Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 9177–9181, August 1996
Medical Sciences

Trials of the β model for complex inheritance

A. COLLINS†, C. J. MACLEAN‡, AND N. E. MORTON†

†Human Genetics, University of Southampton, Level G, Princess Anne Hospital, Coxford Road, Southampton, Hants, SO16 5YA United Kingdom; and
‡Psychiatric Genetics Research Program, Department of Psychiatry, Medical College of Virginia, P.O. Box 980710, Richmond, VA 23298-0710

Contributed by N. E. Morton, March 25, 1996

ABSTRACT **Theoretical advantages of nonparametric logarithm of odds to map polygenic diseases are supported by tests of the β model that depends on a single logistic parameter and is the only model under which paternal and maternal transmissions to sibs of specified phenotypes are independent. Although it does not precisely describe recurrence risks in monozygous twins, the β model has greater power to detect family resemblance or linkage than the more general Δ model which describes the probability of 0, 1, or 2 alleles identical by descent (ibd) with two parameters. Available data on ibd in sibs are consistent with the β model, but not with the equally parsimonious but less powerful γ model that assumes a fixed probability of ½ for 1 allele ibd. Additivity of loci on the liability scale is not disproven. A simple equivalence extends the β model to multipoint analysis.**

Genes of unknown structure and function can be identified by linkage mapping to a small region that may be refined and ultimately cloned preparatory to sequencing. This process has been called *positional cloning*. A central problem in genetic epidemiology is to extend positional cloning from major loci to polygenes. Two directions have been pursued: parametric models of two or more disease loci and nonparametric tests that depend only on marker loci that may be linked to disease loci. The latest development in nonparametric tests is the β model that specifies identity by descent at a linked marker locus in terms of a single parameter (1). It predicts recurrence risks, allelic interaction, locus additivity, and favorable comparison with alternative nonparametric models (Table 1 and Fig. 1). We shall now test these predictions, leaving comparison with parametric models to work in progress.

Recurrence Risks

Under the β model recurrence risks are functions of two parameters a,$\beta > 0$ and the following relations hold exactly.

Population risk: $R_0 = e^{-a}$

Recurrence risk in parents or children: $R_c = e^{-a+\beta}$

Recurrence risk in MZ twins: $R_m = e^{-a+2\beta}$

To a good approximation the recurrence risk in sibs and other relatives with kinship φ is $R_\varphi = e^{-a+4\varphi\beta}$, where φ is ¼ for sibs, ⅛ for second degree relatives (uncle–niece), and 1/16 for the third degree (first cousins). The assumptions are that alleles and loci act independently (i.e., multiplicatively on risk) and that family environment has a negligible effect on recurrence or is confounded with genetic relationship. Not only are these assumptions difficult to test directly, but estimates of recurrence risk vary among populations and are sensitive to definition of affection, age distribution, ascertainment, and reproduction rate. Therefore statistical significance must be examined cautiously. To do this let $R = A/N$, where A is the number of affected and N is sample size. Maximum likelihood scores are

$$U_\theta = \sum_r \left[\frac{A_r}{R_r} - \frac{N_r - A_r}{1 - R_r}\right]\left(\frac{\partial R_r}{\partial \theta}\right),$$

with information matrix

$$K_{\theta\theta'} = \sum_r \left[\frac{N_r}{R_r(1 - R_r)}\right]\left(\frac{\partial R_r}{\partial \theta}\right)\left(\frac{\partial R_r}{\partial \theta'}\right),$$

where θ, θ' denote a,β. Newton–Raphson iteration on $\theta + UK^{-1}$ converges to maximum likelihood (ML) estimates with standard error $\sqrt{K^{-1}_{\theta\theta}}$ if $\chi^2/df < 1$ or $df = 0$. Otherwise we take the empirical standard error $\sqrt{\chi^2 K^{-1}_{\theta\theta}/df}$ based on residual χ^2 and its degrees of freedom (df = number of values of $r - 2$).

We applied this method to published data. When the population risk was given as an interval, the mean was used; when the size of the random sample was omitted, it was arbitrarily taken as 10,000. When N for a sample of relatives was not given it was arbitrarily taken as 100 for MZ twins and 1000 for other relatives. The data were not collected with this analysis in view and do not yield precise estimates (2–10). Notwithstanding this caveat, they are of some interest (Table 2).

The smallest β is for atopy as hay fever, which is less than half as much as for asthma. The largest estimate is for hereditary genius based on a study from the last century that does not discriminate between genes and family environment (10). Even if the latter were negligible, the value of β and therefore the number of contributing genes does not seem to be much greater than for the other traits. A substantial study is likely to map a susceptibility locus if the closest marker has a value of β as great as 0.25 (1). If the model is true, we might hope to detect several leading factors for most of these traits (11), but perhaps only two for atopy.

When residual χ^2 is significant, possible explanations are random errors in the data and systematic departure from the β model. This can be tested by fitting the more general Δ model. Ignoring differences between recurrence risks in sibs and parent–child pairs, which may have several causes, the effect of the Δ model is to change recurrence risks in MZ twins to $e^{-a+\beta+\Delta}$ so that the β model corresponds to the subhypothesis $\Delta = \beta$ (1). We fitted the Δ model under the constraint $\beta > 0$, $\Delta \geq \ln(2 - e^{-\beta})$, which confines identity by descent estimates to the "possible triangle." The Δ model fits all five traits with residual df, the total χ^2 for goodness of fit to disease-specific parameters being 5.96 with 10 df (Table 3). Three of these traits gives significant deviation from the β model, as $\chi^2_1 = \chi^2(\beta) - \chi^2(\Delta)$. The most extreme deviation is for Alzheimer disease, reflecting rare dominant genes in a considerable proportion of early onset cases. Schizophrenia also gives a much better fit to the Δ model, indicating non-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: ML, maximum likelihood; ibd, identical by descent; lod, logarithm of odds; df, degrees of freedom.

Table 1. Alternative models of allelic interaction

Model	Description	Probability of 0, 1, 2 alleles ibd in affected sib pair		
		ζ_0	ζ_1	ζ_2
β	No allelic interaction on logistic scale	$1/(1 + e^{\beta})^2$	$2e^{\beta}/(1 + e^{\beta})^2$	$e^{2\beta}/(1 + e^{\beta})^2$
γ	No allelic interaction on penetrance scale	$1/2(1 + e^{\gamma})$	1/2	$e^{\gamma}/2(1 + e^{\gamma})$
Δ	Allelic interaction on logistic scale ($\Delta \neq \beta$)	$1/1 + 2e^{\beta} + e^{\beta+\Delta}$	$2e^{\beta}/1 + 2e^{\beta} + e^{\beta+\Delta}$	$e^{\beta+\Delta}/1 + 2e^{\beta} + e^{\beta+\Delta}$

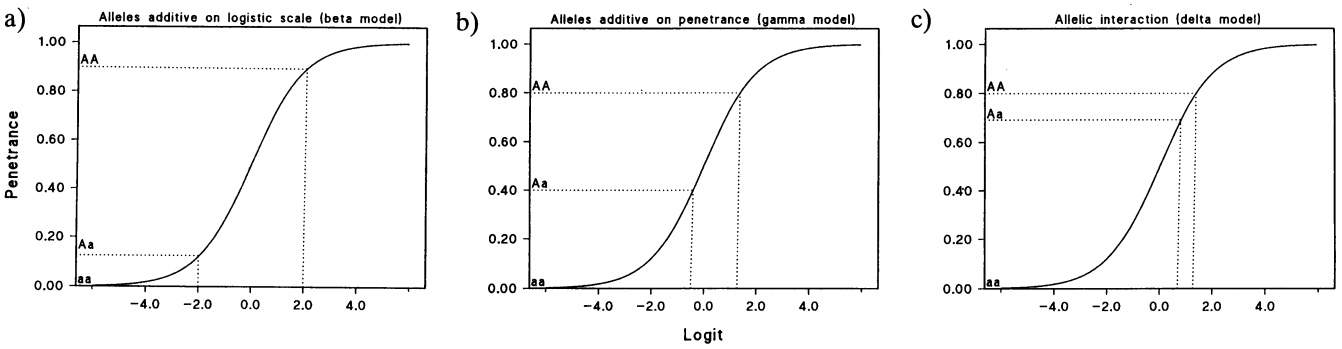

FIG. 1. Alternative models of allelic interaction. (*a*) Alleles additive on logistic scale (β model). (*b*) Alleles additive on penetrance (γ model). (*c*) Allelic interaction (Δ model).

Table 2. Recurrence risks per 10,000 and β estimate

Trait	Kinship φ					β	SE(β)	Ref.
	0	0.0625	0.125	0.25	0.5			
Cleft lip ± cleft palate	10	30	70	400	4000	2.79	0.20	2
Talipes equinovarus	10	20	50	250	3000	2.78	0.14	2
Congenital dislocation of hip	20	40	60	500	4000	2.58	0.19	2
Pyloric stenosis	50	80	250	500	4000	2.16	0.11	2
Schizophrenia	85	156	275	879	4559	1.91	0.14	3
Asthma	380	—	—	930	3700	1.06	0.14	4
Hay fever	1480	—	—	2417	4000	0.49	0.04	4
Multiple sclerosis	10	—	—	147	2714	2.85	0.15	5
Type 1 diabetes	40	—	—	600	3600	2.22	0.26	6
Type 2 diabetes	500	—	—	2600	9000	1.43	0.08	7
Coeliac disease	5	—	—	300	6800	3.39	0.27	8
Alzheimer disease	5	—	—	707	4118	2.78	0.88	9
Hereditary genius	3	74	279	1269	—	3.81	0.64	10

Table 3. Tests of hypotheses on recurrence risks

Trait	β model		Δ model					μ_β/μ_Δ
	χ^2 (β)	df	χ^2 (Δ)	Δ	SE(Δ)	β	SE(β)	
Cleft lip ± cleft palate	9.36	3	0.12	2.29	0.19	3.65	0.32	0.978
Talipes equinovarus	2.00	3	0.02	2.49	0.25	3.23	0.36	0.993
Congenital diclocation of hip	10.74	3	1.65	2.11	0.18	3.28	0.26	0.978
Pyloric stenosis	5.04	3	4.16	2.01	0.26	2.31	0.27	0.997
Schizophrenia	29.58	3	0.01	1.65	0.07	2.33	0.09	0.977
Asthma	4.05	1	0	1.38	0.16	0.90	0.11	0.974
Hay fever	0.01	1	0	0.50	0.13	0.49	0.05	1.000
Multiple sclerosis	0.27	1	0	2.91	0.19	2.69	0.34	0.999
Type 1 diabetes	7.68	1	0	1.79	0.18	2.71	0.20	0.979
Type 2 diabetes	10.84	1	0	1.24	0.06	1.65	0.07	0.987
Coeliac disease	2.92	1	0	3.12	0.23	4.09	0.48	0.989
Alzheimer disease	61.63	1	0	1.76	0.30	4.95	0.45	0.879
Hereditary genius	62.71	2	—	—	—	3.81	0.64	—

Table 4. Analysis of identity by descent in affected pairs of sibs by the β model

N_k pairs with k alleles ibd (k = 0, 1, 2)
$T = \Sigma N_k$
ζ_k = probability of k marker alleles identical by descent
Under H_0, $\zeta_k = c_k$ where $c_0 = ¼$, $c_1 = ½$, $c_2 = ¼$
Likelihood $L = \exp(\beta N_1 + 2\beta N_2)/(1 + e^{\beta})^{2T}$
ML score $U = \partial \ln L/\partial\beta = N_1 + 2N_2 - 2Te^{\beta}/(1 + e^{\beta})$
Information $K = -E(\partial U/\partial\beta) = 2Te^{\beta}/(1 + e^{\beta})^2$
Under H_0 ($\beta = 0$): $\chi^2_1 = 2 \Sigma N_k \ln (N_k/c_k T)$
Under H_1 ($\beta > 0$): Newton–Raphson iteration of $\beta \rightarrow \beta + U/K$; at convergence to the ML
Estimate $\hat{\beta}$, the standard error is $\sqrt{1/K}$
Residual $\chi^2_1 = 2 \Sigma N_k \ln (N_k/\zeta_k T)$

multiplicative risks and perhaps a proportion with a dominant leading factor. Liability models (logit or probit) are not confined to additivity of alleles and loci and therefore approximate but do not correspond to multiplicative risks, which are also violated by etiological heterogeneity. The estimate of Δ is less than β for six of the seven samples in which the Δ model is significantly better, giving no indication of bias toward ascertainment of concordant twins. The opposite bias is possible by misdiagnosis of zygosity, preferential publication of discordant pairs, or confusion between concordance and recurrence risk. Concordance falls progressively below recurrence risk as the ascertainment probability increases. For recurrence risk each co-twin should be counted once for each proband, and so concordant pairs with two probands should be counted twice. Although probands are rarely indicated, most work with twin concordance approaches single selection for which concordance is equivalent to recurrence risk.

Because of the complexity of possible deviations from the β model, we cannot be certain of the cause for any particular trait. However, there is no doubt that fit of the Δ model is impressive. Assuming that it will continue to be superior to the β model for other traits and in better data, it does not follow that it should be used to detect deviations from the null hypothesis that $\beta = \Delta = 0$. A basic principle of statistical analysis is that power to reject a null hypothesis may be increased by use of a parsimonious alternative, since power depends on both df and noncentrality parameter μ (12). The expected value of noncentral χ^2 with 1 df is $\nu + \mu$, where μ is the noncentrality parameter for the observed sample size. Assuming the sample frequencies as parameters, the noncentrality parameter for the β model with 1 df is $\mu_\beta = \chi^2(\alpha) - \chi^2(\beta)$ where $\chi^2(\alpha)$ is the value of χ^2 when $\beta = \Delta = 0$ (no family resemblance) and $\chi^2(\beta)$ is the value when $\Delta = \beta$. The noncentrality parameter for the Δ model with 2 df is $\mu_\Delta = \chi^2(\alpha) - \chi^2(\Delta)$ where $\chi^2(\Delta)$ is the value for ML estimates of Δ and β in the possible triangle. The ratio μ_β/μ_Δ in these samples approaches 1, showing that the parsimony of the β model increases power (Fig. 1). Even for the smallest value of μ_β/μ_Δ (Alzheimer disease) the β test is more powerful than the Δ test (13).

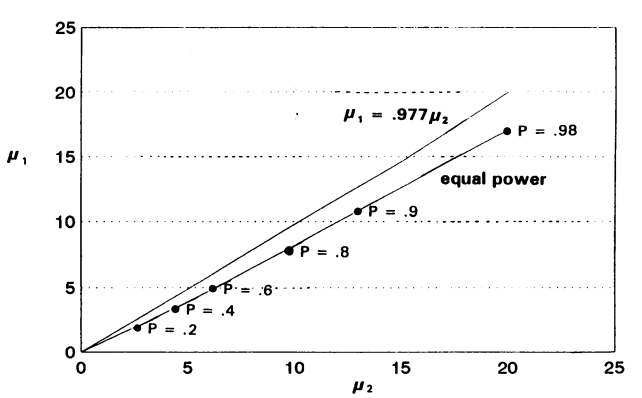

FIG. 2. The power of β and Δ models.

Allelic Interaction

Kruglyak and Lander (14) asserted that the probability ζ_1 of 1 allele identical by decent (ibd) in two sibs is ½ on the assumption of no dominance variance. Actually, the assumptions are more restrictive: $\zeta_1 = ½$ if there is no dominance on the penetrance scale or if there is free recombination between marker and closest disease locus. On the contrary, the assumption of no dominance on an underlying liability or logistic scale leads to $\zeta_1 < ½$ for a linked marker (Fig. 1). This is reflected in the β model as $\zeta_1 = 2e^{\beta}/(1 + e^{\beta})^2$, which is less than ½ unless $\beta = 0$ (Table 4). To test the β model we searched the literature for data on affected sib pairs classified by 0, 1, 2 marker alleles identical by descent. Unfortunately most recent studies give only the 0, 1 distribution for maternal and paternal alleles separately (15–18). All of the six distributions we found have $\zeta_1 < ½$, as predicted (Table 5). Moreover they agree quantitatively with ML estimates of β.

On the contrary, the γ model of Kruglyak and Lander (14) fits poorly (Table 6). Summing χ^2 values for goodness of fit to locus-specific parameters, the ratio of noncentrality parameters for the γ and β tests is 0.9 (Table 7). If these data are representative, the β test is more powerful than the γ and Δ tests (Fig. 2).

Locus Additivity

Under the β model the probability that a heterozygous parent transmits the same allele to a pair of affected children (1) is $\rho = e^{\beta}/(1 + e^{\beta})$, and $\beta = \ln[\rho/(1 - \rho)]$. If the observed numbers of concordant and discordant transmissions are N_1, N_0, respectively, the ML estimate of ρ is $N_1/(N_1 + N_0)$ and the corresponding lod $z = N_1 \log [2N_1/(N_1 + N_0)] + N_0 \log [2N_0/(N_1 + N_0)]$. In large-sample theory (2 ln 10) z has a χ^2 distribution with 1 df on the null hypothesis of no linkage between the marker and a disease locus. This test would be optimal if all parents were typed and fully informative, whereas the β model is general and includes this as a special case.

Table 5. Tests of allelic interaction (0, 1, 2 ibd)

					β model		γ model		α model	
Disease	Marker	0 ibd	1 ibd	2 ibd	β	$\chi^2_1(\beta)$	γ	$\chi^2_1(\gamma)$	$\chi^2_2(\alpha)$	Ref.
Type 1 diabetes	HLA	55	283	373	0.963	0.02	1.914	29.78	294.81	15
Type 1 diabetes	FGF3	20	59	40	0.339	0.05	0.693	0.01	6.80	16
Coeliac disease	HLA	0	16	20	1.253	4.64	13.077	0.45	28.17	17
Haemochromatosis	HLA	0	3	15	2.398	0.27	14.509	8.73	29.53	17
Psoriasis	HLA	3	24	37	1.184	0.13	2.512	4.04	38.18	17
Multiple sclerosis	HLA	33	110	110	0.629	0.44	1.204	4.32	48.06	18

Table 6. Analysis of identity by descent in affected pairs of sibs by the γ model

Likelihood $L = e^{\gamma N_2}/(1 + e^{\gamma})N_0 + N_2$
$U = \partial \ln L/\partial\gamma = (N_2 - N_0 e^{\gamma})/(1 + e^{\gamma})$
$K = -E(\partial U/\partial\gamma) = T(e^{2\gamma} + e^{\gamma})/2(1 + e^{\gamma})^3$
Under H_0 ($\gamma = 0$), $\chi^2_1 = 2 \Sigma N_k \ln (N_k/c_k T)$
Under H_1 ($\gamma > 0$), Newton–Raphson iteration of $\gamma \rightarrow \gamma + U/K$; at convergence to the ML
Estimate γ the standard error is $\sqrt{1/K}$ and residual $\chi^2_1 = 2 \Sigma N_k \ln (N_k/\zeta_k T)$

If disease loci act additively on the liability or logistic scale their effects on risk are nearly multiplicative and therefore the value of β estimated from recurrence risks should be nearly the sum of the values for contributory loci, or $\beta = \Sigma\beta_i$. On the contrary, β tends to be greater than $\Sigma\beta_i$ if there is recombination or a significant effect of family environment or if some loci have not been detected, while β tends to be less than $\Sigma\beta_i$ if type 1 errors are misinterpreted as contributory loci or if gene effect is less than additive on the logistic scale. To test locus additivity we may estimate β either by Table 4 or from transmission frequencies summed over informative parents (Table 8). The β model is unique in making this information additive if correctly tabulated, and so the γ and Δ models cannot be applied. The χ^2 (α) test for $\beta = \Delta = 0$ should agree closely with the χ^2 (MLS) test (20) on the same data, and so the larger differences presumably reflect nonidentical samples.

Insulin-dependent diabetes (IDDM) currently provides the best test of additivity, although many of the loci tentatively identified are not supported by the conventional level of $z > 3$, corresponding to $P < 0.0001$ in large samples and to $P < 0.001$ in general (23, 24). In samples large enough to detect leading factors with high power, the probability that a significant test is a type 1 error is small (25), in contrast to the extremely conservative Ornstein–Uhlenbeck model under which *all* significant tests are type 1 errors (26). However, both approaches justify distrust of a weaker significance level. It should be noted that several tentative IDDM loci give evidence of allelic association ("transmission disequilibrium"), which by an efficient test would increase significance. On these data the hypothesis of locus additivity is not rejected, since estimates of $\Sigma\beta$ for significant loci are in reasonable agreement with the estimate from recurrence risks (Table 9). If locus additivity is true, nearly all the leading factors for IDDM have been detected. If it is false, heterogeneous etiology and synergism on the logistic scale are two possible mechanisms. Estimates of relative effects of contributory loci are model-dependent but in rough agreement (Table 10).

Table 9. Summary of locus additivity for IDDM

Ref. or source	No. of loci	$\Sigma\beta$
20	10	2.52
19, 21, 22	10	3.24
19, 21, 22, $\chi^2 > 4$	6	2.28
Recurrence risk	∞	2.22
λ_S (20)	∞	2.71

Table 10. Relative effect of HLA on IDDM

Ref. or source	β_{HLA}	$\Sigma\beta$	$\beta_{HLA}/\Sigma\beta$
19, 21, 22	1.01	2.22	0.45
20	0.80	2.71	0.30
ln λ_S (20)	0.96	2.71	0.35

Discussion

Cotterman (27) enumerated all possible identity coefficients between two pairs of alleles, showed that they reduced to a trinomial when each pair was independent (corresponding to noninbred or "regular" relatives), introduced the symbol c_k for the probability of k alleles identical by descent, and derived conditional phenotype probabilities for single loci. Even for a diallelic locus this requires four parameters (one gene frequency and three penetrances), and so extension to complex inheritance must be simplified. James (28) expressed recurrence risks in terms of variance components for a binary trait. Risch (29) adopted this convention, dividing risk by population prevalence K to give the risk ratio λ_R for a relative of type R. For a particular locus the risk ratios for parent–offspring pairs (o) and monozygous twins (m) are

$$\lambda_o = 1 + A/2K^2$$

$$\lambda_m = 1 + (A + D)/K^2$$

Table 7. Noncentrality parameters

	α test			β test				γ test	
Source	χ^2	df	μ_Δ	χ^2	df	$\mu_{\Delta/\beta}$	μ_β/μ_Δ	χ^2	μ_γ/μ_β
Recurrence risks*	5316.08	34	5282.08	144.12	22	122.12	0.977	—	—
ibd†	445.55	12	433.55	5.55	6	0	1.000	47.33	.905

*Table 3, 12 diseases with MZ recurrence.
†Table 5, 6 diseases with ibd in sibs.

Table 8. Tests of locus additivity for type 1 diabetes

Locus	0 ibd	1 ibd	β	χ^2 (α)	Chromosome	MLS	χ^2 (MLS)	Todd β	Ref.
IDDM1	98	268	1.01	82.08	6p21	19.3	88.9	0.80	19, 20
IDDM2	227	287	0.23	7.02	11p15	1.6	7.4	0.24	19, 20
IDDM3	271	329	0.19	5.62	15q	—	—	—	21
IDDM4	175	208	0.17	2.85	11q13	1.3	6.0	0.07	19, 20
IDDM5	199	249	0.22	5.59	6q25	1.5	6.9	0.14	19, 20
IDDM6	55	72	0.27	2.28	18q	1.1	5.1	0.09	19, 20
IDDM7	177	226	0.24	5.97	2q31	1.3	6.0	0.12	20, 22
IDDM8	—	—	—	—	6q27	1.2	5.5	0.32	20
IDDM9	—	—	—	—	3q21-q25	—	—	0.22	20
DXS1068	32	24	0.29	1.15	Xq	—	—	0.18	20, 22
IDDM10	48	71	0.39	4.47	10p11.2-q11.2	1.3	6.0	0.34	19, 20
GCK	120	151	0.23	3.55	7p	—	—	—	19

where A, D are the additive and dominance genetic variances for the binary trait. An established principle of biometrical genetics is that gene effects should be studied on a scale that minimizes interaction (30). Statistical experience has shown that interaction is usually less on a liability (probit) or logistic scale than for the corresponding binary trait, which is more distribution-sensitive (12). This motivated our Δ model in which $\lambda_o = e^{\beta}$, $\lambda_m = e^{\beta+\Delta}$, where $\beta = CA^*/2$, $\Delta = C(A^*/2 + D^*)$, C is a constant, and asterisks signify that variance components are defined on the logistic scale. The hypothesis of no dominance ($D^* = 0$) corresponds to the β model ($\Delta = \beta$).

Risch and coworkers (31, 32) have preferred the γ model in which $\zeta_1 = c_1 = ½$. This is incompatible with the β model except for the trivial case $\beta = \Delta = 0$. More seriously, we have shown that ζ_1 is usually less than ½ for loci with a real effect and would equal ½ only for a rare dominant with no sporadic cases, which would have expected frequencies 0, ½, ½ for 0, 1, 2 genes ibd, and so would fit the γ model better. However, this is an unrealistic model for complex inheritance and would readily be detected by any method. The γ model cannot be extended to unilineal relatives, for whom $c_1 = 1 - c_o \neq \zeta_1$. Like the Δ model it makes parental transmission frequencies dependent, whereas the β model can be factored into independent maternal and paternal contributions.

We have found the β model to be more powerful than γ and Δ alternatives. Discrepancies revealed by the Δ model in MZ twins, whatever their cause, are not supported by tests of allelic interaction in sibships and do not abrogate the superior power of the β model. It therefore becomes interesting to extend the β model to multiple markers. Fortunately there is a simple equivalence. One formulation of multipoint mapping expresses probabilities in terms of λ_0, λ_s, and λ_m, the recurrence risks for offspring, sibs, and monozygous twins (equation 3 in ref. 14). The β model specifies $\lambda_0 = e^{\beta}$, $\lambda_s = (1 + e^{\beta})^2/4$, $\lambda_m = e^{2\beta}$. With these substitutions, multipoint logic generates for a standard map a table of lods as a function of β, with map location of the disease locus as an ancillary parameter. There are 2 df for the sex-average map and 3 df for the sex-specific map. Then χ^2 with the appropriate df provides a test of the null hypothesis of no linked locus, no assumption being made about disease loci outside the region of interest.

The properties demonstrated herein make the β model superior to other nonparametric methods. It remains to be determined how it compares with parametric alternatives.

1. Morton, N. E. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 3471–3476.
2. Carter, C. O. (1969) *Br. Med. Bull.* **25,** 52–57.
3. Rao, D. C., Morton, N. E., Gottesman, I. I. & Lew, R. (1981) *Hum. Hered.* **31,** 325–333.
4. Lubs, M. L. (1972) *J. Paediatr.* **80,** 26–31.
5. Ebers, G. C., Bulman, D. E., Sadovnick, A. D., Paty, D. W., Warren, S., Hader, W., Murray, T. J., Seland, T. P., Duquette, P., Grey, T., Nelson, R., Nicolle, M. & Brunet, D. (1986) *N. Engl. J. Med.* **315,** 1638–1642.
6. Olmos, P., A'Hern, R., Heaton, D. A., Millward, B. A., Risley, D., Pyke, D. A. & Leslie, R. D. G. (1988) *Diabetologia* **31,** 747–750.
7. Rotter, J. I., Vadheim, C. M. & Rimoin, D. L. (1992) in *The Genetic Basis of Common Disease*, eds. King, R. A., Rotter, J. I. & Motulsky, A. G. (Oxford Univ. Press, Oxford, U.K.), pp. 413–481.
8. Strober, W. (1992) in *The Genetic Basis of Common Disease*, eds. King, R. A., Rotter, J. L. & Motulsky, A. G. (Oxford Univ. Press, Oxford, U.K.), pp. 279–304.
9. Vassilopoulos, D. (1990) in *Principles and Practice of Medical Genetics*, eds. Emery, A. E. H. & Rimoin, D. L. (Churchill Livingstone, Edinburgh), Vol. 1, pp. 473–480.
10. Morton, N. E. & Rao, D. C. (1980) *Soc. Biol.* **27,** 48–52.
11. Wright, S. (1968) *Evolution and the Genetics of Populations* (Univ. of Chicago Press, Chicago), Vol. 1, pp. 411–417.
12. Agresti, A. (1990) *Categorical Data Analysis* (Wiley, New York), p. 98.
13. Weir, B. (1990) *Genetic Data Analysis* (Sinauer, Sunderland, MA), p. 295.
14. Kruglyak, L. & Lander, E. S. (1995) *Am. J. Hum. Genet.* **57,** 439–454.
15. Payami, H., Thomsen, G., Motro, U., Louis, E. J. & Hudes, E. (1985) *Ann. Hum. Genet.* **49,** 303–314.
16. Hashimoto, L., Habita, C., Beressi, J., Delepine, M., Besse, C., Cambon-Thomsen, A., Deschamps, I., Rotter, J., Djoulah, S., James, M., Froguel, P., Weissenbach, J., Lathrop, G. & Julier, C. (1994) *Nature (London)* **371,** 161–164.
17. Dausset, J. (1982) *HLA 1982: Complex majeur d'histocompatibilité de l'homme* (Flammarion, Paris).
18. Bundy, S. (1990) in *Principles and Practice of Medical Genetics*, eds. Emery, A. E. H. & Rimoin, D. L. (Churchill Livingstone, Edinburgh), Vol. 1, pp. 447–456.
19. Davies, J. L., Kawaguchi, Y., Bennett, S. T., Copeman, J. B., Cordell, H. J., Pritchard, L. E., Reed, P. W., Gough, S. C. L., Jenkins, S. C., Palmer, S. M., Balfour, K. M., Rowe, B. R., Farrall, M., Barnett, A. H., Bain, S. C. & Todd, J. A. (1994) *Nature (London)* **371,** 130–135.
20. Todd, J. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 8560–8565.
21. Field, L., Tobias R. & Magnus, T. (1994) *Nat. Genet.* **8,** 189–194.
22. Copeman, J. B., Cucca, F., Hearne, C. M., Cornall, R. J., Reed, P. W., Ronningen, K. S., Undlien, D. E., Nistico, L., Buzzetti, R., Tosi, R., Pociot, F., Nerup, J., Cornelis, F., Barnett, A. J., Bain, S. C. & Todd, J. A. (1995) *Nat. Genet.* **9,** 80–84.
23. Wald, A. (1947) *Sequential Analysis* (Wiley, New York).
24. Haldane, J. B. S. & Smith, C. A. B. (1947) *Ann. Eugen.* **14,** 10–31.
25. Rao, D. C., Keats, B. J. B., Morton, N. E., Yee, S. & Lew, R. (1978) *Am. J. Hum. Genet.* **30,** 516–529.
26. Lander, E. & Krugylak, L. (1995) *Nat. Genet.* **11,** 241–247.
27. Cotterman, C. W. (1940) Dissertation (Ohio State Univ., Columbus); reprinted in Ballanoff, P. (1974) *Genetics and Social Structure* (Dowden, Hutchinson & Ross, Stroudsburg, PA), pp. 157–272.
28. James, J. W. (1971) *Ann. Hum. Genet.* **35,** 47–49.
29. Risch, N. (1990) *Am. J. Hum. Genet.* **46,** 222–228.
30. Mather, K. & Jinks, J. L. (1982) *Biometrical Genetics* (Chapman & Hall, London).
31. Risch, N. (1992) *Am. J. Hum. Genet.* **51,** 673–675.
32. Schwab, S., Albus, M., Hallmayer, J., Hönig, S., Borrmann, M., Lichtermann, D., Ebstein, R. P., Ackenheil, M., Lerer, B., Risch, N., Maier, W. & Wildenauer, D. B. (1995) *Nat. Genet.* **11,** 325–327.