

Risk Factors for Atherosclerosis in Twins

David L. Duffy, Dianne L. O'Connell, Richard F. Heller, and Nicholas G. Martin

Queensland Institute of Medical Research (D.L.D., N.G.M.), Brisbane; Centre for Clinical Epidemiology and Biostatistics, University of Newcastle (D.O'C., R.F.H.), Newcastle, Australia

We performed multivariate genetic analyses of cardiovascular risk factors from two sets of data on US and Australian female twins. Similar models for body-mass index (BMI), serum low density (LDL) and high density (HDL) lipoproteins, including age as a covariate, were fitted successfully to both groups. These suggested that BMI, or genes responsible for a significant proportion of the variance of BMI, explained correlations between lipid subfractions, as well as those between blood pressure and lipid subfractions, especially HDL.

© 1993 Wiley-Liss, Inc.

Key words: heritability, lipoproteins, blood pressure, triglyceride, LISREL

INTRODUCTION

Two twin data sets examining risk factors for atherosclerosis were provided for analysis at GAW8. We chose to examine the data from the Kaiser-Permanente Women Twin Registry (NHLBI female twins). In addition, we looked at a twin data set on cardiovascular risk factors kindly provided by Dr. Dianne O'Connell [O'Connell et al., 1988] describing Australian female adult twins. We thought it interesting to attempt cross-validation of models in these different groups.

METHODS

The design of both studies is described elsewhere. Preliminary statistical analyses were performed using SAS 6.07 [SAS Institute, 1991] and SPSS-PC [Norusis, 1988]. Genetic path models were fitted using LISREL 7.20 [Jöreskog & Sörbom, 1990]. The path modelling included phenotypic multivariate regression performed in a (forward) stepwise fashion and multivariate genetic factor models.

Address reprint requests to Dr. D. Duffy, Epidemiology Unit, Queensland Institute of Medical Research, 300 Herston Road, Herston, Queensland, Australia 4029.

© 1993 Wiley-Liss, Inc.

These latter included "ACE Choleski" factor models where we fit Choleski (triangular) decompositions of n (corresponding to the n variables to be analyzed) orthogonal additive genetic factors (correlated within each twin pair 1.0 in MZ twins and 0.5 in DZ twins), n orthogonal shared environmental factors (correlated within each twin pair 1.0 in MZ twins and in DZ twins), and n orthogonal unshared environmental factors (uncorrelated within each twin pair - see Neale and Cardon, 1992). We also fitted models that included phenotypic direct causative pathways between variables where the literature suggested these might be appropriate. Model goodness-of-fit has been summarized as an overall χ^2 and as the Akaike Information Criterion (AIC), here expressed as $\chi^2 - 2 \cdot (\text{degrees of freedom})$.

RESULTS

Kaiser-Permanente Twins

The Kaiser-Permanente data set contained data from 434 pairs of like-sex female twins - 255 monozygotic (MZ) pairs, and 179 dizygotic (DZ) pairs. The mean age of the twins was 42 years. From this group, we excluded nine pairs of twins with extreme lipid values, as did Austin et al. [1987]. For analysis, we selected resting systolic and diastolic arterial blood pressure (sBP and dBP), and serum levels of triglyceride (sTG), high density and low density lipoprotein (HDL and LDL). We log-transformed sTG because of marked skewing. Regressions of MZ pair-sums on pair-differences for each of these traits (i.e., phenotypic means on intrapair variances) were not significant, suggesting an absence of significant gene by environment interaction [Jinks and Fulker, 1970].

We screened factors that might influence the chosen biological variables. Only age and (log transformed) body mass index (BMI, kg/m^2), a measure of obesity, explained more than 1% of variation in these traits, with the exception of the relationship between alcohol intake and HDL level. Significant residual covariation for risk factors was found, notably sBP and dBP, and HDL, LDL, and sTG.

We then created a phenotypic path model using all the twins and ignoring relatedness. The model was selected by first fixing the direction of relationship between a number of variables, and then performing forward stepwise selection of subsequent paths using the AM option in LISREL. The path directions that were fixed *a priori* include those from age to all variables; from BMI to blood pressure; from sTG to blood pressure [Reaven, 1991], and to HDL [Clay et al., 1991]; from antihypertensive use to sTG and LDL; from blood pressure to antihypertensive use; from exercise to BMI, and from BMI to exercise [Lindon Eaves, pers. comm.]; and from alcohol intake to HDL [Barter, 1991]. We also started with correlations between sBP and dBP, and between cigarette consumption, alcohol intake, and exercise levels, where more explicit models cannot be decided upon.

To make the genetic models more wieldy, we chose only a subset of covariates. The first multivariate genetic model therefore included the five risk factors, BMI, and age. We first fitted an "intra-class" LISREL model which constrains variances to be the same in the first and second twin of each pair in both zygosity groups, and covariances between different traits within each twin pair to be equal within each zygosity group separately. This model fitted adequately ($\chi^2_{105} = 112.8$, $p = 0.28$). Examination of residuals found that most of the lack of fit was due to inequalities of within-pair cross-trait covariances.

This suggested that a number of assumptions of the twin method were not violated and that model fitting to correlation matrices was appropriate.

Prior to attempting a genetic model for all seven selected variables, we fitted a four variable model that included HDL, LDL, BMI, and age. A good fitting genetic model ($\chi^2_{34} = 28.0, p = 0.75$) based on the phenotypic path model was found (Fig. 1). Dropping either of the reciprocal paths between HDL and LDL led to a significant deterioration in goodness-of-fit.

We then fitted the full Choleski model to HDL, LDL, sTG, sBP, dBP, and BMI, with age included as a single factor. The goodness-of-fit of this ACE Choleski model was tolerable ($\chi^2_{95} = 111.0, p = 0.12$).

This model was further examined by varimax rotation of the factor loadings (the results of which were not greatly different in structure), and simplified by setting small loadings to zero (Table I).

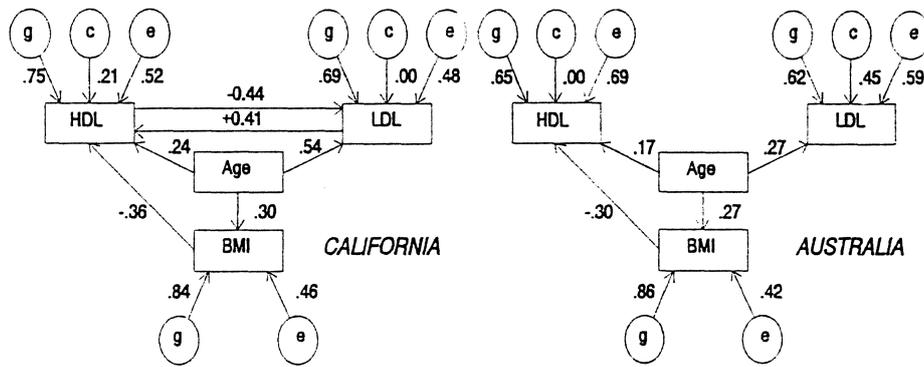


Fig. 1 Trivariate path models for HDL, LDL, BMI with age as a covariate for US and Australian female twins.

TABLE I. Loadings ($\times 100$) on Additive Genetic Factors in Simplified Choleski Genetic Factor Model for Kaiser-Permanente Female Twin Atherosclerosis Risk Factors

	Additive genes					Shared env.			Unshared environment					
	I	II	III	IV	V	I	II	III	I	II	III	IV	V	VI
BMI	80					24			46					
sBP	11	28					-43		18	64				
dBP	21	33	09				-26	36	17	40	58			
HDL	-37		-16	67		23		-9	-16			53		
LDL	20	34	52				14	22					48	
sTG	25	46		-39	25		19		10					51

The Choleski model has only limited explanatory power, so we were interested if we could create a full sized model based on the phenotypic path model. This model (Fig. 2) did not fit the data as well as the Choleski model ($\chi^2_{129} = 164.2$, $p = 0.02$), but makes falsifiable predictions about relationships between physiological variables under experimental manipulation.

Australian Twins

The Australian twin data set contained complete data from 65 pairs of MZ and 67 pairs of DZ like-sex female twins (age 17-60 yrs, mean 37 yrs). Unfortunately, blood pressure and sTG were not measured. However, data on HDL subtypes was included, so that there were independent measures of LDL, total HDL, HDL₂, and HDL₃. Phenotypic correlations are shown in Table II.

We obtained a good fit ($\chi^2_{35} = 29.5$, $p = 0.73$) for the same genetic model fitted to HDL, LDL, BMI, and age, though the direct HDL-LDL paths were not required (Fig. 1). A model including HDL subtypes (HDL₂ and HDL₃) was then elaborated. The "ACE" Choleski model gave a goodness-of-fit $\chi^2_{95} = 129.7$ (AIC = -60). A genetic model based on biologically plausible phenotypic paths (Fig. 2) led to an improvement in AIC to -86.5 ($\chi^2_{132} = 178.9$). The most interesting finding is that covariation between the HDL subfractions was mainly environmental.

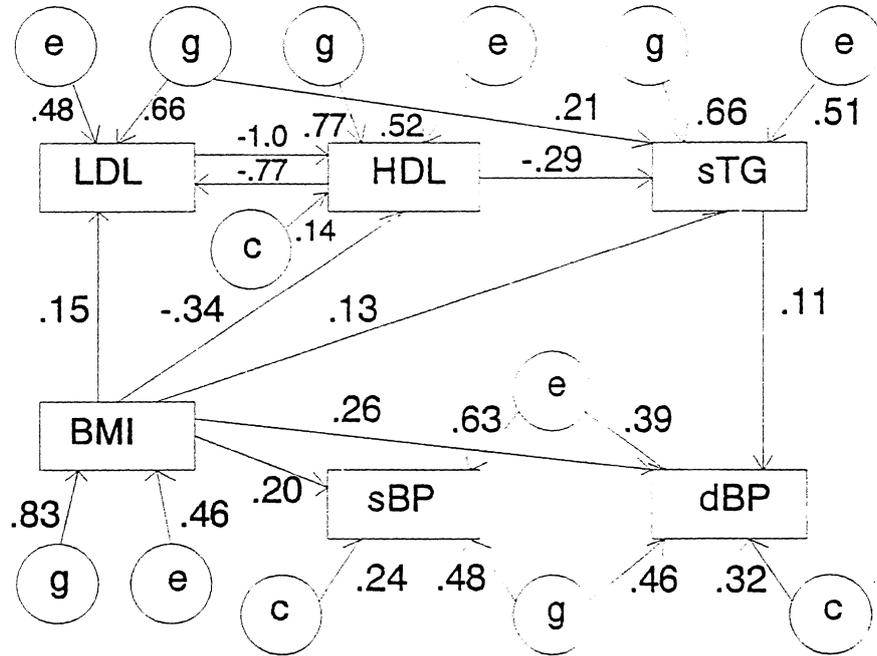


Fig 2. Non-factor genetic model for Kaiser-Permanente women twins. Age regression not shown: standardized regression coefficient for BMI 0.29; sBP 0.47; dBP 0.20; HDL 0.23; LDL 0.50; sTG 0.47.

TABLE II. Phenotypic and Univariate Within-pair Correlations (and Standard Errors) from Australian Data Set

	AGE	BMI	CHOL	LDL	HDL	HDL2	HDL3	r _{MZ}	r _{DZ}
AGE	1.00								
BMI	0.26 (0.07)	1.00						0.81 (0.13)	0.40 (0.11)
CHOL	0.39 (0.06)	0.19 (0.07)	1.00					0.83 (0.13)	0.50 (0.10)
LDL	0.27 (0.06)	0.17 (0.07)	0.79 (0.03)	1.00				0.86 (0.12)	0.32 (0.12)
HDL	0.16 (0.07)	-0.21 (0.06)	0.08 (0.07)	-0.05 (0.06)	1.00			0.58 (0.12)	0.24 (0.09)
HDL2	0.14 (0.07)	-0.17 (0.06)	0.14 (0.07)	-0.02 (0.07)	0.71 (0.03)	1.00		0.69 (0.11)	0.47 (0.18)
HDL3	0.19 (0.07)	-0.09 (0.07)	0.21 (0.07)	0.19 (0.06)	0.39 (0.05)	0.15 (0.06)	1.00	0.64 (0.10)	0.59 (0.18)

DISCUSSION

We have fitted two types of multivariate genetic path models in the present paper. The first type is the factor (e.g., Choleski decomposition) models. These are more equivalent to conventional exploratory (phenotypic) factor analyses, in that no particular causal relationships between variables are posited, aside from the usual genetic (and environmental) ones inherent in the classical twin design. The resulting genetic factors represent genetic correlations between traits.

The other type of model incorporates existing scientific knowledge about covariation and causation of the traits of interest. It has the disadvantage that existing biochemical/physiological/epidemiologic models may be incorrect. It is also more sensitive to errors of measurement of variables, as it is akin to genetic "direction-of-causation" modelling [Duffy and Martin, in press]. That is, attenuation of true associations between traits can actually lead to incorrect genetic models fitting the observed data better than the correct model. It does makes explicit predictions about relationships between traits.

The interpretation of the factors extracted in the Choleski models for the Kaiser-Permanente data is fairly straightforward. The first genetic component is a body mass/overweight factor which, as it increases BMI, also increases blood pressure and lowers HDL. The second genetic component represents further genetic correlations between blood pressure and lipid risk factors. The third, fourth, and fifth are specific genetic factors for HDL, LDL, and sTG, though a further correlation between HDL and sTG levels is seen. The negative correlations between HDL and LDL, and HDL and sTG are consistent throughout.

The second model reaches different conclusions from different priors. The genetic and environmental correlations between LDL and HDL are mediated

phenotypically through physiological mechanisms induced either via changes in BMI, a cause of change in both, or by hypothetical feedback mechanisms in the metabolism of HDL and LDL, which each have specific genetic controls.

Because the full Australian model differs in the variables included, it cannot be profitably compared to the Kaiser-Permanente results. We can examine the trivariate models (HDL, LDL, BMI, age as covariate). These differ first in that the negative correlation between HDL and LDL is not significant in the Australian data. This means that phenotypic paths from HDL to LDL were discarded. In both data sets a direct path from BMI to HDL, but not from BMI to LDL, was required. The heritabilities of all three variables are approximately the same in US and Australian samples (for HDL 0.56 and 0.42; LDL 0.48 and 0.38; BMI 0.71 and 0.74). The role of shared environment for LDL differed, not being required in the US data, but explaining 20% of the variation in LDL in the Australian females. The effects of age on LDL also differed slightly, from 29% of standardized LDL variance in the US to 8% in the Australians.

We can conclude by saying that covariation between LDL and HDL seems to be largely mediated either by genes that influence BMI or by BMI itself. The same is largely true for the relationship between blood pressure and lipids, aside from the conclusion that once a direct path from sTG to dBP is posited, then a reciprocal path is not needed. We find that the broad pattern of determination of these risk factors seems similar in samples from two populations of European descent.

REFERENCES

- Austin MA, King M-C, Bawol RD, Hulley SB, Friedman GD (1987): Risk factors for coronary heart disease in adult female twins. *Am J Epidemiol* 125:308-318.
- Barter PJ (1991): High density lipoproteins and coronary heart disease. *Aust NZ J Med* 21:299-231.
- Clay MA, Newnham HH, Barter PJ (1991): Hepatic lipase promotes the loss of apolipoprotein A-I from triglyceride-enriched human high density lipoproteins during incubation *in vitro*. *Arterioscler Thromb* 11:415-422.
- Duffy DL, Martin NG (in press). Inferring the direction of causation in cross-sectional twin data: Theoretical and empirical considerations. *Genet Epidemiol*.
- Jinks JA, Fulker DW (1970): Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behaviour. *Psychol Bull* 73:311-349.
- Jöreskog KG, Sörbom D (1990): LISREL 7. A Guide to the Program and Applications. Chicago: SPSS Inc.
- Krauss RM (1991): The tangled web of coronary risk factors. *Am J Med* 90(Supp 2A):36S-41S.
- Norusis MJ, SPSS Inc. (1988): SPSS/PC+ V2.0 Advanced Manual. Chicago: SPSS Inc.
- Neale MC, Cardon LR (1992): "Methodology for Genetic Studies of Twins and Families." Dordrecht: Kluwer Academic Publishers BV.
- O'Connell DL, Heller RF, Roberts DCK, Allen JR, Knapp JC, Steele PL, Silove D (1988): Twin study of genetic and environmental effects on lipid levels. *Genet Epidemiol* 5:323-342.
- Reaven GM (1991) Insulin resistance, hyperinsulinemia, and hypertriglyceridemia in the aetiology and clinical course of hypertension. *Am J Med* 90(Supp 2A):7S-12S.
- SAS Institute Inc. (1990): The SAS System, Version 6. Cary, NC: SAS Institute Inc.