

Does smoking among friends explain apparent genetic effects on current smoking in adolescence and young adulthood?

VM White^{*,1}, GB Byrnes², B Webster¹ and JL Hopper²

¹The Centre for Behavioural Research in Cancer, The Cancer Council Victoria, Carlton, Victoria 3053, Australia; ²The Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, School of Population Health, The University of Melbourne, Victoria 3010, Australia

We used data from a prospective cohort study of twins to investigate the influence of unmeasured genetic and measured and unmeasured environmental factors on the smoking behaviour of adolescents and young adults. Twins were surveyed in 1988 (aged 11–18 years), 1991, 1996 and 2004 with data from 1409, 1121, 732 and 758 pairs analysed from each survey wave, respectively. Questionnaires assessed the smoking behaviour of twins and the perceived smoking behaviour of friends and parents. Using a novel logistic regression analysis, we simultaneously modelled individual risk and excess concordance for current smoking as a function of zygosity, survey wave, parental smoking and peer smoking. Being concordant for having peers who smoked was a predictor of concordance for current smoking ($P < 0.001$). After adjusting for peer smoking, monozygotic (MZ) pairs were no more alike than dizygotic pairs for current smoking at waves 2, 3 and 4. Genetic explanations are not needed to explain the greater concordance for current smoking among adult MZ pairs. However, if they are invoked, the role of genes may be due to indirect effects acting through the social environment. Smoking prevention efforts may benefit more by targeting social factors than attempting to identify genetic factors associated with smoking.

British Journal of Cancer (2008) **98**, 1475–1481. doi:10.1038/sj.bjc.6604250 www.bjcancer.com

Published online 4 March 2008

© 2008 Cancer Research UK

Keywords: smoking; adolescence; young adults; twins; longitudinal

Smoking is an important cause of preventable mortality and morbidity in later life (Ezzati and Lopez, 2004), so there is a need to understand the factors associated with its uptake and establishment. Twin studies have the potential to identify whether genetic factors might play a role in explaining individual variation in smoking behaviours. The observation that identical (monozygotic; MZ) twin pairs are more similar than same-sex non-identical (dizygotic; DZ) twin pairs is often interpreted as showing that genetic factors play a role, because this finding is consistent with such an explanation under the assumptions of the classic twin model (CTM). One of the main assumptions of the CTM is that the effects of the shared environment on the relevant trait are the same for MZ and DZ pairs (the equal environments assumption (EEA)). Under this assumption, any greater similarity for MZ pairs compared with DZ pairs is attributed to their greater genetic similarity.

Using the CTM, studies of the smoking behaviour of adult twins have been interpreted as showing that genetic factors play a major role in both initiation and persistence of smoking (Carmelli *et al*, 1992; Heath and Martin, 1993; Heath *et al*, 1993, 1999; Madden *et al*, 1999, 2004; Maes *et al*, 2006). Environmental and lifestyle factors shared by twins have been found to play only a small role in adult smoking (Sullivan and Kendler, 1999; Li *et al*, 2003), although they may be more important in the smoking behaviours

of adolescents and young adults (Boomsma *et al*, 1994; Han *et al*, 1999; Koopmans *et al*, 1999; McGue *et al*, 2000; Hopfer *et al*, 2003; Rhee *et al*, 2003; White *et al*, 2003).

There is increasing recognition that violations of the EEA for smoking may influence heritability estimates for smoking and therefore its adequacy needs examination (Kendler and Gardner, 1998; Rende *et al*, 2005; Pergadia *et al*, 2006; Prescott *et al*, 2006; Kaprio, 2007; Tishler and Carey, 2007). Several genetically informative sibling studies have found that the role of the common environment in explaining variation in smoking is greater among siblings (including twins) who share friends than those who do not (Madden *et al*, 2004; Rende *et al*, 2005). Although various interpretations have been given for these findings, there is some agreement that considering social influences on smoking in genetically informative designs may increase our understanding of the aetiology of smoking (Conger, 2005; Merikangas, 2005; Rende *et al*, 2005; Dick *et al*, 2007). We have conducted a prospective, longitudinal study of smoking from adolescence to adulthood using MZ and DZ twin pairs, utilising a social influence framework to understand smoking uptake. Twins were measured at four times (waves 1–4) spanning 17 years, with the median age of the twins increasing from 15 years at wave 1 to 31 years at wave 4. During this time, the twins began to live apart and spend substantially less time with one another. Data on the use of tobacco as well as information on the smoking behaviours of friends and parents, factors found to be influential in adolescent smoking behaviour (Conrad *et al*, 1992; Tyas and Pederson, 1998), were collected at each wave.

*Correspondence: Dr VM White; E-mail: Vicki.White@cancervic.org.au
Received 24 May 2007; revised 16 January 2008; accepted 16 January 2008; published online 4 March 2008

We used data from this study to investigate more fully the relative influence of friends' smoking, genes and other non-genetic factors on smoking during adolescence and young adulthood. We also used a novel analytic method that allows us to study factors that may modify both individual behaviours and the similarity of behaviour within twin pairs. We focus on current smoking, rather than whether participants had ever smoked, as it is a person's continued current smoking that is of greatest relevance to their future health.

MATERIALS AND METHODS

Procedure

The procedures for recruiting the sample into wave 1 have been described previously (Hopper *et al*, 1992; White *et al*, 2003). In brief, during 1988, questionnaires were mailed to adolescent twins (then aged 9–19 years) registered with the Australian Twin Register via their parents, and completed questionnaires were received from 2863 twins, of whom 1417 were pairs, representing a 97% pairwise response. During 1991, the parents and twins participating at wave 1 were approached by letter and asked to complete the wave 2 survey. At wave 2, 2356 completed questionnaires were returned. Five years later, the wave 3 questionnaire was mailed to all twins participating in the wave 1 survey and a total of 1841 were returned. In 2004, the 2726 twins from wave 1 still registered with the Australian Twin Register were approached and 1884 participated in the study (259 were not contactable and 329 withdrew). Based on the number of contactable twins, a 79% response rate at wave 4 was achieved with 66% of individuals participating at wave 1 also participating at wave 4. At wave 4, 773 twin pairs participated. For this paper, we

used data from the pairs who participated at wave 1, were aged 11–18 years (comprising 99% of wave 1 pairs) and who participated in a subsequent wave. Table 1 shows the number of pairs by type at each wave used in these analyses. Twin pair-type predicted the probability of the twins participating in the subsequent survey wave, with same-sex DZ twins (odds ratio (OR) = 0.81, $P = 0.025$) and opposite-sex DZ twins (OR = 0.65, $P < 0.001$) being less likely to return than MZ twins. These effects were approximately constant across waves, as indicated by the lack of significant interaction between wave and pair-type.

Dependent variable: current smoking

At waves 1, 2 and 3, respondents indicating that they had smoked in the week before the survey were defined as current smokers. At wave 4, respondents indicating that they smoked daily or at least weekly were classified as current smokers.

Smoking behaviours of parents, peers and co-twins

At each wave, respondents indicated the perceived smoking status of their mother, father, co-twin and for each of up to four friends. Respondents classified their parents as being a 'non-smoker', an 'ex-smoker' or a 'smoker'. We focussed on current smoking of parents. Due to small numbers reporting that both parents smoked (see Table 1), parental smoking was classified into two groups: neither parent currently smoked or at least one parent currently smoked.

At each wave, twins reported on the smoking status (non-smokers, ex-smokers, occasional, light or heavy smokers) of up to four of their closest friends. The proportion of friends who engaged in any kind of smoking was determined by dividing the

Table 1 Characteristics of participants in each survey wave (base: individual twins)

Characteristic (n)	Wave 1 (2818)	Wave 2 (2242)	Wave 3 (1464)	Wave 4 (1516)
Average age	14.91	18.00	22.51	30.75
<i>Twin type</i>				
MZ (pairs)	1310 (655)	1076 (538)	752 (376)	746 (373)
DZS (pairs)	824 (412)	652 (326)	406 (203)	436 (218)
DZO (pairs)	684 (342)	514 (257)	306 (153)	334 (167)
<i>Gender</i>				
Males (%)	46	44	41	41
Females (%)	54	56	59	59
<i>Smoking status</i>				
Never smoked (%)	37	23	22	13
Experimenter (%)	49	54	49	55
Current smoker (%)	14	23	29	20
Ex-smoker (%)	NA	NA	NA	11
<i>Friend's smoking</i>				
No friends smoke (%)	65	44	39	46
<50% smoking (%)	20	32	42	39
≥50% smoking (%)	15	25	19	16
<i>Parent's current smoking</i>				
None (%)	65	69	77	77
One parent (%)	26	24	19	18
Both parents (%)	10	7	4	5
<i>Meet with twin</i>				
Everyday (%)	96	81	46	11
Weekly (%)	2	10	22	32
Less than weekly (%)	2	10	33	57

DZO = opposite-sex dizygotic or fraternal twins; DZS = same-sex dizygotic or fraternal twins; MZ = monozygotic or identical twins; NA = not applicable. NB: % may not add to 100 due to rounding.

number of smoking friends by the total number of friends listed. For correlations and logistic regressions, friends' smoking was classified into two groups: no smokers among friends or at least one smoker.

At each wave, twins were asked if they were living with their twin, as well as the frequency of meeting their twin, with responses classified into every day, at least weekly and less than weekly.

Statistical analyses

Proportions of individual twins who were current smokers and who had friends who smoked were calculated by wave and pair-type. As there is no *a priori* reason to prefer tetrachoric correlations to other measures, we estimated both the Pearson and tetrachoric twin-pair correlations (r) for current smoking status and friends' smoking status for each wave and zygosity group. Under the assumptions of the CTM (see Introduction), we estimated naïve heritability as $2(r_{MZ} - r_{DZ})$ and examined its dependence on correlation type and wave.

We estimated parameters in a single model to describe both the probability of current smoking for individual twins and the probability of concordance of current smoking for pairs. Pairs were defined as being concordant for their smoking behaviour at a given wave if both were current smokers or if both were not current smokers. Note that specifying the probability of each twin smoking and the probability that the pairs are concordant for smoking is equivalent to specifying the probabilities of each of the four possible pair-smoking states: both smoke, two combinations where only one smokes and neither smoke (see Appendix).

We used logistic regression to model both individual smoking and pair concordance, and estimated the two sets of parameters simultaneously by maximum likelihood. As the probability of an individual smoking is not independent of the probability of pair concordance (e.g., if everyone smokes, all pairs will be concordant), we included a compensating term in the predictor of concordance such that if all coefficients (log-odds for concordance) were zero, the predicted probability of concordance would be exactly as if twins in the same pair were uncorrelated in their smoking behaviours. Hence, in effect, we estimated predictors of excess concordance. To allow for correlation between observations on the same pairs in separate waves, robust (Huber-White) estimates of standard errors were used. The optimisation procedure was coded in Stata 8 and Stata 9 (StataCorp, 2003, 2006), using the ML package. Multiple starting points were used for each run and the progress of the fitting algorithm was

monitored to determine, so far as possible, that convergence was to the global maximum likelihood. Convexity of the likelihood surface was also checked using simulated data, by plotting in the neighbourhood of the analytically determined maximum. The order of twins was randomly permuted at the start of each run to avoid the possibility of bias due to any unplanned systematic ordering within pairs. Further details of the method are given in the Appendix. Predictors of individual current smoking could be either individual-specific (sex, smoking among peers) or pair-specific (i.e., zygosity, parental smoking). Predictors of concordance were necessarily pair-specific (zygosity, parental smoking, same or different sex and age). To capture any time dependence of effects, the wave number was included in the initial model, both as a main effect and in interaction with zygosity.

To minimize bias due to unequal duration of participation, only those variables that were measured at each wave or that remained constant with time (e.g., sex) were used in the analysis. Only those observations for which both twins participated at the wave could be retained, due to the pairwise nature of the analytic method.

RESULTS

Sample description

Smoking status of the twin pair was associated with subsequent survey participation. Pairs in which one ($OR = 0.62$, $P < 0.001$) or both ($OR = 0.49$, $P < 0.001$) smoked were less likely to participate in a subsequent survey than were pairs in which neither smoked. However, pairs concordant for smoking status were not significantly more likely to remain in the study ($OR = 1.12$, $P = 0.25$) than non-concordant pairs.

The proportion of participating twins who were current smokers increased from wave 1 to wave 3, whereas the proportion of twins with no smokers among their friends decreased (see Table 1). The proportion of twins indicating that neither parent currently smoked also decreased between waves 1 and 3.

Figure 1 shows that the proportion of twins cohabiting decreased over the period of the study. The frequency of contact within pairs also decreased for all twin pair-types as the twins aged. There was an association between zygosity and contact, with more MZ twins reporting daily contact with their co-twin than DZ twins.

Table 2 shows that, at each wave, the prevalence of current smoking for MZ twins was lower than for DZ twins. At each wave, and regardless of twin pair-type, twins who were current smokers

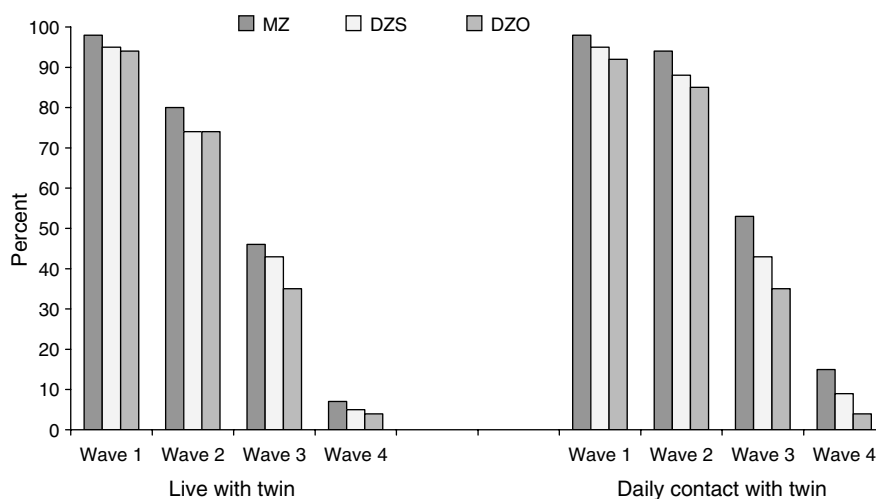


Figure 1 Proportion of twins by zygosity who live with their twin (left) and meet with their twin everyday (right), by survey wave. DZO, opposite-sex dizygotic or fraternal twins; DZS, same-sex dizygotic or fraternal twins; MZ, monozygotic or identical twins.

Table 2 For each twin type, percentage of current smokers at each survey wave and percentage with any smokers among their friends by smoking status, at each survey wave (base: individual twins)

	Twin type			Total
	MZ	DZS	DZO	
Current smokers				
Wave 1 (%)	12	14	17	14
Wave 2 (%)	20	25	26	23
Wave 3 (%)	26	31	31	29
Wave 4 (%)	19	21	24	20
Any smokers among friends				
Wave 1				
Non-smokers (%)	24	30	27	26
Current smokers (%)	90	94	84	90
Total (%)	32	39	37	35
Wave 2				
Non-smokers (%)	46	43	48	46
Current smokers (%)	93	91	91	92
Total (%)	56	54	59	56
Wave 3				
Non-smokers (%)	51	49	49	50
Current smokers (%)	90	88	86	88
Total (%)	61	61	61	61
Wave 4				
Non-smokers (%)	44	47	45	45
Current smokers (%)	90	91	92	91
Total (%)	53	56	56	54

DZO = opposite-sex dizygotic or fraternal twins; DZS = same-sex dizygotic or fraternal twins; MZ = monozygotic or identical twins.

Table 3 Intraclass correlations (s.e.) for current smoking in twins and for smoking among friends, by zygosity and wave (base: twin pairs)

	Wave 1	Wave 2	Wave 3	Wave 4
Current smoking in twins				
<i>Pearson correlation</i>				
MZ	0.65 (0.05)	0.58 (0.04)	0.52 (0.05)	0.44 (0.06)
DZS	0.35 (0.05)	0.29 (0.04)	0.25 (0.05)	0.24 (0.06)
$2(r_{MZ}-r_{DZ})^a$	0.61 (0.13)	0.59 (0.12)	0.53 (0.14)	0.39 (0.17)
<i>Tetrachoric correlation</i>				
MZ	0.89 (0.03)	0.82 (0.04)	0.75 (0.05)	0.69 (0.06)
DZS	0.59 (0.06)	0.47 (0.06)	0.41 (0.08)	0.42 (0.08)
$2(r_{MZ}-r_{DZ})^a$	0.61 (0.06)	0.71 (0.07)	0.68 (0.09)	0.55 (0.10)
Smoking among friends				
<i>Pearson correlation</i>				
MZ	0.58 (0.04)	0.45 (0.04)	0.38 (0.04)	0.39 (0.06)
DZS	0.42 (0.03)	0.33 (0.04)	0.19 (0.06)	0.11 (0.05)
$2(r_{MZ}-r_{DZ})^a$	0.33 (0.10)	0.23 (0.11)	0.38 (0.14)	0.57 (0.16)
<i>Tetrachoric correlation</i>				
MZ	0.80 (0.03)	0.65 (0.05)	0.57 (0.07)	0.58 (0.06)
DZS	0.62 (0.04)	0.50 (0.06)	0.30 (0.08)	0.42 (0.08)
$2(r_{MZ}-r_{DZ})^a$	0.37 (0.05)	0.30 (0.07)	0.54 (0.11)	0.32 (0.10)

DZS = same-sex dizygotic or fraternal twins; MZ = monozygotic or identical twins.

^a $2(r_{MZ}-r_{DZ})$ = provides naïve heritability estimates.

were much more likely to have smokers among their friends than twins who were not current smokers: about 90% of twins who smoked had smokers among their friends at each survey wave.

Correlations

Table 3 shows that for current smoking, the tetrachoric correlations were in general higher than Pearson correlations, which translated into higher naïve heritability estimates at waves 2, 3 and 4. There appeared to be a decline in naïve heritability defined in these terms between waves 3 and 4, more so when measured by Pearson correlation.

Table 3 also shows the Pearson and tetrachoric correlations for friends' smoking status and the naïve heritability estimate for each wave. Correlations were higher for MZ pairs than for DZ pairs and again the tetrachoric correlations were higher than the Pearson correlations.

Associations with smoking and concordance

Table 4 shows the estimates of multivariate ORs for both the probability of individuals smoking and the probability of excess concordance of current smoking under two models. Model 1 includes zygosity, parental smoking and wave as predictors of individual smoking. For excess concordance, it includes zygosity and its interaction with wave (shown in Table 4 as the effect of zygosity at each wave). All estimates shown are adjusted for other variables in the model. Overall, MZ twins were less likely to smoke ($P < 0.001$), and this association was consistent across all waves (there was no interaction with wave number). There was a significant effect of wave indicating the greater probability of being a current smoker with increasing age. Regarding pair concordance, MZ pairs were more alike than DZ pairs for their smoking behaviours at waves 1, 2 and 3 ($P = 0.003$, 0.004 and 0.03, respectively).

Model 2 adds to Model 1 the effect of friends' smoking at the individual level as well as the concordance of friends' smoking at the pair level. There was a strong association between an individual's smoking status and that of their peers ($OR = 10.9$, $P < 0.001$), which was far greater than the association with parental smoking ($OR = 1.75$, $P < 0.001$). There was no evidence that the parental or peer associations varied by wave.

When modelling the probability of pair concordance for current smoking, concordance of friends' smoking status was a significant predictor and this association did not differ significantly across waves. Including concordance for friends' smoking in Model 2 reduced the difference between MZ and DZ concordance for smoking estimated under Model 1 at waves 2, 3 and 4, such that the effect of zygosity was no longer significant at these waves.

DISCUSSION

Consistent with most other investigators, we found greater correlation for current smoking in MZ twin pairs than in DZ twin pairs using both Pearson and tetrachoric correlation estimates. The corresponding heritability estimates under the assumptions of the CTM were consistent with values reported from various twin studies (Carmelli *et al*, 1992; Heath and Martin, 1993; Heath *et al*, 1993, 1999; Boomsma *et al*, 1994; Han *et al*, 1999; Koopmans *et al*, 1999; Madden *et al*, 1999, 2004; McGue *et al*, 2000; Rhee *et al*, 2003; Maes *et al*, 2006). However, using a novel analysis, which allowed us to adjust for both the smoking status of an individual's friends and the concordance for friends' smoking status within pairs, we found greater concordance for smoking in MZ pairs only at wave 1 when most twins were living together.

There is considerable evidence in the literature that the smoking behaviour of friends has a major influence on the current and future smoking of adolescents and young adults (Flay *et al*, 1994, 1998; Distefan *et al*, 1998; Engels *et al*, 1999; Chassin *et al*, 2000; Leatherdale *et al*, 2005). If, as our study found, MZ pairs are more similar in their friends' smoking behaviours than are DZ pairs, this could explain some or all of the greater correlation of smoking

Table 4 Multivariate ORs jointly estimated for individual current smoking and excess concordance of smoking within twin pairs

Factor	Model 1			Model 2		
	OR	95% CI	P-value	OR	95% CI	P-value
<i>Individual current smoking</i>						
Monozygosity	0.77	0.63–0.94	0.01	0.77	0.62–0.94	0.01
Parents smoke	1.86	1.63–2.11	<0.001	1.75	1.45–2.09	<0.001
Wave 2	2.72	2.29–3.22	<0.001	1.20	0.99–1.46	0.06
Wave 3	1.70	1.42–2.04	<0.001	1.61	1.29–2.02	<0.001
Wave 4	1.93	1.64–2.27	<0.001	1.08	0.87–1.34	0.5
Friends' smoking	NA			10.90	7.90–15.10	<0.001
<i>Excess concordance of current smoking</i>						
Monozygosity (wave 1)	1.99	1.26–3.15	0.003	2.60	1.47–4.60	0.001
Monozygosity (wave 2)	1.83	1.22–2.76	0.004	1.12	0.78–1.60	0.5
Monozygosity (wave 3)	1.68	1.07–2.66	0.03	1.13	0.79–1.62	0.5
Monozygosity (wave 4)	1.30	0.86–1.98	0.2	1.18	0.66–1.68	0.5
Concordance of friends' smoking	NA			1.55	1.30–1.85	<0.001

CI = confidence interval; OR = odds ratio; NA = not applicable. Excess concordance also adjusted for main effect of wave (no significant association). Estimates shown adjusted for all variables in model.

observed in MZ twins. Our study found this to be the case, at least during the time when twins start to live apart.

If MZ pairs share a more similar environment than DZ pairs, then either the EEA of the CTM is invalid, or the environment must be considered a manifestation of the twins' genes. The first possibility would result in biased heritability estimates; the second would imply a broad notion of heritability, part of which may be subject to environmental modification. This is the standard interpretation of heritability used in zoology, where it is used to determine the response to selection (Mulder *et al*, 2007). The EEA has been tested only under a limited number of circumstances (Loehlin, 1992) and for only some substances, and the findings have been mixed (Prescott *et al*, 2006). Although we did not formally assess this assumption, we found that, compared with DZ twins, MZ twins had more frequent contact with each other at all survey waves and their friends were more similar in smoking behaviours. Other work has also found that adolescent MZ twin pairs spend more time together and share more friends than do DZ twin pairs (Rende *et al*, 2005), and that MZ pairs are more dependent on their co-twin than DZ pairs (Penninkilampi-Kerola *et al*, 2005). These findings are in line with the suggestion that a 'special MZ environment' might contribute to the greater similarities in the smoking behaviours of MZ twin pairs (Stallings *et al*, 1999). Directly measuring and adjusting for differences in shared environment may help reduce any resultant bias on heritability estimates. Several recent studies involving adolescents have commenced this investigation (Rende *et al*, 2005; Slomkowski *et al*, 2005; Pergadia *et al*, 2006; Dick *et al*, 2007). Dick *et al* (2007) found that adjusting for parental contact influenced heritability estimates for smoking and suggested that different environments moderate genetic effects on the variability in tobacco use. Rende and colleagues found that adjusting for shared friends and amount of contact between twins influenced the role of the shared environment, but not genes, on smoking variability, suggesting to them a sibling 'contagion effect' that operates through environmental processes (Rende *et al*, 2005; Slomkowski *et al*, 2005). None of these studies modelled the influence of concordance of friends' smoking in their models.

To disentangle the possible confounding of the effects of genes and friends, we developed an analytic method that could adjust for measured covariates of smoking. Our approach allowed us to include both the smoking status of each twin's friends and the concordance of friends' smoking status of a twin pair. After adjusting for these factors, there was no evidence for increased concordance in MZ pairs at waves 2, 3 or 4.

A necessary consequence of a genetic contribution to behaviour is that MZ pairs are more alike in that behaviour than are DZ pairs. If this difference is not observed then it is problematic to accept the hypothesis of a genetic contribution. A more difficult question is whether greater similarity in MZ pairs is sufficient to conclude a role for genes. Our findings suggest that it is possible to explain the greater concordance in smoking for MZ pairs compared with DZ pairs at waves 2–4 without reference to unmeasured genetic factors. Monozygotic pairs may be more alike in their smoking for the simple reason that their friends are more alike in their smoking.

Two outstanding issues remain: what is the source of greater concordance for smoking in MZ pairs at wave 1; and what is the source of greater concordance for the smoking status of the friends for MZ pairs at all waves.

Regarding the first issue, although the pattern of results could suggest genetic factors influencing smoking 'turn-on' during adolescence and 'turn-off' during young adulthood, it may also suggest that there is unmeasured confounding due to greater shared environment for MZ pairs while living together.

For the second issue, it is possible that the greater social contact or connectedness between MZ twins than DZ twins (Penninkilampi-Kerola *et al*, 2005; Rende *et al*, 2005) simply leads to twins sharing more friends. However, as the literature based on the CTM provide some evidence that the choice of friends (Baker and Daniels, 1990; Iervolino *et al*, 2002) and exposure to friends who smoke (Cleveland *et al*, 2005) are influenced by genetic factors, it is also possible to suggest that the influence of friends on smoking behaviour is, at root, genetic. From this position and assuming that the EEA is correct, our findings could suggest that genes influence the smoking behaviours of adolescents and young adults indirectly by influencing friendship selection. One motivation for seeking genes influencing smoking behaviour is that their discovery could provide a target for pharmaceutical interventions, by either blocking or enhancing the action of the proteins encoded by the genes (Tyndale, 2003). However, a gene that modifies smoking indirectly by influencing the selection of friends would be a less likely target for pharmaceutical interventions. If genetic explanations of smoking are to be made, research needs to distinguish the contribution of direct and indirect genetic effects, as this will determine whether searching for specific genes associated with smoking is likely to be fruitful. If genetic effects on smoking mainly act through environmental mechanisms as is suggested by our results, then social interventions may be the most effective means at reducing smoking.

Although this study has a number of strengths, including its longitudinal nature and assessing friends' smoking status at each survey wave, several limitations need to be kept in mind. First, there was attrition from waves 1 to 4. Smokers were less likely to participate in the study at later waves than non-smokers, as were DZ twins. This could lead to a false association between zygosity and individual smoking status, with the progressive concentration with each wave expected to produce an increasingly strong association between zygosity and smoking. However, this was not observed. Second, we studied current smoking status at each wave rather than studying the status of ever having been a current or regular smoker or ever smoking, as has been usual in the behavioural genetics smoking literature. We adopted this strategy due to the young age of our sample at the first wave. The notion of an adolescent as a smoker is more ambiguous than it is among adults, with research suggesting that a non-trivial proportion of older adolescents who describe themselves as a smoker refer to themselves as a 'non-smoker' 6 months later (Schofield *et al*, 1998). Finally, smoking was more common among our DZ twins than MZ twins at all survey waves. Although this finding must be considered with caution due to differential retention in later waves, as discussed above, it may indicate a violation of the EEA leading to inflated heritability estimates (Tishler and Carey, 2007). Although the veracity of this suggestion is being debated (Kaprio, 2007), increasing concern about the appropriateness of the EEA for smoking led us to describe our heritability estimates as 'naïve'.

Recently, there has been a call for researchers studying twins to investigate how environmental factors might mediate or influence the putative role of genes (Pennikilampi-Kerola *et al*, 2005). Our study has attempted to do this by using a novel analytic approach that does not rely on the assumptions of the CTM to examine the influence of zygosity and peer smoking on the current smoking behaviours of adolescent twins as they grow to adulthood. Once

concordance for peer smoking was adjusted for, MZ pairs were no more likely to be concordant for smoking than DZ pairs in late adolescence and adulthood. Our results suggest that genetic explanations do not need to be invoked to explain the greater concordance for smoking in MZ pairs. Although further research is needed to confirm our findings, we believe that they are in line with results from genetically informative designs showing the importance of the social environment on smoking uptake (Rende *et al*, 2005; Slomkowski *et al*, 2005) and in moderating the influence of genes (Dick *et al*, 2007). This growing body of work provides support for the suggestion that smoking prevention efforts may benefit more by targeting social influences than attempting to identify genes associated with smoking (Merikangas, 2005). This public health approach should include well-funded mass media anti-tobacco advertising campaigns, restrictions on smoking in public, increased prices for cigarettes and removal of all tobacco product advertising (Laugesen *et al*, 2000). Since the late 1980s, Australia has adopted many of these policies and the prevalence of smoking among adolescents aged 12–17 years has fallen from 22% in 1984 to 9% in 2005 (White and Hayman, 2006).

ACKNOWLEDGEMENTS

This project was funded by grants from the National Health and Medical Research Council (NHMRC) (Grant no. 262604) and from the Victorian Health Promotion Foundation. Professor John Hopper is an NHMRC Senior Principal Research Fellow. The authors are grateful to the Australian Twin Registry for facilitating access to twins through their parents at recruitment and directly at follow-up. Particular thanks go to the twins.

REFERENCES

- Baker LA, Daniels D (1990) Nonshared environmental influences and personality differences in adult twins. *J Pers Soc Psychol* **58**: 103–110
- Boomsma DI, Koopmans JR, Van Doornen LJ, Orlebeke JF (1994) Genetic and social influences on starting to smoke: a study of Dutch adolescent twins and their parents. *Addiction* **89**: 219–226
- Carmelli D, Swan G, Robinette D, Fabsitz R (1992) Genetic influence on smoking: a study of male twins. *N Engl J Med* **327**: 829–833
- Chassin L, Presson CC, Pitts SC, Sherman SJ (2000) The natural history of cigarette smoking from adolescence to adulthood in a midwestern community sample: multiple trajectories and their psychosocial correlates. *Health Psychol* **19**: 223–231
- Cleveland HH, Wiebe RP, Rowe DC (2005) Sources of exposure to smoking and drinking friends among adolescents: a behavioral-genetic evaluation. *J Genet Psychol* **166**: 153–169
- Conger RD (2005) Sibling effects on smoking in adolescence: evidence for social influence from a genetically informative design: comment on Slomkowski *et al*, 2005. *Addiction* **100**: 441–442; discussion 443–444
- Conrad KM, Flay BR, Hill D (1992) Why children start smoking cigarettes: predictors of onset. *Br J Addict* **87**: 1711–1724
- Dick DM, Pagan JL, Viken R, Purcell S, Kaprio J, Pulkkinen L, Rose RJ (2007) Changing environmental influences on substance use across development. *Twin Res Hum Genet* **10**: 315–326
- Distefan JM, Gilpin EA, Choi WS, Pierce JP (1998) Parental influences predict adolescent smoking in the United States, 1989–1993. *J Adolesc Health* **22**: 466–474
- Engels RC, Knibbe RA, de Vries H, Drop MJ, van Breukelen GJ (1999) Influences of parental and best friends' smoking and drinking on adolescent use: a longitudinal study. *J Appl Soc Psychol* **29**: 337–361
- Ezzati M, Lopez AD (2004) Regional, disease specific patterns of smoking-attributable mortality in 2000. *Tob Control* **13**: 388–395
- Flay BR, Hu FB, Richardson J (1998) Psychosocial predictors of different stages of cigarette smoking among high school students. *Prev Med* **27**: A9–A18
- Flay BR, Hu FB, Siddiqui O, Day E, Hedeker D, Petraitis J, Richardson J, Sussman S (1994) Differential influence of parental smoking and friends' smoking on adolescent initiation and escalation of smoking. *J Health Soc Behav* **35**: 248–265
- Han C, McGue MK, Iacono WG (1999) Lifetime tobacco, alcohol and other substance use in adolescent Minnesota twins: univariate and multivariate behavioral genetic analyses. *Addiction* **94**: 981–993
- 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, Kirk KM, Meyer JM, Martin NG (1999) Genetic and social determinants of initiation and age at onset of smoking in Australian twins. *Behav Genet* **29**: 395–407
- Heath AC, Martin NG (1993) Genetic models for the natural history of smoking: evidence for a genetic influence on smoking persistence. *Addict Behav* **18**: 19–34
- Hopfer CJ, Crowley TJ, Hewitt JK (2003) Review of twin and adoption studies of adolescent substance use. *J Am Acad Child Adolesc Psychiatry* **42**: 710–719
- Hopper JL, White VM, Macaskill GT, Hill DJ, Clifford CA (1992) Alcohol use, smoking habits and the Adult Eysenck Personality Questionnaire in adolescent Australian twins. *Acta Genet Med Gemellol (Roma)* **41**: 311–324
- Iervolino AC, Pike A, Manke B, Reiss D, Hetherington EM, Plomin R (2002) Genetic and environmental influences in adolescent peer socialization: evidence from two genetically sensitive designs. *Child Dev* **73**: 162–174
- Kaprio J (2007) Differences in smoking habits of MZ and DZ twins: a commentary on Tishler and Carey. *Twin Res Hum Genet* **10**: 718–720
- Kendler KS, Gardner Jr CO (1998) Twin studies of adult psychiatric and substance dependence disorders: are they biased by differences in the environmental experiences of monozygotic and dizygotic twins in childhood and adolescence? *Psychol Med* **28**: 625–633

- Koopmans JR, Slutske WS, Heath AC, Neale MC, Boomsma DI (1999) The genetics of smoking initiation and quantity smoked in Dutch adolescent and young adult twins. *Behav Genet* 29: 383–393
- Laugesen M, Scollo M, Sweanor D, Shiffman S, Gitchell J, Barnsley K, Jacobs M, Giovino GA, Glantz SA, Daynard RA, Connolly GN, Difranza JR (2000) World's best practice in tobacco control. *Tob Control* 9: 228–236
- Leatherdale ST, McDonald PW, Cameron R, Brown KS (2005) A multilevel analysis examining the relationship between social influences for smoking and smoking onset. *Am J Health Behav* 29: 520–530
- Li MD, Cheng R, Ma JZ, Swan GE (2003) A meta-analysis of estimated genetic and environmental effects on smoking behavior in male and female adult twins. *Addiction* 98: 23–31
- Loehlin JC (1992) *Genes and Environment in Personality Development*. Newbury Park, CA: SAGE Publications
- Madden PA, Heath AC, Pedersen NL, Kaprio J, Koskenvuo MJ, Martin NG (1999) The genetics of smoking persistence in men and women: a multicultural study. *Behav Genet* 29: 423–431
- Madden PA, Pedersen NL, Kaprio J, Koskenvuo MJ, Martin NG (2004) The epidemiology and genetics of smoking initiation and persistence: crosscultural comparisons of twin study results. *Twin Res* 7: 82–97
- Maes HH, Neale MC, Kendler KS, Martin NG, Heath AC, Eaves LJ (2006) Genetic and cultural transmission of smoking initiation: an extended twin kinship model. *Behav Genet* 36: 795–808
- McGue M, Elkins I, Iacono WG (2000) Genetic and environmental influences on adolescent substance use and abuse. *Am J Med Genet* 96: 671–677
- Merikangas KR (2005) The significance of social connectedness: comment on Slomkowski et al, 2005. *Addiction* 100: 442–443; discussion 443–444
- Mulder HA, Bijma P, Hill WG (2007) Prediction of breeding values and selection responses with genetic heterogeneity of environmental variance. *Genetics* 175: 1895–1910
- Penninkilampi-Kerola V, Kaprio J, Moilanen I, Rose RJ (2005) Co-twin dependence modifies heritability of abstinence and alcohol use: a population-based study of Finnish twins. *Twin Res Hum Genet* 8: 232–244
- Pergadia ML, Heath AC, Agrawal A, Bucholz KK, Martin NG, Madden PA (2006) The implications of simultaneous smoking initiation for inferences about the genetics of smoking behavior from twin data. *Behav Genet* 36: 567–576
- Prescott CA, Madden PA, Stallings MC (2006) Challenges in genetic studies of the etiology of substance use and substance use disorders: introduction to the special issue. *Behav Genet* 36: 473–482
- Rende R, Slomkowski C, Lloyd-Richardson E, Niaura R (2005) Sibling effects on substance use in adolescence: social contagion and genetic relatedness. *J Fam Psychol* 19: 611–618
- Rhee SH, Hewitt JK, Young SE, Corley RP, Crowley TJ, Stallings MC (2003) Genetic and environmental influences on substance initiation, use, and problem use in adolescents. *Arch Gen Psychiatry* 60: 1256–1264
- Schofield PE, Borland R, Hill DJ, Pattison PE, Hibbert ME (1998) Instability in smoking patterns among school leavers in Victoria, Australia. *Tob Control* 7: 149–155
- Slomkowski C, Rende R, Novak S, Lloyd-Richardson E, Niaura R (2005) Sibling effects on smoking in adolescence: evidence for social influence from a genetically informative design. *Addiction* 100: 430–438
- Stallings MC, Hewitt JK, Beresford T, Heath AC, Eaves LJ (1999) A twin study of drinking and smoking onset and latencies from first use to regular use. *Behav Genet* 29: 409–421
- StataCorp (2003) *Stata Statistical Software: Release 8.2*. College Station, TX: Stata Corporation
- StataCorp (2006) *Stata Statistical Software: Release 9.0*. College Station, TX: Stata Corporation
- Sullivan PF, Kendler KS (1999) The genetic epidemiology of smoking. *Nicotine Tob Res* 1(Suppl 2): S51–S57
- Tishler PV, Carey VJ (2007) Can comparison of MZ- and DZ-twin concordance rates be used invariably to estimate heritability? *Twin Res Hum Genet* 10: 712–717
- Tyas SL, Pederson LL (1998) Psychosocial factors related to adolescent smoking: a critical review of the literature. *Tob Control* 7: 409–420
- Tyndale RF (2003) Genetics of alcohol and tobacco use in humans. *Ann Med* 35: 94–121
- White V, Hayman J (2006) *Smoking Behaviours of Australian Secondary Students in 2005* Australian Government Department of Health and Ageing: Canberra. Available from <http://www.nationaldrugstrategy.gov.au/internet/drugstrategy/publishing.nsf/Content/mono59>
- White VM, Hopper JL, Wearing AJ, Hill DJ (2003) The role of genes in tobacco smoking during adolescence and young adulthood: a multivariate behaviour genetic investigation. *Addiction* 98: 1087–1100

APPENDIX

We demonstrate below that the joint distribution of smoking (or other binary features) within a pair of twins is equivalently captured by the marginal probability of smoking, together with the probability that the pair are concordant. We also provide further details on the combined regression model we used to estimate the odds ratios of marginal smoking and concordance with respect to various predictors.

As the twins within a pair are assumed to be exchangeable, the joint distribution of a pair is described by the probabilities that both, one or none of the twins in the pair smoke. We denote these by p_2 , p_1 and $p_0 = 1 - p_2 - p_1$, respectively.

Now consider the marginal probability p_m that a twin smokes, which is the probability that a twin chosen at random from the population is a smoker. This random draw can be conducted by first choosing a pair, and then selecting one of the twins in the pair at random. It is easy to see that $p_m = p_2 + 0.5p_1$. Similarly, the probability that a pair of twins is concordant is just $p_c = p_0 + p_2 = 1 - p_1$. This relationship can be inverted to yield $p_2 = p_m + 0.5(p_c - 1)$ and $p_1 = 1 - p_c$, establishing that they are equivalent representations.

It follows that if we have a model for p_m and another for p_c , then we have a model for the observed joint distribution. We chose to

use logistic models for both p_m and p_c largely to avoid confusion with the often-assumed probit model, but any generalised linear model for a binary outcome would suffice.

The remaining problem is to disentangle the effect of the marginal probability of smoking from the probability of concordance. Suppose that each twin made an independent decision to smoke, with probability p . Then $p_2 = p^2$ and $p_1 = 2p(1-p)$, so that $p_c = 1 - 2p + 2p^2$. This increases to 1 as p approaches either 0 or 1, as would be expected. To compensate for this, we write the generalised linear models in the form

$$p_m = g_m(\alpha_m + \beta_m X_m)$$

$$p_c = g_c(g_c^{-1}(1 - 2p_m + 2p_m^2) + \beta_c X_c).$$

Here g_m and g_c are the link functions for the marginal and concordance models, whereas X_m and X_c are the corresponding covariates.

Hence, in the null model where $\beta_c = 0$, the probability of concordance is as it would be if the smoking status of twins was independent.