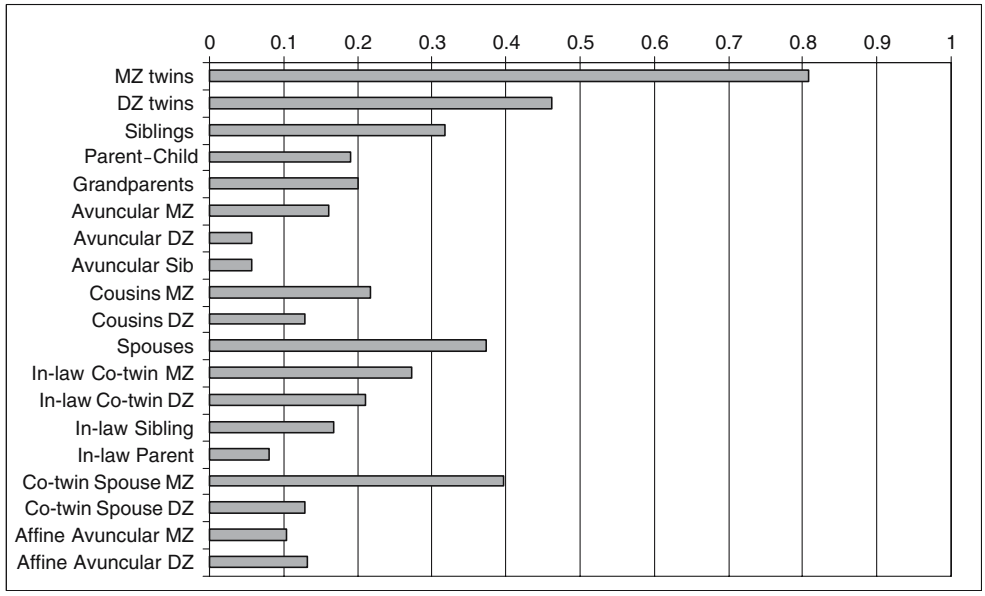

Fig. 2 Maximum likelihood correlations for smoking behavior in the VA30,000, grouped by degree of genetic and environmental similarity, constrained to be equal across sex

confidence intervals would require an impractical amount of computer time. Therefore, we opted to fit a range of submodels which allows us to test the significance of individual parameters or a group of parameters simultaneously. Results of the likelihood ratio tests are presented in Table 5. The first five tests assess whether different genes or different environmental factors operate in males and females (referred to as non-scalar sex limitation). There is evidence of some form of non-scalar sex limitation, most likely for the non-parental sibling shared environmental factors. This is not surprising, since the opposite-sex twin and sibling correlations are considerably lower than their respective same-sex correlations. Some indication also exists for sex-dependent cultural transmission with greater environmental contribution from parents to female than to male offspring. The second five tests address scalar sex limitation, testing whether the sources of variance contribute equally to males and females. With respect to the magnitude of the genetic and environmental factors, no significant differences were observed. The third group of seven tests consecutively evaluates the significance of each of the sources of variance. Assortment, non-parental shared, and special twin environment are all highly significant. The contribution of additive genetic factors was borderline significant (P=.059), while that of cultural transmission approached significance (P=.10). Given the estimates

Table 4 Parameter estimates and variance components from the ET model for smoking behavior in the VA 30,000 sample

	Before age regression					After age regression				
	Parameter estimates			Variance components		Parameter estimates			Variance components	
	male	female	mf	male	female	male	female	mf	male	female
Assortative mating			.382	.099	.107			.385	.129	.118
Common additive genetic	.699	.727		.489	.529	.763	.732		.582	.535
Male-specific genetic	.118		.019	.014		.021		.003	.000	
Dominance	.000	.004	1.000	.000	.000	.040	.000	1.000	.002	.000
Unique environment	.491	.487		.168	.204	.512	.502		.177	.207
Shared environment	−.481	.368	.001	.231	.136	−.489	.381	.094	.240	.145
Twin environment	.369	.386	.485	.136	.149	.238	.379	.999	.057	.144
Cultural transmission from father	−.093	−.149				−.192	−.202		.085	.045
Cultural transmission from mother	−.508	−.289		.073	.033	−.468	−.301			
Genotype–Environment correlation a	−.308	−.222		−.211	−.157	−.347	−.264		−.271	−.194
Genotype–Environment correlation b	−.019	−.017		.002		−.004	−.004		.000	

mf: male–female; genotype–environment correlation a: between common additive genetic factors and environment; genotype–environment correlation b: between male-specific genetic factors and environment

Table 5 Model fitting results for fitting the ET model and submodels to smoking behavior in the VA30,000 sample

	Full model				No dominance model			
	−2LL	χ^2	df	*P*	−2LL	χ^2	df	*P*
Full ET model	30496.796		24207		30496.858		24210	
No male-specific genes	30496.916	0.1	1	0.729	30496.917	0.1	1	0.810
Same shared environment in two sexes	30510.135	13.3	1	0.000	30510.033	13.2	1	0.000
Same twin environment in two sexes	30498.545	1.7	1	0.186	30498.326	1.5	1	0.226
Same dominance in two sexes	30496.858	0.1	1	0.803	NA			
Cultural transmission no sex diff	30506.684	9.9	3	0.020	30506.779	9.9	3	0.019
Same effect of environment in two sexes	30496.860	0.1	1	0.800	30496.801	−0.1	1	
Same effect of genes in two sexes	30496.865	0.1	1	0.793	30496.833	0.0	1	
Same effect of dominance in two sexes	30496.858	0.1	1	0.803	NA			
Same effect of shared env in two sexes	30498.563	1.8	1	0.184	30498.647	1.8	1	0.181
Same effect of twin env in two sexes	30496.925	0.1	1	0.719	30496.914	0.1	1	0.815
No non-parental shared environment	30524.118	27.3	3	0.000	30524.152	27.3	3	0.000
No special twin environment	30515.493	18.7	3	0.000	30515.506	18.6	3	0.000
No cultural transmission	30504.383	7.6	4	0.108	30510.268	13.4	4	0.009
No shared environment	30569.151	72.4	10	0.000	30678.625	181.8	10	0.000
No additive genetic effects	30504.255	7.5	3	0.059	30576.907	80.0	3	0.000
No dominance effects	30496.858	0.1	3	0.996	NA			
No assortment	30769.531	272.7	1	0.000	30743.298	246.4	1	0.000

−2LL: minus twice the log-likelihood

for dominance were very close to zero, we repeated the set of likelihood ratio tests, fixing the dominance parameters for males and females to zero and their correlation to one. Results for the reduced model were similar to those for the full model, except that the contributions of additive genes and cultural transmission were significant.

The finding of negative cultural transmission led us to investigate more closely the possible mechanism for the lower parent–offspring correlations compared to the sibling and DZ twin correlations. At least two explanations seem plausible: (i) age × gene interaction, i.e. the genes accounting for liability to smoking are not the same across the age range of the subjects and/or do not account for the same amount of the variance; (ii) environment to environment transmission from parents to offspring (social homogamy) and environmental instead of phenotypic assortment. A feature of the social homogamy model is the assumption that the observed spousal correlation is entirely due to their assortment for the cultural environment. Given the complexity of the extended kinship model, and therefore its limited flexibility, we evaluated the alternative explanations by reducing the model to manageable subsets. We used the twin and sibling data to test for age × gene interaction. The social homogamy hypothesis was tested using the twin-parent and twin-spouse designs. In addition to interaction with age, we also allowed the thresholds to vary as a function of age (cohort).

Age Regression and Moderation

Thresholds and twin correlations were estimated by maximum likelihood for each of the five zygosity groups. Age was allowed to moderate both thresholds and the twin correlation. The thresholds and age regression could be equated across birth order and zygosity without loss of fit, but not across sex. Both linear and quadratic effects of age on the prevalence of smoking were highly significant for males ($P <= .001$) but not for females ($P = .900$), and were significantly different between males and females ($P \leq .001$). Similarly, the twin correlations changed with age for male twins ($P \leq .001$ and $P = .001$ respectively for MZ and DZ twins) but not for female ($P = .218$ and .186 for MZ and DZ) or opposite sex twins ($P = .531$), such that the estimated correlations increased from .77 at age 20 to .87 at age 90 for MZ male twins and from .38 to .71 for DZ male twins. When partitioning the twin correlations into additive, shared and specific environmental sources of variance, both the genetic and shared environmental variance decreased with age in females, but neither moderation was significant. In contrast, increases in the genetic and shared environmental variance with age were observed for males. Although dropping the moderation of the additive variance or that of the shared environmental variance did not result in a significant decrease of fit, both moderation effects could not be dropped simultaneously. The magnitude of the contribution of genes and environment on the liability

to smoking was thus similar across the age range for females but not for males. Since we did not have repeated measures of the twins' smoking behavior, we used data from sibling pairs to test whether the same genetic (and environmental) factors accounted for liability to smoking at different ages. This was modeled by allowing the sibling correlations to vary as a function of increasing age difference, using an exponential decay function exp($-y$ * abs(ages1 − ages2)) where y is the decay parameter, and ages1 and ages2 the respective ages for sibling 1 and 2, and abs the absolute value). Results indicated that the sibling correlations did not vary significantly as a function of age difference for same-sex or opposite sex pairs (P=.882). Similar to the results from the twin analysis, the regression on age was significant for males (P=.000) but not females (P=.087). Our results thus suggest that the effect of the genetic and environmental factors is the same across the age range.

Social Homogamy Versus Phenotypic Assortment and Cultural Transmission

Rather than modeling assortment as a function of the phenotype and transmission as a direct effect from the phenotypes of the parents on the shared environment in the offspring (P to C transmission), alternative mechanisms exist. One such alternative is that shared social background factors account for the spousal correlation and that these social background factors are transmitted to the offspring generation (C to C transmission). When just data on twins and their parents were selected, the social homogamy model fitted almost as well as the phenotypic assortment model. However, the social homogamy model resulted in unreasonable estimates for the spousal correlation and cultural transmission. We fitted both twin-parent models allowing either different genes(free r_g) or allowing different shared environmental factors (free r_c) in males and females. In all four models, assortment, modeled either as phenotypic assortment or social homogamy, was highly significant, as were genetic factors and the total of shared environmental factors. C to C transmission was not significant. P to C transmission was significant only in the phenotypic assortment model allowing for different shared environmental factors for males and females, where it explained 3% of the variance in female children and 10% in male children. The data thus provide more support for the hypothesis that parent–offspring similarity is due to phenotypic cultural transmission than to environment to environment transmission.

Data from twins and their spouses permit the simultaneous evaluation of alternative mechanisms of assortment, under the assumption of no genotype × environment covariance. Under phenotypic assortment, the spousal correlation is expected to be higher than the correlation between a twin and the spouse of his/her cotwin (cotwin-spouse correlation) which is expected to be higher than the correlation between the respective spouses of the twin pair (spouse1–spouse2 correlation). Equal spousal and cotwin-spouse correlations are consistent with social homogamy. Although both types of assortment appeared to be significant—consistent with the twin-parent analyses—a model with only phenotypic assortment fitted better than one with only social homogamy.

Given the results from these separate analyses in subsets of the data, we considered phenotypic assortment and phenotypic cultural transmission to be reasonable mechanisms for assortment and environmental transmission from parents to offspring respectively. Furthermore, it was deemed important to extend the extended twin kinship model with age as a moderator of the thresholds, thus accounting for the significant decrease of the smoking prevalence in males. To make this feasible, we only included a maximum of two male and female siblings and/or children of the twins in the analyses. This reduced the total sample size by only 267 individuals. The results obtained when including age regression were very similar to those without, even though including the age regression significantly improved the fit of the model. The major difference was a slight increase in the additive genetic variance components and the shared environmental component for females only, with a resulting increase in negative genotype × environment covariance (see Table 5 and Fig. 3).

Discussion

Results for the extended twin kinship analyses demonstrate that both genetic and environmental factors play a significant role in the liability to smoking initiation. Major influences on individual differences include additive genetic and unique environmental effects, with smaller contributions from assortative mating, shared sibling environment, twin environment, cultural transmission and resulting genotype–environment covariance. Although several large scale twin studies have investigated the sources of variability in smoking liability, to our knowledge this is the first report to include kinships of twins, thereby allowing extended partitioning of the variance and tests of

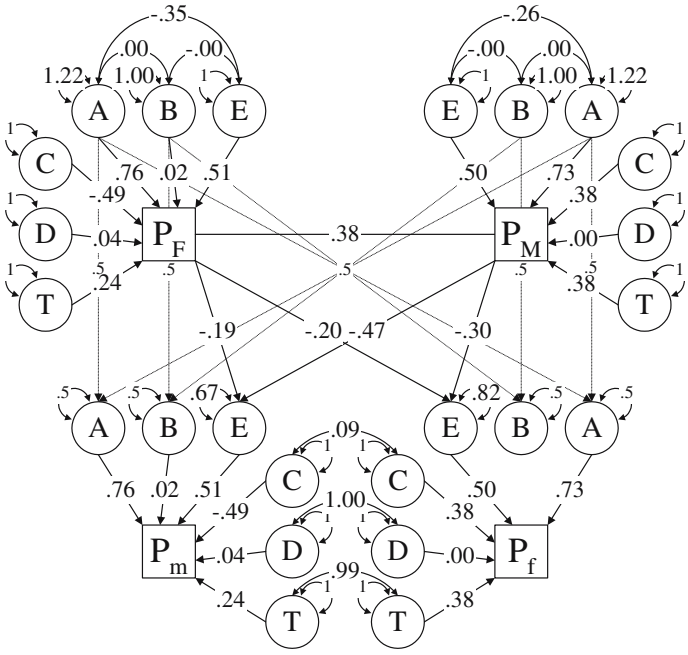

Fig. 3 Maximum likelihood estimates of parameters for the full extended family resemblance model for smoking behavior in the VA30,000

assumptions of the twin study. The overall heritability in the VA30,000 data was estimated to be 71% in males and 65% in females. These estimates are very close to the unweighted mean (56% for males, 50% for females) calculated from published reports on adult Scandinavian, Australian and US samples (Prescott et al. 2005). Note that the US samples are mostly overlapping with the VA30,000. With the availability of spousal information, we established that about 10% of the total variance was due to the genetic consequences of assortative mating. A possible reason for the slightly higher heritability estimate in the current analysis, is the presence of negative genotype × environment covariance, which amounted to −27% in males and −19% in females. This may also explain the somewhat higher estimate of the combined effects of all sources of shared environment (sibling, twin & cultural transmission), which total 38% in males and 33% in females, compared to the unweighted mean of 24–28% from published reports of adult samples. The estimates of the specific environmental variance are consistent across the current analysis (18–21%) and published reports (18%).

It is important to note that the estimates obtained here are not just based on twin data, but on a wide range of relatives with different degrees of genetic similarity and shared environments. Furthermore, the estimates result from taking the effects of sex, assortment, and genotype × environment covariance and age regression into account. The spousal correlation estimate of .38 is in line with published spousal correlations which range mostly from .18 to .43 based on US, Swedish and Dutch samples (Price and Vandenberg 1980; Boomsma et al. 1994). Regardless of the mechanism of assortment, given that it is substantial, the estimates of the genetic and environmental parameters will be biased, if assortment is not taken into account. In the current sample, about 15% of the genetic variance can be ascribed to the genetic consequences of assortment.

Although twin studies have consistently reported significant contributions of the shared environment to the liability of smoking, data on twins reared together cannot by themselves distinguish between the environmental effects shared with their cotwin, siblings, and peers versus those shared with their parents., nor the combined effects of additive genetic, dominance and shared environmental factors. In the current analyses, we observed significant contributions of the non-parental shared environment (30% in males, 29% in females) of which one fourth to one half could be accounted for by additional similarity between twins compared to other siblings. The latter could be related to the fact that adolescents typically socialize with people of similar age, which would in the case of twins include their cotwin. However, when we compared sibling correlations with differing degrees of age difference between them, we did not observe any significant decrease or increase in the correlations, suggesting that same age peers do not necessarily account for the significant twin environment. An alternative explanation may be the lingering effects of the

intra-uterine environment. Furthermore, it appears that the shared sibling (and possibly twin) environmental factors are different in males and females, indicated by the significantly lower opposite sex versus same sex correlations. This observation is in line with previous results. Boomsma et al. (1994) estimated the correlation between the shared environmental factors of males and females (r_c) at .66; Heath et al. (1993) reported an improved fit of the model when r_c was allowed to be less than one, but the effect did not reach statistical significance.

In addition to the non-parental shared environment, parental shared environmental factors (or cultural transmission) were also—albeit borderline—significant, but accounted for a relatively small proportion of the total variance (between 5% and 9%). Furthermore, the paths from parents to children's environment were estimated to be negative, suggesting that parents would have an inhibiting effects on their children's smoking behavior. These results are consistent with the only other available twin-parent data which also showed negative, but non-significant cultural transmission (Boomsma et al. 1994). The vast literature on parental smoking as a positive risk factor for adolescent smoking (Li et al. 2002; Peterson et al 2006; Shakib et al. 2003; Vitaro et al. 2004) and parental non-smoking or smoking cessation as a protective factor (Andersen 2004; Bricker 2005; den Exter Blokland et al. 2004) does not allow for the separation of the genetic effects of parents on their children from the environmental influences. Typically these studies report that children of smokers are more likely to smoke themselves and vice versa. This holds true in our data, where in households where both parents smoke, 56% of sons and 52% of daughters have ever smoked. In families where neither parent ever smoked, 36% of sons and 31% of daughters smoke. If one parent smokes but the other does not, smoking prevalence is 46% for male offspring and 40% for females if the father smokes; the corresponding numbers for mothers who smoke are 35% and 43%. The negative cultural transmission estimates might indicate a failure of the current model to appropriately account for the much lower parent–offspring than twin/sibling correlations. But the marginal significance of cultural transmission compared to the non-parental shared environmental sources of variance corresponds to the finding that adolescent smoking is more strongly associated with friends' and siblings' smoking than parents' smoking (de Vries et al. 2003; Rose et al. 2003; Simons-Morton et al. 2004). Vink et al. (2003a) noted that the relative risk to smoke is highest when having smoking friends, somewhat lower when having smoking younger/older siblings and lowest when having smoking parents. However, the risk to initiate smoking when having smoking friends was similar to that of having a same age smoking twin, with significantly higher risks given an MZ versus DZ smoking twin (Vink et al. 2003b). This points to the likelihood of genetic influences, as well as the importance of age difference or age similarity of the social influence. Although it is possible that children whose parents smoke are less likely to initiate smoking—similar to more teetotalers among offspring of alcoholics than non-alcoholics—it appears more plausible that the environmental impact of the parents on their offspring smoking behavior is limited and that the observed association is primarily accounted for by shared genes.

If genetic factors account significantly for the variability in liability to smoking, correlations are expected to decrease with decreasing degree of genetic relatedness of the pairings. This is observed for the MZ, DZ, sibling and other within-generation correlations, but parent–offspring and other across-generation correlations appear too low. Possible genetic explanations for the lower than expected parent–offspring than sibling correlations are genetic dominance or gene × age interaction. The results from fitting the full model showed no evidence for dominance. In effect, the dominance variance was estimated very close to zero. The alternative explanation of gene × age interaction implies that the genetic variance changes as a function of age and/or that different genes account for variability at different ages, sometimes called reduced genetic transmission. Although the study was cross-sectional, we were able to examine the change in genetic variance with age in two ways. First by comparing the twin correlations at different ages, and second by comparing siblings with different degree of age difference between them. The twin correlations further allow us to evaluate shared environment × age interaction. Although the magnitude of the genetic and shared environmental effects showed some variation as a function of age in male twins, there was no evidence of this in females. In addition, prevalence changed significantly with age in males but not females. There was little evidence that sibling correlations declined with age difference, in fact, while correlations slightly decreased with increasing age in males, they increased in females. Neither of these comparisons suggested a large impact of age on the genetic architecture of smoking initiation. These results are consistent with those reported by Madden et al. (2004) who reported that the additive genetic variance could be constrained across three age groups between age 18 and 46, as well as three countries (Australia, Sweden & Finland), in women and in men. However,

they did report a significant decrease in the impact of shared environmental factors with increasing age. This finding is in accordance with the appreciably higher estimates for shared environment in adolescent versus adult samples (Sullivan and Kendler 1997; Prescott et al. 2005) and adds to the evidence of greatest social influence of individuals close in age (Vink et al. 2003a, b). Finally, given the large age range of our sample, cohort and age effects may be confounded. Although prevalence of tobacco use has decreased in both males and females since the data were collected, the estimates of the contribution of genetic and environmental factors are consistent with estimates from more recently collected samples. We tested for age effects by evaluating both age regression and interaction. Significant age regression was found for males, but not females. Furthermore, there was little evidence for genotype and/or environment by age interaction in males or females. Given the prevalence of smoking decreased more rapidly in the cohorts captured in the Virginia 30,000 sample than since then, results are expected to be influenced to a limited extent. We hope to be able to replicate these results in more recently collected samples.

In summary, the data from a large range of biological and social relationships confirm that genetic factors account for the majority of individual differences in liability to smoking initiation, with a small proportion resulting from the consequences of assortative mating. Shared environmental factors do play a significant role, but are primarily due to within-generational influences, e.g. siblings and co-twins. The association between smoking behavior in parents and their children can be most likely accounted for by their genetic relatedness. The idea of social learning in smoking may apply to siblings or peers but does not appear to apply to children learning by modeling from their parents.

Limitations

Given the complexity of the model and the large number of estimated parameters, caution is needed in the interpretation of the results. Even with a large sample as the Virginia 30,000, information may be limited to estimate some parameters, especially those that are highly correlated or only identified by one or few relationships. For example, the correlation between the male and female special twin environmental parameters is derived from the difference between the opposite sex dizygotic twin correlation and the two same sex twin correlations.

Second, the sample was entirely Caucasian, and we do not know whether the pattern of results holds for other ethnic groups. Third, the sample comprises twins and relatives ascertained through a population based registry, as well as subjects ascertained through advertisement in the AARP newsletter. Although this combined sample departs from being representative of the US population, the impact on correlations between relatives is likely to be relatively small. The prevalence rates could, however, be more substantially affected.

Acknowledgments This research has been supported by grants AG04954, GM30250, GM32732, AA06781, AA07728, AA07535 and MH40828 from NIH, grant 941177 from the NH&MRC, a gift from R.J.R. Nabisco and grants from the JM Templeton Foundation. The authors would also like to thank the twins and their families for their participation in this project. The first author is supported by grants HL60688, MH45268, DA016977, MH068521, DA018673 and the Virginia Tobacco Youth Project.

References

Andersen MR, Leroux BG, Bricker JB, Rajan KB, Peterson AV Jr (2004). Antismoking parenting practices are associated with reduced rates of adolescent smoking. Arch Pediat Adol Med 158:348–352

Boomsma DI, Koopmans JR, Van Doornen LJ, Orlebeke JF (1994). Genetic and social influences of starting to smoke: a study of Dutch adolescent twins and their parents. Addiction 89:219–226

Bricker JB, Rajan KB, Andersen MR, Peterson AV Jr (2005). Does parental smoking cessation encourage their young adult children to quit smoking? A prospective study. Addiction 100:379–386

Centers for Disease Control and Prevention (2004) http://www.cdc.gov/tobacco/sgr/sgr_2004/sgranimation/flash/index.html, accessed Jul y 2005

den Exter Blokland EA, Engels RC, Hale WW III, Meeus W, Willemsen MC (2004). Lifetime parental smoking history and cessation and early adolescent smoking behavior. Prev Med 38:359–368

de Vries H, Engels R, Kremers S, Wetzels J, Mudde A (2003). Parents' and friends' smoking status as predictors of smoking onset: findings from six European countries. Health Educ Res 18:627–636

Eaves LJ, Eysenck HJ (1980). The causes and effects of smoking. Maurice Temple Smith, Ltd, London

Eaves LJ, Eysenck HJ, Martin NG (1989). Genes, culture and personality: an empirical approach. Oxford University Press, London

Eaves LJ, Heath AC, Martin NG, Neale MC, Meyer JM, Silberg JL, Corey LA, Truett K, Walters E (1999). Comparing the biological and cultural inheritance of stature and conservatism in the kinships of monozygotic and dizygotic twins. In: Cloninger CR (ed) Proceedings of 1994 APPA Conference, pp 269–308

Fulker DW (1988). Genetic and cultural transmission in human behavior. In: Weir SB, Eisen EJ, Goodman MM, Namkoong G (eds) Proceedings of the second international conference on quantitative genetics, pp 318–340

Heath AC, Eaves LJ (1985). Resolving the effects of phenotype and social background on mate selection. Behav Genet 15:15–30

Heath AC, Cates R, Martin NG, Meyer J, Hewitt JK, Neale MC, Eaves LJ (1993). Genetic contribution to risk of smoking initiation: comparisons across birth cohorts and across cultures. J Subst Abuse 5:221–246

Heath AC, Madden PA, Martin NG (1998). Statistical methods in genetic research on smoking. Stat Methods Med Res 7:165–186

Kendler KS, Karkowski LM, Pedersen NC (2000). Tobacco consumption in Swedish twins reared-apart and reared-together. Arch Gen Psychiat 57:886–892

Li C, Pentz MA, Chou CP (2002). Parental substance use as a modifier of adolescent substance use risk. Addiction 97:1537–1550

Li MD (2003). The genetics of smoking related behavior: a brief review. Am J Med Sci 326:168–173

Madden PA, Pedersen NL, Kaprio J, Koskenvuo MJ, Martin NG (2004). The epidemiology and genetics of smoking initiation and persistence: cross-cultural comparisons of twin study results. Twin Res 7:82–97

Maes HH, Neale MC, Martin NG, Heath AC, Eaves LJ (1999). Religious attendance and frequency of alcohol use: same genes or same environments: a bivariate extended twin kinship model. Twin Res 2:169–179

Neale MC, Boker SM, Xie G, Maes HH (2003). Mx: statistical modelling. Department of Psychiatry Virginia, Commonwealth University, Richmond Virginia

Neale MC, Cardon LR (1992). Methodology for genetic studies of twins and families. Kluwer Academic Publishers, Dortrecht

Neale MC, Miller MB (1997). The use of likelihood-based confidence intervals in genetic models. Behav Genet 27:113–120

Ooki S, Yamada K, Asaka A, Hayakawa K (1990). Zygosity diagnosis of twins by questionnaire. Acta Genet Med Gemel (Roma) 39:109–115

Peterson AV Jr, Leroux BG, Bricker J, Kealey KA, Marek PM, Sarason IG, Andersen MR (2006). Nine-year prediction of adolescent smoking by number of smoking parents. Addict Behav 31:788–801

Prescott CA, Maes HH, Kendler KS (2005). Genetics of substance use disorders. In: Kendler KS, and Eaves LJ (eds) Psychiatric genetics (Review of psychiatry Vol 24), Washington, DC: American Psychiatric Publishing Inc.

Price RA, Vandenberg SG (1980). Spouse similarity in American and Swedish couples. Behav Genet 10:59–71

Ries LAG, Eisner MP, Kosary CL, Hankey BF, Miller BA, Clegg L, Mariotto A, Feuer EJ, Edwards BK (eds), SEER cancer statistics review, 1975-2001, National Cancer Institute. Bethesda MD, (http://www.seer.cancer.gov/csr/1975_2001/) Accessed June 2005

Rose RJ, Viken RJ, Dick DM, Bates JE, Pulkkinen L, Kaprio J (2003). It does take a village: nonfamilial environments and children's behavior. Psychol Sci 14:273–277

Shakib S, Mouttapa M, Johnson CA, Ritt-Olson A, Trinidad DR, Gallaher PE, Unger JB (2003). Ethnic variation in parenting characteristics and adolescent smoking. J Adol Health 33:88–97

Simons-Morton B, Chen R, Abroms L, Haynie DL (2004). Latent growth curve analyses of peer and parent influences on smoking progression among early adolescents. Health Psychol 23:612–621

Sullivan PF, Kendler KS (1997). The genetic epidemiology of smoking. Nicotine Tob Res 1:S51–S57

Truett KR, Eaves LJ, Walters EE, Heath AC, Hewitt JK, Meyer JM, Silberg J, Neale MC, Martin NG, Kendler KS (1994). A model system for analysis of family resemblance in extended kinships of twins. Behav Genet 24:35–49

Vink JM, Willemsen G, Boomsma DI (2003). The association of current smoking behavior with the smoking behavior of parents, siblings, friends and spouses. Addiction 98:923–931

Vink JM, Willemsen G, Engels RC, Boomsma DI (2003). Smoking status of parents, siblings and friends: predictors of regular smoking? Findings from a longitudinal twin-family study. Twin Res 6:209–217

Vitaro F, Wannera B, Brendgena M, Gosselinb C, Gendreaua PL (2004). Differential contribution of parents and friends to smoking trajectories during adolescence. Addict Behav 29:831–835

World Health Organization, http://www.who.int/tobacco/health_priority/en/index.html, accessed July 2005