Am J Hum Genet 35:1196–1210, 1983

Multivariate Models for Human Genetic Analysis: Aggregation, Coaggregation, and Tracking of Systolic Blood Pressure and Weight

CRAIG L. HANIS,[1,2] CHARLES F. SING,[1] WILLIAM R. CLARKE,[3] AND HELMUT G. SCHROTT[3]

SUMMARY

A multivariate path model parameterizing the sources of familial aggregation and coaggregation of systolic blood pressure and weight, as well as their tracking across time, is applied to longitudinal data collected in Muscatine, Iowa. Genetic, common household, and individual environmental effects, pleiotropy, and a direct regression effect of blood pressure on weight are parameterized. The sample consisted of 998 individuals distributed in 261 families of whom 601 were measured on four successive occasions. The data were divided with times 1 and 2 forming group 1, and times 3 and 4, group 2. Model fitting and estimation was performed using group 1, followed by testing the model and estimates using the data in group 2. Heritability estimates for systolic blood pressure and weight were .15 and .54, respectively. The genetic correlation between these traits was nonsignificant, but there was a significant direct regression effect. The results indicate that 30% of the full-sib correlation for systolic blood pressure is attributable to the aggregation of weight. In terms of tracking, 59% and 60% of the predicted systolic blood pressure and weight correlations, respectively, were attributable to genetic effects. Testing the model from group 1 in group 2 indicates that the qualitative relationships between blood pressure and weight are stable with time.

Received September 21, 1982; revised April 26, 1983.

This work was supported in part by predoctoral training grant 1-T32-GM07554-04, D.O.E. contract E (2828), grants 74-G-28 and 75-G-26 from the Iowa Heart Association, grant NIH-NHLBI-N01-HV-2-2913-L from the Lipid Research Clinics Program, and grants HL-14230 and HL-20124 from SCOR: Atherosclerois, Lipids, and Thrombosis.

[1] Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109.

[2] Current address: Center for Demographic and Population Genetics, University of Texas Health Science Center at Houston, Houston, TX 77025.

[3] Department of Preventive Medicine, University of Iowa College of Medicine, Iowa City, IA 52242.

INTRODUCTION

The familial aggregations of blood pressure and weight (each considered separately) have been well documented in population studies that have included several age groups (for blood pressure, see [1–5]; for weight, see [6, 7]). A consistently significant correlation between blood pressure and weight has also been demonstrated [8–12]. Furthermore, it has been shown that blood pressure and weight track over time [13], where tracking refers to the correlation within individuals of measures of a trait over time. While several investigators estimated the genetic and environmental contributions to familial aggregation of each of these traits, there have been no studies that systematically investigate the causes of correlation between blood pressure and weight and that assess the extent to which the aggregation of weight can explain the aggregation of blood pressure. Nor have there been any studies that have separated the tracking correlations of these traits into genetic and environmental components. To address these questions, we have applied a class of multivariate models recently developed by Hanis and Sing [14] to longitudinal data collected in Muscatine, Iowa. Here we examine only the bivariate case and consider the aggregation, coaggregation, and tracking of systolic blood pressure and weight.

Studies to date have demonstrated that genetic and environmental factors both contribute significantly to the separate familial aggregations of blood pressure and weight. Annest et al. [15] reported a heritability for systolic blood pressure in a sample from Montreal of .34 and the proportion of variability attributable to common household environmental variability to be .11. In terms of the correlations between full-sibs, this translates to 61% of the correlation being attributable to shared genes and 39% to shared environments. Other published heritabilities for systolic blood pressure include: .41 in a sample from northeastern Brazil [16] and .24 in Japanese-Americans [17], and heritabilities of .12 to .28 in Tokelau Islanders [18]. Similarly for weight, Annest et al. [7] report a heritability of .43 in Montreal and Rao et al. [19] report .44 in pedigrees from northeastern Brazil. We begin by accepting these studies as prior evidence that both genetic and environmental factors contribute to the separate aggregations of blood pressure and weight, and now turn to a consideration of factors that lead to their coaggregation in households and tracking over time.

Hanis and Sing [14] modeled coaggregation and tracking by parameterizing the correlation between two traits within and across time in terms of: (1) the pleiotropic action of genes, that is, genes that jointly affect weight and blood pressure; (2) environments that affect both traits; and (3) a direct regression relationship between the phenotypes of the two traits. We hypothesize that each of these sources may contribute to the correlation between blood pressure and weight. Such a parameterization assumes a class of mechanisms that determine the direct regression relationship between blood pressure and weight irrespective of the correlated genotypic and environmental effects; that is, those genetic and environmental effects contributing to the regression act on weight directly and only secondarily contribute to blood pressure variability through the regression. In a recent review of the association between obesity and hypertension, Dustan [20] suggested several mechanisms that could cause a directional relationship

(measured by regression) between blood pressure and weight throughout their distributions. The mechanisms proposed include the following possible alterations associated with higher weights: (1) increased cardiac output, (2) increased blood volume, (3) dietary differences across the weight distribution (particularly sodium intake), (4) changes in steroid production, and (5) alterations in receptors associated with various pressor substances. Other effects may include: increased capillary resistance and increased vascularization associated with higher weights. Through such a regression effect, variability in genetic and environmental factors that affect weight only are "translated" into sources of variability of blood pressure in addition to those factors that directly affect blood pressure. Thus, the familial aggregation of weight may contribute to the aggregation of blood pressure even in the absence of a genetic correlation and/or environments shared by individuals within households that affect both traits.

MATERIALS AND METHODS

Sample

The data used in this study represent a subset of that collected as part of the "Muscatine Study," an ongoing longitudinal survey beginning in 1970 of cardiovascular risk factors in children aged 5 to 19 attending school in Muscatine, Iowa. Among other variables, systolic blood pressure and weight were determined at approximately 2-year intervals. Detailed demographic information about the reference population has been published [21]. Clarke et al. [13], Lauer et al. [22], and Schrott et al. [23] summarized the measurement procedures employed, have given a statistical description of the variables measured, estimated the degree of tracking of variables, and related risk factor levels in children to cardiovascular disease experience in adult relatives.

The Muscatine Study has served further to identify individuals as index cases for the Muscatine Hyperlipidemia Family Study (MHFS). We used the pedigree information obtained in the course of the MHFS to identify the sample of genetically related children that had participated in the longitudinal study. Because the longitudinal study was limited to school-age children, the sample obtained included very few second-degree relatives (uncles, aunts, nephews, and nieces of the index cases). Consequently, we limited our sample to families consisting of full-sibs (residing in the same household) and first cousins. Those included had been measured on up to nine different occasions (some individuals were included in random recalls or were otherwise measured more frequently than 2-year intervals), but only results from the first four examinations (hereafter referred to as times 1 through 4) for each individual were used because of the small number measured more than four times. The sample used consisted of 998 individuals distributed in 261 families at time 1, of which 601 were measured on all four occasions. We assume that ascertainment of these families on the basis of an index case's lipid phenotype was independent of the distribution of weight and blood pressure in these families.

Model

The model used to address the sources of coaggregation and tracking of systolic blood pressure and weight is given in two parts by the path diagrams presented in figure 1. In both diagrams, the measured phenotypes, y_{ij} and x_{ij}, represent systolic blood pressure and weight, respectively. The subscript, i, identifies an individual as either an index (I), sib (S), or cousin (C), while the second subscript, j, denotes time. Note that under the assumption that ascertainment is independent of weight and blood pressure, the assignment of an index case is purely arbitrary. In the diagrams, only two times are represented for an index and sib pair. In the upper diagram, additive genetic effects (G_y and G_x) and residual effects (E_y and E_x) are parameterized with appropriate path coefficients as indicated alongside

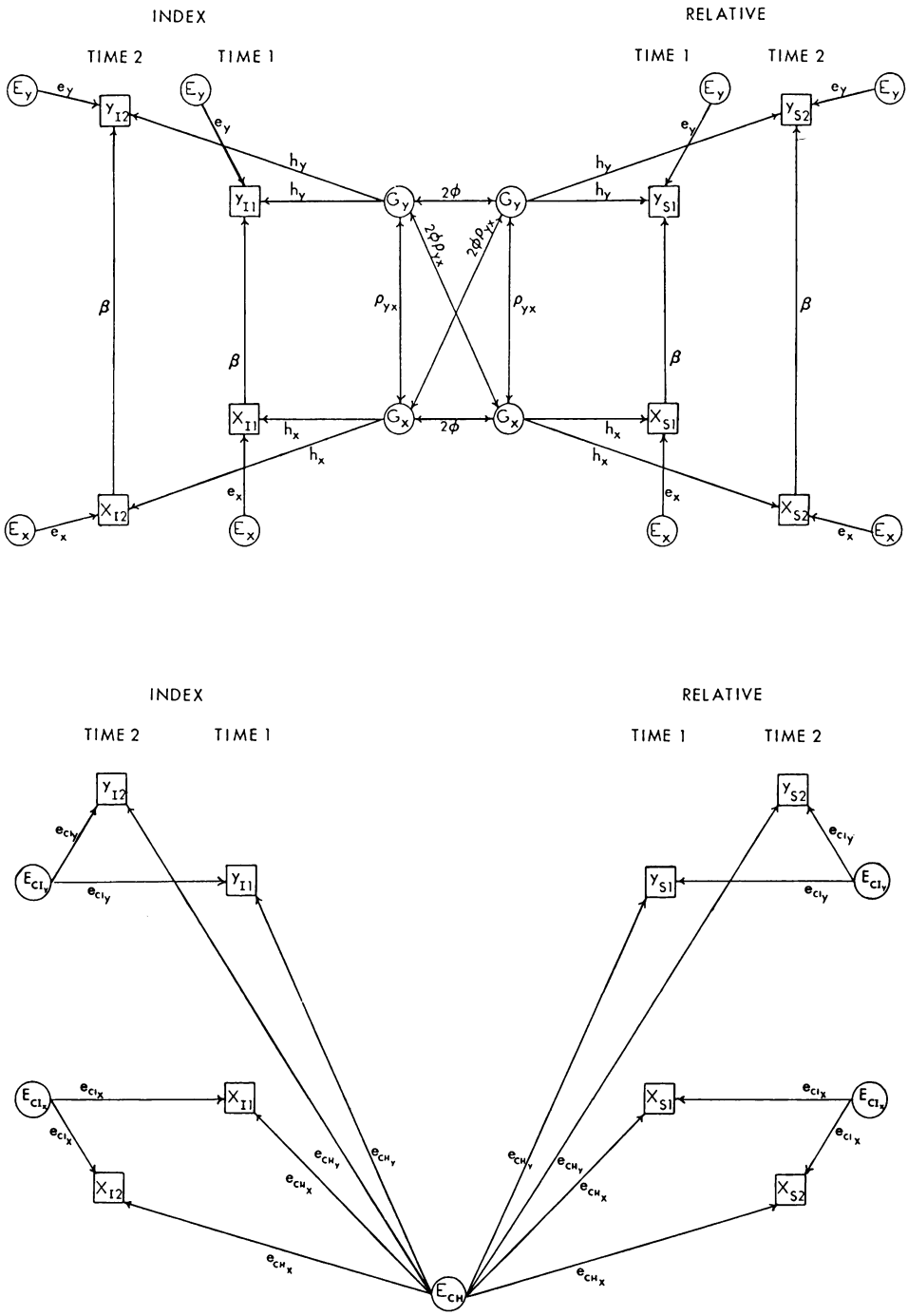

Fig. 1.—Path diagrams for the correlation between y and x in an index and relative for a longitudinal sample. *Upper diagram* presents the genetic, regression, and residual sources of variability, while *lower diagram* parameterizes the household and individual environmental effects.

the paths leading to the measured phenotypes. The direct regression effect is represented by β, and the correlation of additive gene effects is indicated by ρ_{yx}. In the lower diagram are household and individual environmental effects. E_{CH} represents those environments that are shared by all individuals within a household and that vary randomly among households. E_{CH} is constant in time and affects both blood pressure and weight, but the magnitude of the effect on each is measured by a different path coefficient. E_{CI_y} and E_{CI_x} represent environments that are unique to an individual yet are constant through time. We have assumed that E_{CI_y} and E_{CI_x} are uncorrelated.

When the two diagrams are combined into a "complete" model, one can obtain expressions for the expectations of all possible correlations in terms of eight parameters (a total of 38 correlations when considering sibs and cousins at two times). Many of the 38 correlations have equivalent expectations. In fact, each correlation falls into one of 10 equivalence classes as detailed in table 1. In table 2, the expectations of these correlation classes are given. Although other bivariate path models have been presented for multivariate data [24–26], none have dealt with longitudinal data and the parameterization of a direct regression effect as defined here.

Estimation

As reviewed in Hanis and Sing [14], estimation of the parameters in path analysis generally, although not necessarily, employs two steps. First, the correlations between measured phenotypes are estimated. Second, the relationships of these estimates to their expectations (table 2) are submitted to a numerical analysis to obtain estimates of the parameters. Correlations within individuals (between variables and across time) are estimated as product-moment correlations, while all sib-sib correlations are calculated as intraclass correlations. Cousin correlations were estimated as product-moment correlations based on a random pair from each family. The minimum chi-square procedure outlined in Annest et al. [5] is used to obtain parameter estimates from the set of equations that relate the correlation statistics to their expectations.

TABLE 1

EQUIVALENCE OF EXPECTED CORRELATIONS (y = SYSTOLIC BLOOD PRESSURE, x = WEIGHT)

			Class
$E[cor(y_{I1},y_{I2})]$	= $E[cor(y_{C1},y_{C2})]$	.....	(1)
$E[cor(y_{I1},y_{S1})]$	= $E[cor(y_{I1},y_{S2})]$	= $E[cor(y_{I2},y_{S2})]$	(2)
$E[cor(y_{I1},y_{C1})]$	= $E[cor(y_{I1},y_{C2})]$	= $E[cor(y_{I2},y_{C1})]$	
	= $E[cor(y_{I2},y_{C2})]$	.....	(3)
$E[cor(y_{I1},x_{I1})]$	= $E[cor(y_{I2},x_{I2})]$	= $E[cor(y_{C1},x_{C1})]$	
	= $E[cor(y_{C2},x_{C2})]$	.....	(4)
$E[cor(y_{I1},x_{I2})]$	= $E[cor(y_{I2},x_{I1})]$	= $E[cor(y_{C1},x_{C2})]$	
	= $E[cor(y_{C2},x_{C1})]$	.....	(5)
$E[cor(y_{I1},x_{S1})]$	= $E[cor(y_{I1},x_{S2})]$	= $E[cor(y_{I2},x_{S1})]$	
	= $E[cor(y_{I2},x_{S2})]$	.....	(6)
$E[cor(y_{I1},x_{C1})]$	= $E[cor(y_{I1},x_{C2})]$	= $E[cor(y_{I2},x_{C1})]$	
	= $E[cor(y_{I2},x_{C2})]$	= $E[cor(x_{I1},y_{C1})]$	
	= $E[cor(x_{I1},y_{C2})]$	= $E[cor(x_{I2},y_{C1})]$	
	= $E[cor(x_{I2},y_{C2})]$	.....	(7)
$E[cor(x_{I1},x_{I2})]$	= $E[cor(x_{C1},x_{C2})]$	.....	(8)
$E[cor(x_{I1},x_{S1})]$	= $E[cor(x_{I1},x_{S2})]$	= $E[cor(x_{I2},x_{S2})]$	(9)
$E[cor(x_{I1},x_{C1})]$	= $E[cor(x_{I1},x_{C2})]$	= $E[cor(x_{I2},x_{C1})]$	
	= $E[cor(x_{I2},x_{C2})]$	.....	(10)

TABLE 2

FORMULAS FOR EXPECTED CORRELATIONS (y = SYSTOLIC BLOOD PRESSURE, x = WEIGHT)

	Class
$E[\text{cor}(y_{I1},y_{I2})] = h_y^2 + e_{CI_y}^2 + e_{CH_y}^2 + \beta^2 h_x^2 + \beta^2 e_{CI_x}^2 + \beta^2 e_{CH_x}^2 + 2\beta h_y \rho_{yx} h_x + 2\beta e_{CH_y} e_{CH_x}$	(1)
$E[\text{cor}(y_{I1},y_{S1})] = \frac{1}{2}h_y^2 + e_{CH_y}^2 + \frac{1}{2}\beta^2 h_x^2 + \beta^2 e_{CH_x}^2 + 2\beta e_{CH_y} e_{CH_x} + \beta h_y \rho_{yx} h_x$	(2)
$E[\text{cor}(y_{I1},y_{C1})] = \frac{1}{8}h_y^2 + \frac{1}{8}\beta^2 h_x^2 + \frac{1}{4}\beta h_y \rho_{yx} h_x$	(3)
$E[\text{cor}(y_{I1},x_{I1})] = \beta + h_y \rho_{yx} h_x + e_{CH_y} e_{CH_x}$	(4)
$E[\text{cor}(y_{I1},x_{I2})] = \beta(e_{CH_x}^2 + e_{CI_x}^2) + h_y \rho_{yx} h_x + e_{CH_y} e_{CH_x} + \beta h_x^2$	(5)
$E[\text{cor}(y_{I1},x_{S1})] = \frac{1}{2}h_y \rho_{yx} h_x + e_{CH_y} e_{CH_x} + \frac{1}{2}\beta h_x^2 + \beta e_{CH_x}^2$	(6)
$E[\text{cor}(y_{I1},x_{C1})] = \frac{1}{8}h_y \rho_{yx} h_x + \frac{1}{2}\beta h_x^2$	(7)
$E[\text{cor}(x_{I1},x_{I2})] = h_x^2 + e_{CH_x}^2 + e_{CI_x}^2$	(8)
$E[\text{cor}(x_{I1},x_{S1})] = \frac{1}{2}h_x^2 + e_{CH_x}^2$	(9)
$E[\text{cor}(x_{I2},x_{C1})] = \frac{1}{8}h_x^2$	(10)

Analytical Strategy

The availability of data measured on four different occasions provides considerable flexibility in choosing an analytical strategy. Our first step was to adjust for concomitant variability using standard analysis of covariance techniques [27]. Age, sex, and examination number were treated as concomitants. The ascertainment group to which each family was assigned as part of the MHFS was also treated as a concomitant. Families were assigned to an ascertainment group based on the lipid level criteria specified in the MHFS (e.g., twice high cholesterol, middle cholesterol, high triglycerides, or low triglycerides). Following adjustment for concomitants, we divided the data into two sets (times 1 and 2 comprising set 1 and times 3 and 4 comprising set 2), calculated the 38 observed correlations for each set, and followed the three steps detailed below. First, a test of the homogeneity of the correlations within each equivalence class (table 2) was carried out separately for the two sets of data and a pooled estimate for each class obtained. Inasmuch as the equivalence relationships are defined by the model in figure 1, these homogeneity tests provided an initial test of whether the parameterization of the equivalence class was consistent with the data. Heterogeneity within an equivalence class would indicate either random variability (in which case one would not expect the pattern of heterogeneity to be repeated in the second set of data) or it would indicate a departure from the assumptions under which the model was formulated. In the latter case, it would be necessary to alter one or more of the assumptions and reformulate the model accordingly. The second step of the analysis consisted of estimating the parameters of the complete model and a selected subset of reduced models using the 10 pooled correlations from times 1 and 2. To test whether one or a combination of parameters was significantly different from zero, we used the difference between the chi-square measure of poorness of fit of the reduced model and the chi-square measure of poorness of fit of the complete model as discussed by Hanis [28]. Based on these results, a parsimonious model was chosen. A parsimonious model was defined as having the fewest parameters (when compared to the complete model or other alternatives), yet which did not fit the data significantly poorer than the complete model. Third, we tested the hypothesis that the model from times 1 and 2 could explain

the data from times 3 and 4. For this comparison, we set the heritabilities for the analysis of times 3 and 4 at those obtained from times 1 and 2 and allowed all other parameters to float.

RESULTS

Table 3 summarizes the descriptive statistics for the unadjusted age, weight, and systolic blood pressure data at the four examinations considered in this study. The average age, in years, increases from 9.2 years at time 1 to 13.3 years at time 4 by increments of 1.7, 1.3, and 1.1 years, respectively. That these are not 2-year intervals is due to the study's loss of older individuals who became ineligible for further participation through graduation from high school. Coupled with this is the absence of an influx of new individuals at the lower ages who were subsequently measured on four occasions. Overall, the data appear consistent with a group of children undergoing growth as indicated by mean changes with time in the variables that reflect size. The pattern of mean changes is similar to that observed in the United States Health Survey [29]. There is no general pattern of change in the blood pressure means across time, nor are there any obvious trends or inconsistencies in the measures of variation, skewness, and kurtosis.

The relationships among systolic blood pressure, weight, and the two variables on which ascertainment was based are summarized in the correlation matrices of table 4. These correlations are based on individuals within each time period that have all four variables measured. Correlations greater than .08 at time 1 (no. = 962) are significant at the .01 level, while for times 2–4 (nos. = 842, 706, and 491, respectively), the .01 significance levels are .09, .11 and .13, respectively. Of particular relevance to the analysis presented here is the consistency of the correlation between systolic blood pressure and weight within each time: .60, .59, .63, and .57, respectively. This indicates that even though the mean of each variable is increasing with time, there is little evidence to suggest that the relationship between these variables is not homogeneous across time. As seen, the

TABLE 3

DESCRIPTIVE STATISTICS OF THE UNADJUSTED DATA: TOTAL SAMPLE

	No.	Mean	SD	Skewness	Kurtosis
Age (yr):					
Time 1	998	9.24	2.639	0.524	−0.463
Time 2	930	10.89	2.880	0.418	−0.681
Time 3	805	12.19	2.622	0.208	−0.814
Time 4	601	13.27	2.441	0.050	−0.800
Weight (lb):					
Time 1	993	76.62	29.446	1.191	1.205
Time 2	875	92.34	34.705	0.889	0.445
Time 3	765	105.80	35.552	0.827	0.790
Time 4	599	117.65	35.232	0.701	1.023
Systolic blood pressure (mm Hg):					
Time 1	992	106.97	13.801	0.500	1.139
Time 2	785	108.18	13.804	0.514	0.939
Time 3	768	109.19	13.495	0.446	0.201
Time 4	598	112.12	13.074	0.497	0.584

TABLE 4

CORRELATION STATISTICS: UNADJUSTED DATA

	TIME 1			TIME 2			TIME 3			TIME 4		
	Weight	SBP*	Cholesterol	Weight	SBP	Cholesterol	Weight	SBP	Cholesterol	Weight	SBP	Cholesterol
SBP	.60	...	...	.59	...	...	.63	...	...	.57	...	...
Cholesterol	.10	.13	...	−.09	.02	...	−.08	.01	...	−.20	−.05	...
Triglycerides	.30	.24	.27	.24	.20	.21	.14	.10	.24	.23	.20	.29

* SBP = systolic blood pressure.

correlations of weight and systolic blood pressure with cholesterol and triglycerides are not homogeneous in time, and although their magnitudes are generally small, several are significantly different from zero. That the correlation between blood pressure and weight is stable across time while their correlations with cholesterol and triglycerides are unstable is taken as justification for the assumption that ascertainment on the latter did not affect the relationship between blood pressure and weight. Ascertainment of the families was not limited to only one tail of the cholesterol or triglycerides distribution, but, rather, involved sampling throughout their distributions, which further supports the assumption that ascertainment was independent of blood pressure and weight.

The primary variables of interest in this analysis, systolic blood pressure and weight, were adjusted by covariance analysis to remove the effects of age, age^2, age^3, sex, time (i.e., examination number), and ascertainment class. All two-way interaction effects were also removed and the residuals obtained used for all subsequent analysis. The concomitants accounted for 69% and 28% of the variability of weight and systolic blood pressure, respectively. Several transfor-

TABLE 5

CORRELATION STATISTICS CALCULATED FROM THE ADJUSTED DATA OF TIMES 1 AND 2 (y = SYSTOLIC BLOOD PRESSURE, x = WEIGHT)

		No.			No.
Interclass correlations					
$r(y_{I1},y_{I2})$	=	.4044	513	$r(y_{I2},x_{I1})$ =	.3618 514
$r(y_{C1},y_{C2})$	=	.2274	353	$r(y_{C2},x_{C1})$ =	.1974 353
$r(y_{I1},x_{I1})$	=	.3672	577	$r(y_{I2},x_{I2})$ =	.4125 515
$r(y_{C1},x_{C1})$	=	.3387	410	$r(y_{C2},x_{C2})$ =	.3179 355
$r(y_{I1},x_{I2})$	=	.3305	513	$r(x_{I1},x_{I2})$ =	.9123 514
$r(y_{C1},x_{C2})$	=	.2639	353	$r(x_{C1},x_{C2})$ =	.8575 353
Intraclass correlations—within sibships					
$r(y_{I1},y_{S1})$	=	.1567	577	$r(y_{I2},x_{S1})$ =	.1932 514
$r(y_{I2},y_{S2})$	=	.2485	516	$r(y_{I2},x_{S2})$ =	.1730 515
$r(y_{I1},y_{S2})$	=	.2197	513	$r(x_{I1},x_{S1})$ =	.3035 578
$r(y_{I1},x_{S1})$	=	.1682	577	$r(x_{I1},x_{S2})$ =	.3321 514
$r(y_{I1},x_{S2})$	=	.1250	513	$r(x_{I2},x_{S2})$ =	.3445 516
Interclass correlations—cross relatives (i.e., sib-cousin)					
$r(y_{I1},y_{C1})$	=	−.0023	88	$r(x_{I1},y_{C1})$ =	.0229 88
$r(y_{I1},y_{C2})$	=	−.0175	84	$r(x_{I1},y_{C2})$ =	−.1377 84
$r(y_{I2},y_{C1})$	=	−.0453	85	$r(x_{I2},y_{C1})$ =	−.1675 84
$r(y_{I2},y_{C2})$	=	.0881	81	$r(x_{I2},y_{C2})$ =	−.0633 79
$r(y_{I1},x_{C1})$	=	−.0402	88	$r(x_{I1},x_{C1})$ =	.0440 88
$r(y_{I1},x_{C2})$	=	.1961	84	$r(x_{I1},x_{C2})$ =	.1538 84
$r(y_{I2},x_{C1})$	=	.1501	85	$r(x_{I2},x_{C1})$ =	.1583 84
$r(y_{I2},x_{C2})$	=	.2564	81	$r(x_{I2},x_{C2})$ =	−.0106 80

TABLE 6

CORRELATION STATISTICS CALCULATED FROM THE ADJUSTED DATA OF TIMES 3 AND 4 (y = SYSTOLIC BLOOD PRESSURE, x = WEIGHT)

		No.			No.
Interclass correlations					
$r(y_{I1},y_{I2})$	=	.4270	357	$r(y_{I2},x_{I1})$ =	.4072 355
$r(y_{C1},y_{C2})$	=	.4812	203	$r(y_{C2},x_{C1})$ =	.3997 203
$r(y_{I1},x_{I1})$	=	.4386	474	$r(y_{I2},x_{I2})$ =	.4732 387
$r(y_{C1},x_{C1})$	=	.3804	288	$r(y_{C2},x_{C2})$ =	.4312 209
$r(y_{I1},x_{I2})$	=	.4491	357	$r(x_{I1},x_{I2})$ =	.9194 355
$r(y_{C1},x_{C2})$	=	.3613	204	$r(x_{C1},x_{C2})$ =	.9269 204
Intraclass correlations—within sibships					
$r(y_{I1},y_{S1})$	=	.2470	476	$r(y_{I2},x_{S1})$ =	.1111 355
$r(y_{I2},y_{S2})$	=	.1658	387	$r(y_{I2},x_{S2})$ =	.1323 387
$r(y_{I1},y_{S2})$	=	.1180	357	$r(x_{I1},x_{S1})$ =	.3885 474
$r(y_{I1},x_{S1})$	=	.2052	474	$r(x_{I1},x_{S2})$ =	.3958 355
$r(y_{I1},x_{S2})$	=	.1895	357	$r(x_{I2},x_{S2})$ =	.4214 387
Interclass correlations—cross relatives (i.e., sib-cousin)					
$r(y_{I1},y_{C1})$	=	−.1120	74	$r(x_{I1},y_{C1})$ =	−.1064 74
$r(y_{I1},y_{C2})$	=	.0157	64	$r(x_{I1},y_{C2})$ =	.2232 64
$r(y_{I2},y_{C1})$	=	−.0737	69	$r(x_{I2},y_{C1})$ =	−.0389 69
$r(y_{I2},y_{C2})$	=	.0869	60	$r(x_{I2},y_{C2})$ =	.1021 60
$r(y_{I1},x_{C1})$	=	−.0204	74	$r(x_{I1},x_{C1})$ =	.0818 74
$r(y_{I1},x_{C2})$	=	.2400	64	$r(x_{I1},x_{C2})$ =	.1125 64
$r(y_{I2},x_{C1})$	=	.1747	69	$r(x_{I2},x_{C1})$ =	.2326 69
$r(y_{I2},x_{C2})$	=	.0337	60	$r(x_{I2},x_{C2})$ =	.2821 60

mations were made on each variable to study the effects of nonnormality on the outcome of the analysis. These transformations included a series of power transforms and an inverse normal transform based on ranks (see [30]). The within-individual, within-time correlation coefficients between systolic blood pressure and weight under these various transformations remained stable [28]. For example, the inverse normal transformation (INORM) removed all skew and kurtosis for each variable, yet the correlation between the untransformed and transformed systolic blood pressure and weight differed by only .0061. For this reason, we calculated all correlations based on the adjusted, but nontransformed data.

The 38 correlation coefficients for each set of data are given in tables 5 and 6. Unlike the correlations in table 4, these take into account the family and longitudinal structure of the sample and quantify the aggregation, coaggregation, and tracking of systolic blood pressure and weight. Significant heterogeneity (at the .01 level) was indicated for only two of the equivalence classes in table 1. These were the comparisons of the tracking correlations (within individuals across time) computed from siblings and those computed from cousins at times 1 and 2. Additionally,

the correlation between blood pressure and weight in cousins showed marginal heterogeneity at times 1 and 2. Because these effects were not detected in the data from times 3 and 4 and because of the number of tests performed, they were interpreted as random findings.

In table 7 we present the parameter estimates, their standard errors, and the chi-square poorness-of-fit criterion for the parsimonious model obtained from times 1 and 2. Also presented are the results of fitting this model to the data of times 3 and 4 (recall that the heritabilities were fixed in this analysis). In the parsimonious model, ρ_{yx} and $e_{Cl_y}{}^2$ have been eliminated (i.e., set to zero). The magnitude of the common household environmental parameters ($e_{CH_y}{}^2$ and $e_{CH_x}{}^2$) and their standard errors suggest that they might also be nonsignificant; however, testing each separately yields a 1 df chi-square (when compared with the parsimonious model) that is significant for each: 7.87 and 3.97, respectively.

As seen in table 7, the fit of the parsimonious model obtained to the data of times 3 and 4 (fixing the heritabilities) is significant at the .05 level. Although there is significant heterogeneity between the set of correlations estimated in set 1 and the set estimated for set 2, their rank ordering and the rank ordering and magnitudes of the parameters estimated for each set are very similar. Returning to the results from times 3 and 4, we note that the estimates of the heritabilities used (those from 1 and 2) were subject to rather large standard errors; hence, we did a grid search about these values using the times 3 and 4 data. Changing $h_y{}^2$ from .151 to .170 and keeping $h_x{}^2$ at .538 results in a nonsignificant test of poorness of fit of this model to the data of times 3 and 4. Thus, we conclude that while the quantitative relationships among variables have changed significantly, the qualitative relationships are the same as evidenced by the model from times 1 and 2 explaining the data at times 3 and 4. Note also the similarities in the magnitude of the parameter estimates and their rankings when comparing the two sets of data.

TABLE 7

PARAMETER ESTIMATES, STANDARD ERRORS, AND POORNESS OF FIT OF PARSIMONIOUS MODEL TO TIMES 1 AND 2 AND OF MODEL TO TIMES 3 AND 4 HOLDING HERITABILITY CONSTANT

Parameters	Times 1 and 2	Times 3 and 4
$h_y{}^2$	.151 ± .077	.151
$h_x{}^2$	.538 ± .177	.538
β	.285 ± .039	.395 ± .030
$e_{CH_y}{}^2$	.071 ± .056	.023 ± .018
$e_{CH_x}{}^2$	.061 ± .083	.111 ± .023
$e_{Cl_x}{}^2$	.293 ± .099	.273 ± .023
Poorness of fit:		
df	4	6
χ^2	1.73	14.67*

* $.01 < P < .05$.

We now turn to the interpretation of these results in terms of the proportion of the familial aggregation of systolic blood pressure that can be explained by the aggregation of weight and the contributions of genetic and environmental effects to the tracking of systolic blood pressure and weight across time. For this discussion, we consider only the results from times 1 and 2, although it is apparent that similar results are obtained from the times 3 and 4 data as well. The predicted full-sib correlation for systolic blood pressure is .21 (observed is .21), of which 30% is attributable to the aggregation of weight. Forty-six percent is attributable to shared genes. Of this 46%, 78% is directly attributable to genes affecting systolic blood pressure and 22% to genes that affect weight directly and blood pressure secondarily through the regression effect. Failure to consider the separate contribution of weight would result in overestimation of the direct effect of genetic variation on blood pressure variation.

In terms of tracking, 59% of the predicted tracking correlation of systolic blood pressure (.33) is attributable to genetic effects. Thirty-two percent is explained by weight. In a similar fashion, 60% of the predicted weight tracking across time (.89) is attributable to genetic effects and 40% to environmental effects.

DISCUSSION

This study represents the first application of a multivariate model that includes genetic, environmental, and regression sources of correlation and employs longitudinal family data. For this reason, we have discussed the major methodological considerations as we have developed the results. We now turn to a discussion of the implications of this research as it relates to the coaggregation, aggregation, and tracking of systolic blood pressure and weight. First, the interpretation of the analysis is model dependent. The model used here represents our best effort to parameterize mathematically the known biological relationships between blood pressure and weight. The model is consistent with biological information and provides an adequate explanation of the data. This, of course, does not imply that it is the only model having such properties. In spite of these caveats, however, the results obtained should prove valuable in: (1) establishing the utility of applying multivariate models that explicitly parameterize the sources of correlation among variables, (2) identification of sources of coaggregation and tracking, and (3) establishing directions for future work on the risk factors of common diseases.

The application of multivariate models is limited by the number and power of the contrasts in the data for estimating values and testing hypotheses about a multiplicity of parameters. For a given set of effects and an appropriate design, the inclusion of additional variables results in an increase in the number of parameters that must be estimated. For the bivariate analysis presented here, implementation and convergence of the estimation algorithm (a Newton-Raphson procedure) was relatively straightforward. As more variables are added with an accompanying increase in the dimension of the parameter space, it is not clear that such will continue to be the case. Even in this case, there were initial starting values that did not lead to convergence. For this reason, we agree with Lange and Boehnke [31] that this class of analysis will likely be most applicable to bivariate situations, at least with present algorithms. Also, the computational

power required for such expanded analyses may prove prohibitive. We believe that these models will prove useful in hypothesis testing and making inferences about the postulated relationships among variables and that such analyses are practical in at least longitudinal bivariate data.

The inferences about coaggregation, aggregation, and tracking of blood pressure and weight from the analysis presented here are of particular interest. Briefly, the results suggest three major points. First, the coaggregation of blood pressure and weight (defined by the sib-sib correlation and the within-individual-between-traits correlation) is explained in these data without inclusion of a genetic correlation. We must qualify this conclusion because the test for a genetic correlation determines whether it is significantly different from zero given the presence of a regression effect. Hanis and Sing [14] point out that these two parameters are parametrically correlated. An exact expression for the expected correlation between parameters has not been formulated. The computed correlation is $-.98$ when the complete model is fit to the data of times 1 and 2. Consequently, there may have been insufficient independent information in the data to reject the hypothesis; $H:\rho_{yx} = 0$, and inclusion of either the regression (β) or correlation (ρ_{yx}), will result in essentially the same results. However, testing the significance of β with ρ_{yx} in the model resulted in a significant chi-square (8.61 on 1 df), indicating that the regression effect is required in the final model. The second point is that a large percentage of the aggregation and tracking of systolic blood pressure can be explained by weight. The third and final point is that the tracking of variables measured by correlations across time can be partitioned into genetic and environmental components in the same fashion as the sources of aggregation and coaggregation.

Inferences from this study are limited to statements about the population from which the sample was drawn. This kind of information has utility in the development of public health measures that are designed to alter a risk-factor distribution. When genetic factors predominate in determining risk-factor variability, it is necessary to formulate specific intervention schemes that account for family structure in the population and that exploit the interaction of genes and environments. A relatively low genetic contribution implies that environmental and behavior modification may be applied without regard for genetic information in the population. This study extends these concepts further to prediction of a response that intervention on one trait will have on a second correlated trait. The study conducted here represents a valuable step in a hierarchy of questions that lead to a full understanding of the etiology of risk-factor variability and its association with disease endpoints [32]; that is, after establishing that a variable is a risk factor for some disease, it is then of interest to establish whether it aggregates in families or, as in this case, whether multiple traits coaggregate in families. If so, an apparent next step is to partition the sources of the aggregation and/or coaggregation into genetic and/or environmental components.

In conclusion, this study represents the first general treatment simultaneously of the aggregation, coaggregation, and tracking of two traits. The results indicate that a large proportion of the aggregation and tracking of systolic blood pressure is explained by weight. This implies that a study of systolic blood pressure

ignoring weight leads to estimates of the contributions of genetic and environmental variability to population varibility, which, in actuality, are largely measures of the contributions of these factors to weight variability. We would hope that the results presented will provide impetus for further research to identify and define specific causative factors that mediate the relationships between systolic blood pressure and weight. With the identification of such factors, inferences and prediction will shift from the populaion level to a level of identifying and intervening on high-risk individuals and families that have the potential for profoundly reducing the burden of common complex diseases.

REFERENCES

1. JOHNSON ES, EPSTEIN FH, KJELSBERG MO: Distributions and familial studies of blood pressure and serum cholesterol in a total community—Tecumseh, Michigan. *J Chronic Dis* 18:147–160, 1965
2. ZINNER SJL, LEVY PS, KASS EH: Familial aggregation of blood pressure in childhood. *N Engl J Med* 284:401–404, 1971
3. KLEIN BE, HENNEKINS CH, JESSE MJ, GOURLEY JE, BLUMENTHAL S: Longitudinal studies of blood pressure in offspring of hypertensive mothers, in *Epidemiology and Control of Hypertension*, edited by PAUL O, Chicago, Yearbook Medical, 1975, pp 387–395
4. ROSE RJ, MILLER JZ, GRIM CE, CHRISTIAN JC: Aggregation of blood pressure in the families of identical twins. *Am J Epidemiol* 109:503–511, 1979
5. ANNEST JL, SING CF, BIRON P, MONGEAU JG: Familial aggregation of blood pressure and weight in adoptive families. I. Comparisons of blood pressure and weight statistics among adoptive families with adopted, natural, or both natural and adopted children. *Am J Epidemiol* 110:479–491, 1979
6. BIRON P, MONGEAU JG, BERTRAND D: Familial resemblance of body weight and weight/height in 374 homes with adopted children. *J Pediatr* 91:555–558, 1977
7. ANNEST JL, SING CF, BIRON P, MONGEAU JG: Familial aggregation of blood pressure and weight in adoptive families. III. Estimation of the relative contributions of genes and common household environmental factors to the correlations between family members for weight, height and selected weight/height indices. *Am J Epidemiol* 117:492–506, 1983
8. CHIANG B, PERLMAN L, EPSTEIN F: Overweight and hypertension. *Circulation* 39:403–421, 1969
9. VOORS AW, WEBBER LS, FRERICHS RR, BERENSON GS: Body height and body mass as determinants of basal blood pressure in children—The Bogalusa Heart Study. *Am J Epidemiol* 106:101–108, 1977
10. KOTCHEN JM: Effect of relative weight on familial blood pressure aggregations. *Am J Epidemiol* 105:214–222, 1977
11. FEINLIEB M, GARRISON RJ, HAVLIK RJ: Environmental and genetic factors affecting the distribution of blood pressure in children, in *Childhood Prevention of Atherosclerosis and Hypertension*, edited by LAUER RM, SHEKELLE RB, New York, Raven Press, 1980, pp 271–290
12. HIGGENS MW, KELLER J, MOORE F, OSTRANDER L, METZNER H, STOCK L: Studies of blood pressure in Tecumseh, Michigan. I. Blood pressure in young people and its relationship to personal and familial characteristics and complications of pregnancy in mothers. *Hypertension* 111:142–155, 1980
13. CLARKE WR, SCHROTT HG, LEAVERTON PE, CONNOR WE, LAUER RM: Tracking of blood lipids and blood pressures in school age children: The Muscatine Study. *Circulation* 58:626–634, 1978
14. HANIS CL, SING CF: Multivariate models for human genetic analysis: development of path analysis and variance components models. Unpublished

15. ANNEST JL, SING CF, BIRON P, MONGEAU JG: Familial aggregation of blood pressure and weight in adoptive families. II. Estimation of the relative contributions of genetic and environmental factors to blood pressure correlations between family members. *Am J Epidemiol* 110:492–503, 1979
16. KRIEGER H, MORTON NE, RAO DC, AZEVEDO E: Familial determinants of blood pressure in northeastern Brazil. *Hum Genet* 53:415–418, 1980
17. MORTON NE, GULBRANDSEN C, RAO DC, RHOADS G, KAGAN A: Determinants of blood pressure in Japanese-American families. *Hum Genet* 53:261–266, 1980
18. WARD RH, CHIN PG, PRIOR IAM: Tokelau Island Migrant Study: effect of migration on the familial aggregation of blood pressure. *Hypertension* 2(Suppl. I):I43–I54, 1980
19. RAO DC, MACLEAN CJ, MORTON NE, YEE S: Analysis of family resemblance. V. Height and weight in northeastern Brazil. *Am J Hum Genet* 27:509–520, 1975
20. DUSTAN HP: Obesity and hypertension, in *Childhood Prevention of Atherosclerosis and Hypertension*, edited by LAUER RM, SHEKELLE RB, New York, Raven Press, 1980, pp 305–312
21. DEPT. HEALTH, EDUCATION AND WELFARE: *Cardiovascular Profile of 15,000 Children of School Age in Three Communities 1971–1976*, DHEW publication no. 78-1472, Bethesda, Md., 1978
22. LAUER RM, CONNOR WE, LEAVERTON PE, REITER MA, CLARKE WR: Coronary heart disease risk factors in school children: The Muscatine Study. *J Pediatr* 86:697–706, 1975
23. SCHROTT HG, BUCHER KA, CLARKE WR, LAUER RM: The Muscatine Hyperlipidemia Family Study Program, in *Genetic Analysis of Common Diseases: Application to Predictive Factors in Coronary Heart Disease*, edited by SING CF, SKOLNICK M, New York, Alan R. Liss, 1979, pp 619–646
24. MOLL PP, SING CF, BREWER GJ, GILROY TE: Multivariate analysis of the genetic effects on red blood cell glycolysis, in *The Red Cell*, edited by BREWER GJ, New York, Alan R. Liss, 1978, pp 385–495
25. PLOMIN R, DEFRIES JC: Multivariate behavioral genetic analysis of twin data on scholastic abilities. *Behav Genet* 9:505–517, 1979
26. MCGUE M, RAO DC, REICH T, LASKARZEWSKI P, GLUECK CJ, RUSSELL JM: The Cincinnati Lipid Research Clinic Family Study: bivariate path analysis of lipoprotein concentrations. Unpublished
27. NETER J, WASSERMAN W: *Applied Linear Statistical Models*. Homewood, Ill., Richard D. Irwin, 1974
28. HANIS CL: Multivariate models for human genetic analysis: development and application to systolic blood pressure and weight. Ph.D. dissertation, Ann Arbor, University of Michigan, 1981
29. DEPT. HEALTH, EDUCATION AND WELFARE: *Growth Curves for Children Birth–18 Years*, DHEW publication no. 78-1650, Hyattsville, Md., 1977
30. ELSTON RC, NAMBOODIRI KK, KAPLAN EB: Resolution of major loci for quantitative traits, in *Genetic Epidemiology*, edited by MORTON NE, CHUNG CS, New York, Academic Press, 1978, pp 223–253
31. LANGE K, BOEHNKE M: Extensions to pedigree analysis. IV. Covariance component models for multivariate traits. *Am J Med Genet* 14:513–524, 1983
32. SING CF, HANIS CL, MOLL PP: Questions, measures and analytical strategies in human genetics, in *Genetic Basis of the Epilepsies*, edited by ANDERSON VE, HAUSER WA, PENRY JK, SING CF, New York, Raven Press, 1982, pp 239–247