Am J Hum Genet 30:28–37, 1978

Efficiency and Robustness of Pedigree Segregation Analysis

R. C. P. Go,[1] R. C. Elston,[1,2] and E. B. Kaplan[1]

The general method of pedigree analysis introduced by Elston and Stewart [1], which has recently been extended to cover a variety of situations [2–4], introduced two novel features. It allows the analysis of multigenerational pedigree data, and it allows quantitative data to be examined under simple Mendelian hypotheses without first being converted to a dichotomy. It is intuitively clear, and has been well confirmed [5], that there is an appreciable loss of information when quantitative data are dichotomized. It is also clear that the study of a single large pedigree will reduce the effects of genetic heterogeneity at the expense of loss of generality. Except in the study of linkage, however, it is perhaps not so clear whether, for a fixed number of persons studied, large pedigrees contain more information about the genetic parameters of interest than small nuclear families. Another difference between pedigree analysis [1] and traditional methods of segregation analysis arises specifically because more than 2 generations are considered. Parameter estimates and tests of hypotheses are based on the likelihood of all the data at hand, whereas the traditional methods of segregation analysis are based on the joint likelihood of all the data on the children *conditional on their parents' phenotypes*.

In the first part of this paper, we address the question of how the structure of the data and the likelihood used affect the information, that is, the efficiency of the estimates, under one-locus models. In the second part, we test the robustness of the method by examining particular situations not covered by the model to see if they are likely to suggest the presence of spurious major loci.

In order to differentiate between polygenic inheritance and segregation due to a single major locus, Elston and Stewart [1] suggested a model in which both effects are present. Morton and MacLean [6] developed this "mixed" model in detail for nuclear families, along with other refinements, and have prepared a computer program to perform the analyses. When more than 2 generations of data are present, however, the amount of computation under this model becomes enormous. Although this problem is being actively tackled by using approximative techniques [7], up to the present time,

Received April 19, 1977; revised July 11, 1977.

This investigation was supported by a U.S. Public Health Service Research Scientist Award (MH31732), and research grants from the National Institute of Mental Health (MH26721), from the National Cancer Institute (CA17107), and from the National Institute of General Medical Sciences (GM16697).

[1] Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina.

[2] Curriculum in Genetics, University of North Carolina at Chapel Hill.

analyses have been restricted to the simpler one-locus models [8–10]. In this paper, data are generated to simulate a wide variety of models and then analyzed under the simple one-locus model to determine its robustness.

THE ONE-LOCUS MODEL

The model used for analysis assumes for simplicity that after suitable sex and age adjustments there are just three types of individuals, *AA*, *Aa*, and *aa*. Data on these individuals come from three normal distributions, with means μ_{AA}, μ_{Aa}, and μ_{aa}, respectively, and common variance, σ_e^2, due to random environmental causes. The proportions of these three types of individuals in the population are taken to be ψ_{AA}, ψ_{Aa}, and ψ_{aa}, respectively, summing to unity; we assume throughout that $\psi_{Aa} = 2\sqrt{\psi_{AA}\psi_{aa}}$, corresponding to Hardy-Weinberg equilibrium under the simple Mendelian hypothesis. Finally the model includes three transmission probabilities, defined as follows: $\tau_{AA\,A}$ = probability that an *AA* individual transmits *A* to his offspring; $\tau_{Aa\,A}$ = probability that an *Aa* individual transmits *A* to his offspring; and $\tau_{aa\,A}$ = probability that an *aa* individual transmits *A* to his offspring. Under the Mendelian hypothesis, these three probabilities are 1, .5, and 0, respectively. If, on the other hand, there is no inheritance of any kind from one generation to the next, these probabilities are all equal.

Under the Mendelian hypothesis, we can define basic genetic parameters derived from the above as follows [6]: (1) (overall) mean: $\mu = \mu_{AA}\psi_{AA} + \mu_{Aa}\psi_{Aa} + \mu_{aa}\psi_{aa}$; (2) displacement: $t = \mu_{AA} - \mu_{aa}$; (3) (degree of) dominance: $d = (\mu_{Aa} - \mu_{aa})/(\mu_{AA} - \mu_{aa})$; (4) gene frequency: $q = \psi_{AA} + \psi_{Aa}/2$; (5) major locus variance: $\sigma_m^2 = \mu_{AA}^2\psi_{AA} + \mu_{Aa}^2\psi_{Aa} + \mu_{aa}^2\psi_{aa} - \mu^2$; (6) total variance: $\sigma_T^2 = \sigma_m^2 + \sigma_e^2$; and (7) (major locus) heritability $h^2 = \sigma_m^2 / \sigma_T^2$. In the next section, we shall investigate how well these seven parameters are estimated by five different strategies.

Efficiency of Parameter Estimation under a Major Gene Model

In this study, data were generated by Monte Carlo methods to simulate quantitative pedigree data that might be obtained under four major gene models. In two models, complete dominance is assumed ($d = 1$), and in two, additive gene action is assumed ($d = 0.5$); in each case, models with major locus heritabilities of 15% and 50% are examined. Table 1 gives details of the parameter values assumed in each of the four models.

TABLE 1

PARAMETERS USED IN THE FOUR MAJOR GENE MODELS SIMULATED

	COMPLETE DOMINANCE		ADDITIVE GENE ACTION	
PARAMETERS	$h^2 = .15$	$h^2 = .5$	$h^2 = .15$	$h^2 = .5$
Displacement, t	1.00	1.00	1.00	1.00
Gene frequency, q	.70	.65	.20	.33
Major locus variance, σ_m^2	.08	.11	.08	.11
Total variance, σ_T^2	.51	.22	.51	.22

Data were generated for 20 pedigrees each with the structure illustrated in figure 1. The parameters were then estimated by five maximum likelihood strategies, differing as to which individuals of the 16 member pedigrees were included in the likelihood calculations and which particular likelihood was maximized, as follows: (1) likelihood of the whole pedigree, 16 members (information on 320 individuals used) [L1]; (2) likelihood of the children, parents and grandparents, 10 members only per pedigree (information on 200 individuals used) [L2]; (3) likelihood of the nuclear families, parents and children, six members only per pedigree (information on 120 individuals used) [L3]; (4) likelihood of the four children conditional on the phenotypes of the rest of the pedigree members (information on 320 individuals used) [L4]; and (5) likelihood of the four children conditional on the parents' phenotypes (information on 120 individuals used) [L5]. Likelihood L3 is the simplest and was calculated for the one-locus model described above exactly as indicated by Elston and Stewart [1]; L1 and L2 are similar but used the extension to complex pedigrees given by Lange and Elston [4]. The conditional likelihoods L4 and L5 were calculated as the ratio of two likelihoods, the numerator in each case being L1, the joint likelihood of all 16 individuals. For L4, the denominator was the likelihood of the 12 individuals comprising the first 2 generations in figure 1—the product of the likelihoods of two nuclear families, since random mating was assumed. For L5 the denominator was the likelihood of the two parents—the product of the likelihoods for the separate parents.

For those data generated under the assumption of dominance, dominance (i. e., $\mu_{AA} = \mu_{Aa}$) was also assumed during the estimation; for those data generated under an additive model, three means and the degree of dominance were estimated.

Under each model, data on 30 replicate sets of the 20 pedigrees were generated, and 5×30 sets of maximum likelihood estimates thus obtained. Thus for each parameter, under each of the four models and for each different likelihood, 30 estimates were available from which to calculate estimates of the biases, standard errors, and root mean square errors (RMSEs). The biases were found to be very small, and hence the standard errors and RMSEs were nearly the same; therefore we need only report in detail here the results in terms of the RMSEs. Each RMSE was adjusted for sample size by dividing it by the square root of the number of individuals used in calculating the corresponding likelihood.

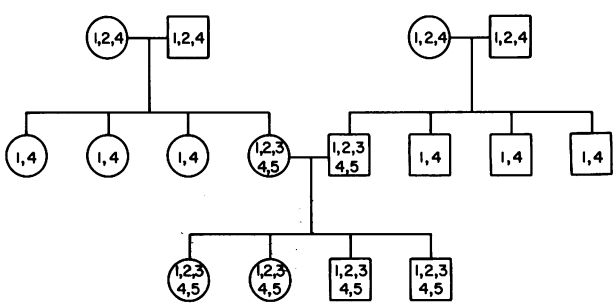

FIG. 1.—Pedigree structure used to study the efficiency of parameter estimation. The numbers 1, 2, 3, 4, and 5 indicate the individuals on whom information is used to calculate the likelihoods of L1, L2, L3, L4, and L5, respectively.

Although the RMSEs were somewhat erratic, as might well be expected because they were based on only 30 replicate sets, certain general trends were quite apparent. A comparison of the RMSEs for the additive and dominant models indicated that the dominant models usually resulted in better estimates; this was also expected, since complete dominance was assumed when analyzing the data from the dominant models, but the degree of dominance was estimated when analyzing the data from the additive models. For all models, with few exceptions, the RMSEs were greatest when the estimates were obtained by maximizing the likelihood of the children conditional on the phenotypes of the rest of the members of the pedigree. This method of utilizing information on the parental and grandparental generations is very inefficient. Similarly, the RMSEs were usually smallest when the first three likelihoods were used (i. e., the three unconditional likelihoods). There was little to choose among these three unconditional likelihoods, so it appears that the size of the family has little effect on the efficiency of parameter estimation. Overall there was a trend for the RMSEs to increase slightly with the size of the family structure, but this trend was reversed for the additive model with 15% heritability. Table 2 gives the detailed results for the complete dominance model with heritability 50%, which typify the overall trend in the four models studied.

ROBUSTNESS OF THE MODEL

Since the size of the family structure analyzed did not appear to be a critical factor in determining the efficiency of estimation, and to obtain results that would be comparable to those obtained by MacLean et al. [5], only nuclear families were simulated to study robustness. The unconditional likelihood (L3) was used throughout. Monte Carlo methods were used to simulate data on six-member nuclear families (two parents and four children). Initially data on 500 such families (i.e., 3,000 individuals) were used for each analysis; this sample size was reduced to one-tenth, however, in the later analyses. Three basic types of data were generated: polygenic data, data with environmentally caused skewness and platykurtosis, and data with both a polygenic component and skewness.

Throughout this study on robustness, it is assumed that there are no age or sex effects; in practice, either the data would be adjusted for such effects or the model would

TABLE 2

ROOT MEAN SQUARE ERRORS (RMSEs) OF THE PARAMETER ESTIMATES: MODEL WITH COMPLETE DOMINANCE, $h^2 = .5$

PARAMETER	LIKELIHOOD MAXIMIZED				
	L1	L2	L3	L4	L5
μ	0.64	0.65	0.62	3.70	1.0
t	1.06	0.96	0.97	1.77	1.0
q	0.76	0.70	0.63	3.05	1.0
σ_m^2	0.61	0.53	0.58	2.82	1.0
σ_T^2	0.61	0.53	0.60	2.78	1.0
h^2	0.63	0.61	0.56	2.37	1.0

NOTE. — RMSEs were adjusted for sample size and relative to the RMSEs for L5.

specifically allow for them. The following criteria developed by Elston et al. [9] to test for a major gene affecting a quantitative trait are used. (1) The data must show that a mixture of two normal distributions fit the data significantly better than one normal distribution. The test for two distributions fitting better than one is based on the likelihood ratio criterion, assuming a completely random sample; twice the negative of the $\log_e$ of the likelihood ratio follows asymptotically a chi-square distribution with 2 df. (2) The comparison of the maximum likelihoods obtained when the transmission probabilities are arbitrary but constrained between 0 and 1, and when they are fixed at their Mendelian values (single gene hypothesis), must be nonsignificant; this test is similarly based on a chi-square distribution. (3) The comparison of the maximum likelihoods obtained when the transmission probabilities are arbitrary but constrained to be between 0 and 1, and when they are constrained to be equal (environmental hypothesis), must be significant; this test is also based on a chi-square distribution. A 5% significance level was used for all tests of significance.

Polygenic Data

Initially 50 sets of polygenic data, 10 with each of the heritabilities .8, .5, .4, .2, and .1, were simulated onto 500 six-member nuclear families. Only additive polygenic variability and random environmental variability were included in this simulation. Let σ_g^2 and σ_e^2 be the additive genetic and environmental variances, respectively, so that the heritability is $\sigma_g^2/(\sigma_g^2 + \sigma_e^2)$. Each parent's genotypic value was chosen at random from the distribution $N(0, \sigma_g^2)$, and each member of the sibship's genotypic value was taken to be the midparent value plus a random number from $N(0, \sigma_g^2/2)$. Then a random number from $N(0, \sigma_e^2)$ was added to each genotypic value to obtain the individual's phenotypic value.

Of these 50 sets of polygenically simulated data, only three satisfied the first criterion; that is, a mixture of two distributions fit significantly better than one. In none of these three cases did a mixture of three distributions fit significantly better than a mixture of two distributions, so dominance was assumed in the subsequent major gene analyses of these three cases (whenever dominance was assumed, care was taken to look for two maxima on the likelihood surface, one corresponding to $\mu_{AA} = \mu_{Aa}$, and the other corresponding to $\mu_{aa} = \mu_{Aa}$).

When a dominant model was fitted to the data only one of these sets, at 40% heritability, fit the data well ($\chi^2 = 0.35$). This set also met the third criterion, namely that an environmental hypothesis resulted in a significantly bad fit. Therefore only one of the 50 original sets of data, based on the criteria developed for the Elston-Stewart model, led to the probable detection of a spurious major locus. It should also be noted that the test for a mixture of two distributions resulted in approximately the appropriate significance level, even though the phenotypic values were correlated, and not independent as assumed for the test. We conclude that this procedure is robust against detecting false major genes when the true mechanism is purely polygenic.

Environmentally Caused Skewness and Platykurtosis

Entirely environmental data were simulated onto sets of 50 nuclear, six-member families. Initially 10 such sets were generated by taking the phenotypes of a random

0.8 of the individuals from the distribution $N(0, 1)$ and those of the other 0.2 of the individuals from $N(3, 1)$, resulting in a skewed distribution overall. As expected, all 10 sets fit two distributions significantly better than one, but none satisfied the second criterion—the fit to a single gene hypothesis: the χ^2 values ranged from 13 to 44 with 3 df.

Ten very platykurtotic data sets were then simulated with 0.5 of the phenotypes coming from $N(0, 1)$ and 0.5 from $N(3, 1)$. Again all 10 sets fit two distributions better than one, but none fit a single gene hypothesis: χ^2 values ranged from 31 to 55. A striking similarity was that all these 20 sets of data fit the environmental hypothesis with equal transmission probabilities.

These two studies were repeated with an environmental correlation induced among siblings in the simulated data by generating, separately for each family, a random number from $N(0, 0.5)$ and then adding it to each of the children's phenotypes within that family. Addition of this sibling environmental correlation to the skewed data failed to change the previous results obtained without sibling correlation; the χ^2 values ranged from 18 to 36. In the case of the platykurtotic data sets, only three of the 10 sets now fit two distributions better than one, and none of these fit a single gene hypothesis; χ^2 values ranged from 11 to 26. These 20 simulated data sets also fit the environmental model in all instances.

Data with a Polygenic Component and Skewness

The final studies on simulated data included all the components discussed so far (i.e., polygenic additive genetic variance, sibling environmental correlation, and skewness). Homogeneous data of this form were simulated onto each of 10 sets of 50 six-member families with varying degrees of polygenic heritability and sibling environmental correlation. Skewness was then created by adding six to all individual values (to insure values greater than zero) and taking logarithms. These data were thus similar to those used by MacLean et al. [5], but the resulting skewness was -0.51 rather than -0.87. The variance components, heritabilities, and correlations quoted in the sequel are those before logs were taken.

The data were simulated as indicated above for the polygenic data, with the addition of a random number from $N(0, \sigma_b^2)$ to each of the siblings—the same number within each sibship, but a different number from sibship to sibship. In each case, values of σ_g^2, σ_e^2 and σ_b^2 were taken that sum to unity, so that heritability can be equated to σ_g^2, and sibling environmental correlation can be equated to σ_b^2.

Table 3 summarizes the results obtained for increasing sibling environmental correlations when heritability is held fixed at .2. When the sibling environmental correlation is .1, six out of 10 sets of data fit our first two criteria for the single gene hypothesis; of these six, two also fit the environmental hypothesis, leaving four remaining sets that satisfy all three criteria for a major gene. Similarly it can be seen in table 3A that when the sibling environmental correlation is .2, six of the data sets satisfy the three criteria for a major gene, and when it is .3, five data sets do so. When the sibling environmental correlation is .4 or greater (or equal to zero), the chance of detecting a spurious major gene using our criteria is no greater than the significance level used in the tests. Thus when both the heritability and the sibling environmental

TABLE 3

RESULTS OF ANALYZING DATA WITH VARYING AND FIXED POLYGENIC HERITABILITY (OF .2) AND VARYING DEGREES OF SIBLING ENVIRONMENTAL CORRELATION, SKEWNESS INDUCED BY TAKING LOGARITHMS

VALUES OF VARIANCE COMPONENTS USED TO SIMULATE DATA			NO. SETS FITTING TWO DISTRIBUTIONS BETTER THAN ONE	NO. SETS FITTING MENDELIAN TRANSMISSION	NO. SETS FITTING AN ENVIRONMENTAL HYPOTHESIS	NO. SETS MEETING THE THREE CRITERIA FOR A MAJOR GENE	MAXIMUM DIFFERENCE IN LIKELIHOODS BETWEEN DOMINANT AND RECESSIVE HYPOTHESES
σ_g^2	σ_e^2	σ_b^2					
				A. Fixed (Polygenic) Heritability			
.2	.7	.1	9	6	2	4	17-fold
.2	.6	.2	9	7	1	6	28-fold
.2	.5	.3	9	5	0	5	10-fold
.2	.4	.4	9	1	1	0	. . .
.2	.3	.5	8	2	0	2	7-fold
.2	.2	.6	8	0	0	0	. . .
.2	.0	.8	9	0	0	0	. . .
				B. Varying Values of (Polygenic) Heritability			
.4	.4	.2	10	5	0	5	132-fold
.4	.2	.4	9	1	0	1	2-fold
.4	.0	.6	9	1	0	1	96-fold
.6	.2	.2	9	3	0	3	4,169-fold
.6	.0	.4	9	1	0	1	4-fold
.8	.0	.2	9	2	0	2	79,682-fold

NOTE. — Each line gives results of analysis of 10 sets of 50 six-member (nuclear) families.

correlation are about .2, and the data are negatively skewed, there is a serious possibility of falsely "detecting" a major locus.

In every data set where a mixture of two distributions was found to fit significantly better than one distribution, it was possible to find two local maxima on the likelihood surface for the case of Mendelian transmission: one of these maxima corresponded to the distribution with the lower mean being due to a recessive, the other to it being due to a dominant gene. The larger likelihood was always taken to test the fit of the model, but these two maxima rarely differed very much; the last column of table 3A gives the maximum difference found, as a multiplicative factor, in each case. This finding is similar to one found by MacLean et al. [5], who note that "The most common danger signal is a U-shaped [likelihood] surface with respect to dominance."

The next procedure was to increase the polygenic heritability of the trait and examine the robustness of the model over a range of heritabilities; again the model contained varying degrees of sibling environmental correlation and random environmental effects. Table 3B summarizes the results of these simulations. The same trend is again seen: for low nonzero sibling environmental correlation, there is a serious danger of detecting spurious major loci, but this danger decreases rapidly as the sibling environmental correlation increases. Furthermore, in none of these data sets where the heritability was .4 or greater, and which met the first two criteria for a major gene, did the data fit the environmental model with equal transmission probabilities. It is also seen from the last column in table 3B that for these simulations the maximum differences in likelihood found between dominant and recessive hypotheses can be quite large.

DISCUSSION

In the first part of this paper we determined that to estimate the parameters of a one-locus model for quantitative data, it is better to use the complete unconditional likelihood rather than the likelihood of sibships conditional on the parental phenotypes. It must be stressed that this finding is relevant when the one-locus model is known (or assumed) to be correct, and we are not concerned with testing that mechanism as a genetic hypothesis. If we had been interested in estimating the transmission probabilities or in testing whether there is indeed Mendelian transmission, maximizing a conditional likelihood would have perhaps been preferable. It is, however, intuitively probable that this would only be the case for data on nuclear families alone; it is difficult to believe that, when 3 or more generations of data are available, conditional likelihoods such as L4 would be more informative about the transmission probabilities.

The studies on robustness were restricted to nuclear families, and we cannot be certain that the results would be the same for extended families. The application of pedigree analysis to single large families has the advantage of minimizing the genetic heterogeneity present in any one study; for this reason, we have not investigated robustness of the methods in the face of such genetic heterogeneity. Even though the robustness and power of the proposed method of analysis has not been systematically investigated for extended pedigrees, results of the method in which the presence of a major gene for hypercholesterolemia has been confirmed by biochemical and linkage analysis are encouraging [9, 11–13]. On the basis of our empirical studies and the

present simulations, we believe the model to be quite robust if the following two criteria are met (in addition to those explicitly stated above) before concluding that a major gene is present: (1) The multiplicative difference between the two maximum likelihoods obtained when dominant and recessive hypotheses are fitted should be large; and (2) the maximum likelihood estimates of the transmission probabilities should be examined. If they are inexplicably very different from the Mendelian values, and yet the single gene hypothesis is not rejected (i. e., the second of the three primary criteria is satisfied), a flat likelihood is indicated. This suggests that a large variety of hypotheses will fit the data, so one should be cautious in concluding the presence of a major gene.

It is difficult to specify the magnitude of the difference that should be required for criterion (1). As a guide, we may note that of all the analyses performed in this study of robustness in which both dominant and recessive Mendelian hypotheses were fitted, 95% of the likelihood ratios (always taking the ratio greater than unity) were less than 28. In fact the ratio was only found to be greater than 28 for the simulated situations in which, simultaneously, the polygenic heritability was .4 or greater, there was a moderate sibling environmental correlation, and the data showed significant skewness; furthermore, such a large ratio virtually only occurred when the sibling environmental correlation was greater than zero and less than .4.

Our studies on robustness have sampled only some of the many situations that may actually occur. We have not considered the effects of assortative mating (which we know are negligible under other methods of analysis [5]), nor have we considered the possibility of genotypic and environmental effects being nonadditive. Nevertheless, the simulation results presented here do suggest that the single gene model that includes as parameters the transmission probabilities defined by Elston and Stewart [1], provided it is used carefully, will often be a useful and robust tool for the detection of major loci. Skewness per se will not lead to the detection of a spurious locus; inclusion of the transmission probabilities in the model allows one to distinguish between the Mendelian hypothesis and environmentally caused skewness. Nor will polygenic inheritance per se lead to the detection of a spurious locus, if we insist that the data fit a mixture of normal distributions significantly better than a single normal distribution prior to the segregation analysis. It is only when these two factors occur together, along with a moderate environmental correlation among sibs, that we have found a serious possibility of falsely detecting the presence of a major gene effect. In any case, it is desirable to confirm the presence of such genes by an elucidation of the basic biochemical defect or by showing a linkage relationship between the hypothesized gene and an established genetic marker.

SUMMARY

Different pedigree structures and likelihoods are examined to determine their efficiency for parameter estimation under one-locus models. For the cases simulated, family size has little effect; estimates based on unconditional likelihoods are generally more efficient than those based on conditional likelihoods. The proposed method of pedigree analysis under a one-locus model is found to be robust in the analysis of nuclear families: skewness of the data and polygenic inheritance will not lead to the

spurious detection of major loci unless they occur simultaneously, and together with a moderate amount of environmental correlation among sibs.

REFERENCES

1. ELSTON RC, STEWART J: A general model for the genetic analysis of pedigree data. *Hum Hered* 21:523–542, 1971
2. ELSTON RC: Ascertainment and age of onset in pedigree analysis. *Hum Hered* 23:105–112, 1973
3. ELSTON RC, YELVERTON KC: General models for segregation analysis. *Am J Hum Genet* 27:31–45, 1975
4. LANGE K, ELSTON RC: Extensions to pedigree analysis. I. Likelihood calculations for simple and complex pedigrees. *Hum Hered* 25:95–105, 1975
5. MACLEAN CL, MORTON NE, LEW R: Analysis of family resemblance. IV. Operational characteristics of segregation analysis. *Am J Hum Genet* 27:365–384, 1975
6. MORTON NE, MACLEAN CJ: Analysis of family resemblance. III. Complex segregation of quantitative traits. *Am J Hum Genet* 26:489–503, 1974
7. ELSTON RC, KAPLAN EB: Polynomial approximations for pedigree analysis. In preparation
8. ELSTON RC, NAMBOODIRI KK, NINO HV, POLLITZER WS: Studies on blood and urine glucose in Seminole Indians: indications for segregation of a major gene. *Am J Hum Genet* 26:13–34, 1974
9. ELSTON RC, NAMBOODIRI KK, GLUECK CJ, FALLAT R, TSANG R, LEUBA V: Study of the genetic transmission of hypercholesterolemia and hypertriglyceridemia in a 195 member kindred. *Ann Hum Genet* 39:67–87, 1975
10. ELSTON RC, NAMBOODIRI KK, SPENCE MA, RAINER JD: A genetic study of schizophrenia pedigrees. II. One-locus hypotheses. *Neuropsychobiology*. In press, 1977
11. BROWN MS, GOLDSTEIN JL: Expression of the familial hypercholesterolemia gene in heterozygotes: mechanism for a dominant disorder in man. *Science* 185:61–63, 1974
12. ELSTON RC, NAMBOODIRI KK, GO RCP, SIERVOGEL RM, GLUECK CJ: Probable linkage between essential familial hypercholesterolemia and C3. *Baltimore Conference (1975): Third International Workshop on Human Gene Mapping, Birth Defects: Orig Art Ser* 12(7), New York, The National Foundation, 1976, pp 294–297
13. BERG K, HEIBERG A: Linkage studies on familial hyperlipoproteinemia with xanthomatosis: normal lipoprotein markers and the C3 polymorphism. *Baltimore Conference (1975): Third International Workshop on Human Gene Mapping, Birth Defects: Orig Art Ser* 12(7), New York, The National Foundation, 1976, pp 266–270