BIOMETRICS *28*, 931–946
December 1972

THE MULTIFACTORIAL MODEL FOR THE INHERITANCE OF LIABILITY TO DISEASE AND ITS IMPLICATIONS FOR RELATIVES AT RISK

R. N. CURNOW

Department of Applied Statistics, University of Reading, Reading, Berkshire, England

SUMMARY

The multifactorial model for the inheritance of disease liability (Falconer [1965]) is discussed. In this model, the probability that an individual has the disease depends on the value of some underlying continuous quantity x. The quantity x is assumed to have a genetic component leading to correlations between relatives. For certain family groups, the probabilities of all possible patterns of disease occurrence are shown to be calculable from single integrals involving only univariate Normal density and cumulative distribution functions. Using these probabilities, the recurrence risk for an individual can be calculated from a knowledge of the occurrence of the disease in the family. Relative recurrence risks are tabulated for individuals belonging to families in which there is information on one or both parents, or on two or three full-sibs (or, equivalently, one parent and one or two full-sibs). Recurrence risks in families containing a pair of monozygous twins are also given.

1. INTRODUCTION

Models have been proposed for disease liability in which the probability that an individual succumbs to the disease depends on the value of some underlying quantity, x. This quantity is generally assumed to have a genetic component leading to correlations between relatives (see, e.g. Carter [1969], Edwards [1969], Falconer [1965; 1967] and Smith [1970; 1971]). By a suitable choice of origin and scale, the mean and standard deviation of the distribution of x over the population at risk can be taken as 0 and 1 respectively. If the distribution of x is continuous, then a suitable transformation can always be found to transform the distribution to Normality. We shall henceforth assume that this transformation has been carried out. The x-values of members of the same family will be assumed to be correlated, and the joint distribution of these x-values to be multi-Normal with the correlation coefficients representing the effects of genetic correlation, common environment, and, possibly, infections within the family. Models of this kind have been suggested for liability to, among others, the following diseases: congenital pyloric stenosis, diabetes mellitus, peptic ulcer, and talipes equinovarus (club foot) [see references]. The model has been tested for some of these diseases but little has been so far achieved in determining the nature of the underlying quantities, x.

We shall write $f_k(x_1, x_2, \cdots, x_k; \varrho)$ for the density function of the k-

variate standardized Normal distribution with correlation matrix ϱ, and $S(x)$ for the probability that an individual with value x succumbs to the disease. The (i, j)th element of ϱ will be written ρ_{ij} . The probability that all k members of the family succumb to the disease is then

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_k(x_1 , x_2 , \cdots , x_k ; \varrho) \cdot S(x_1)S(x_2) \cdots S(x_k)\, dx_1\, dx_2 \cdots dx_k , \qquad (1.1)$$

provided only that the probabilities of any set of relatives succumbing to the disease are independent given their respective x values. The probability that any particular set or number of relatives succumbs, and the remainder do not, can be calculated from integrals of the same form as (1.1) but with the $S(x)$ functions of those not succumbing replaced by $[1 - S(x)]$. This leads to a number of integrals of the same form as (1.1) but of lower, or the same, dimension.

Great simplification results if the $S(x)$ function can be assumed to be of a sigmoid form i.e.

$$S(x) = \Phi\left(\frac{x - \mu}{\sigma}\right),$$

where

$$\Phi(z), = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp - \{\tfrac{1}{2}t^2\}\, dt,$$

is the cumulative standard Normal distribution function. μ is then the value of x at which there is a probability of $\frac{1}{2}$ of succumbing to the disease and $1/\sigma$ is a measure of the sensitivity of the probability to the value of x. The probability that all k relatives succumb is then

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1 , x_2 , \cdots , x_k ; \varrho) \cdot \Phi\left(\frac{x_1 - \mu}{\sigma}\right)\Phi\left(\frac{x_2 - \mu}{\sigma}\right) \cdots \Phi\left(\frac{x_k - \mu}{\sigma}\right) dx_1\, dx_2 \cdots dx_k .$$

Because all the integrals are from $-\infty$ to $+\infty$, this is the probability that

$$z_1 < \frac{x_1 - \mu}{\sigma} , z_2 < \frac{x_2 - \mu}{\sigma} , \cdots , z_k < \frac{x_k - \mu}{\sigma} ,$$

where z_i $(i = 1, 2, \cdots , k)$ are uncorrelated standardized Normal variables distributed independently of $x_1 , x_2 , \cdots , x_k$. This probability is also the probability that

$$y_1 = \frac{\sigma z_1 - x_1}{\sqrt{(1 + \sigma^2)}} < -\Theta,$$

$$y_2 = \frac{\sigma z_2 - x_2}{\sqrt{(1 + \sigma^2)}} < -\Theta, \cdots , y_k = \frac{\sigma z_k - x_k}{\sqrt{(1 + \sigma^2)}} < -\Theta,$$

where

$$\Theta = \frac{\mu}{\sqrt{(1 + \sigma^2)}}$$

and $y_1, y_2, \cdots, y_k$ are standardized Normal variables with correlation matrix ϱ^* where $\rho_{ij}^* = \rho_{ij}/(1 + \sigma^2)(i \neq j)$. Therefore the probability that all k relatives succumb is

$$\int_{-\infty}^{-\Theta} \int_{-\infty}^{-\Theta} \cdots \int_{-\infty}^{-\Theta} f_k(y_1, y_2, \cdots, y_k ; \varrho^*)\, dy_1\, dy_2 \cdots dy_k . \qquad (1.2)$$

The probability that any particular set or number of relatives succumb and the rest do not is of the same form as (1.2) but with the integrals for those who do not succumb having lower and upper limits $-\Theta$ and $+\infty$ instead of $-\infty$ and $-\Theta$. This leads to a number of integrals of exactly the same form as (1.2) but of lower, or the same, dimension. This will be illustrated later with examples. The frequency of the disease in the population at risk corresponds to $k = 1$ in (1.2) and is

$$P_1 = \int_{-\infty}^{-\Theta} f_1(y_1)\, dy_1 = \Phi(-\Theta).$$

In setting $S(x) = \Phi[(x - \mu)/\sigma]$, we have assumed that $S(-\infty) = 0$ and $S(+\infty) = 1$. We could take $S(x) = a + (b - a)\Phi[(x - \mu)/\sigma]$ and so allow $S(x)$ to increase from $a \geq 0$ to $b \leq 1$ as x increases from $-\infty$ to $+\infty$. This would introduce two extra parameters, a and b, that would have to be estimated from population data. This may well be difficult.

The approach in this paper is mathematically equivalent to Falconer's abrupt threshold model (Falconer [1965; 1967]). In Falconer's model, a random quantity z is added to the underlying x value for an individual and all individuals with values of $(x + z)$ greater than a certain threshold, T, succumb while those with $(x + z)$ less than T do not. $S(x)$ is then the probability that $z > (T - x)$ and is therefore related to the cumulative distribution function of the quantity z. The larger the value of the threshold, T, the lower the frequency of the disease in the population. There is some arbitrariness in the division between x and z when, as is generally assumed, x is Normally distributed and the $S(x)$ function sigmoid. All that matters is that the z-values for different individuals must be independent. Although the mathematics is the same, the idea of an abrupt threshold is less acceptable biologically than the idea of a risk function (Edwards [1969] and Smith [1970; 1971]). Edwards [1969] presented a model in which x was Normally distributed but his risk function was exponential and tended to infinity as x tended to infinity, whereas a risk function, by definition, should never exceed one.

All the probabilities derived above depend only on Θ (or equivalently on P_1, the frequency of the disease in the population at risk) and on the correlation coefficients, $\rho_{ij}^* = \rho_{ij}/(1 + \sigma^2)(i \neq j)$. Smith [1970] in considering pairs of related individuals wrote $\rho_{ij} = R$ and $1/(1 + \sigma^2) = h^2$, the "heri-

tability". h^2 is a heritability only if, in the abrupt threshold model, x is entirely genetic, z entirely environmental and correlations between relatives are entirely genetic.

Most papers to date have concentrated on the estimation of h^2, or ρ_{ij}^*, from a comparison of the frequency of the disease among pairs of monozygotic twins, or among first degree relatives of those affected, with the frequency in the general population. The values of h^2, or ρ_{ij}^*, are then used to predict the frequency of the disease among less closely related individuals. These predictions can sometimes be used to check the model against data.

We shall assume that the correlation coefficients of the x-values of any pair of full-sibs or of any parent and an offspring are all equal. This would be true if the genetic variance of x was entirely additive and the environmental and infective correlations between full-sibs or between parents and offspring were all equal. The x-values of parents will be assumed to be uncorrelated. Allowance could be made for a correlation between parents but it would introduce a further parameter that would have to be estimated or specified.

The results depend directly on the values of $\rho_{ij}^* = \rho_{ij}/(1 + \sigma^2)$ rather than on the values of ρ_{ij}. For full-sibs and for parent and offspring, we shall take $\rho_{ij} = \frac{1}{2}$, the value to be expected when the correlations are entirely genetic and due to additive genetic variation. (Dominance variation would affect the full-sib correlations but not the parent-offspring correlations). Other values of ρ_{ij} and σ^2 are also covered because the results do depend only on $\rho_{ij}^* = \rho_{ij}/(1 + \sigma^2)$. All that matters is that the value of ρ_{ij}^* must be estimated appropriately from disease incidence data (see §6).

Smith [1970] divided the range of values of the underlying quantity into a large number of small non-overlapping intervals. For the ith interval he calculated the frequency f_i; the probability, P_i, that an individual with underlying value at the midpoint of the interval succumbs to the disease and the probability, P_i', that a particular relative, e.g. a full-sib, succumbs. The frequency of the disease among these relatives is then

$$\frac{\sum_i f_i P_i P_i'}{\sum_i f_i P_i},$$

where the summation is over the intervals. By increasing the number of intervals any desired accuracy can be achieved. Smith graphed for various values of $\rho_{ij}^* = Rh^2$, the frequency of the disease in relatives of affected individuals against the frequency, P_1, of the disease in the whole population. He also showed the relationship of the concordance rate of monozygotic twins to h^2, for various frequencies of the disease in the general population.

Generally, information is available about more than one close relative and these relatives may themselves be related. In a further paper, Smith [1971] discussed recurrence risks in individual families given varying amounts of information, both positive and negative, about members of the family. Smith's method was an extension of his approach to pairs of relatives. The

distributions of x for all original and all intermediate members of independent branches of the family were split into classes of calculable frequency. The probabilities of patterns of occurrence of the disease for given sets of classes were added over all combinations of classes, weighted by their frequencies. Smith discussed the calculation of confidence limits for recurrence risks; the possibility of taking account of sex and age differences in frequency and "heritability"; approximations to recurrence risks; and the use of disease rates in the whole population and disease rates among relatives to estimate genetic correlations and "heritability" and to test genetic models of inheritance. For $h^2 = 1/(1 + \sigma^2) = 0.8$, he plotted values of recurrence risks against the population frequency, P_1 , when information was available on various sets of first, second, or third degree relatives. Also tabulated were the recurrence risks for sibs when there is information on both parents and on up to four full-sibs. The heritabilities considered were 0.2, 0.5, 0.8 and 1.0 and the population frequencies were 0.001, 0.01 and 0.1.

In this paper, recurrence risks and relative recurrence risks for some of the simpler sets of information about relatives will be obtained rather more directly by showing that the multiple integrals (1.2) can sometimes be written as single integrals that can be evaluated easily on a computer. This reduction to single integral form always occurs when information is available on only two relatives (Curnow and Dunnett [1962], and by a special argument when one correlation, e.g. between parents, is zero). We shall generally be discussing first degree relatives having $\rho_{ij}^* = \rho_{ij}/(1 + \sigma^2) = 1/[2(1 + \sigma^2)]$. To simplify the notation we shall define ρ as $\rho = 1/[2(1 + \sigma^2)]$. Numerical results will be given for a range of values of ρ and P_1 when there is information on two first degree relatives, or on three full-sibs (or equivalently, one parent and two full-sibs). Families containing monozygotic twins will also be considered. Special calculations may be needed for recurrence risks when there is a more complex pattern of disease occurrence in a family. The methods of this paper may be useful when these calculations can be reduced to the evaluation of single integrals. Results obtained here and elsewhere may be useful in deriving approximate formulae for recurrence risks in more complex situations.

Wherever possible, the results in this paper have been compared with those obtained by Smith [1970; 1971]. No important discrepancies have been found.

2. DERIVATION OF FORMULAE FOR PROBABILITIES AND RELATIVE RECURRENCE RISKS

The probability that a random individual succumbs to the disease is the frequency of the disease in the population. This is, from (1.2) with $k = 1$,

$$P_1 = \int_{-\infty}^{-\Theta} f_1(y)\, dy = \Phi(-\Theta). \tag{2.1}$$

The probability that a parent and a child, or two full-sibs, are both affected is

$$P_2 = \int_{-\infty}^{-\Theta} \int_{-\infty}^{-\Theta} f_2(y_1 , y_2 , \rho^*)\, dy_1\, dy_2 , \tag{2.2}$$

$$\text{with} \quad \rho^* = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \text{where} \quad \rho = \frac{1}{2(1 + \sigma^2)} .$$

Now, P_2 = prob $(y_1 < -\Theta, y_2 < -\Theta)$, where y_1 and y_2 are standardized Normal variables with correlation coefficient ρ. Defining x, z_1 , and z_2 as independent standardized normal variables, we can write

$$y_1 = \sqrt{\rho}\, x + \sqrt{(1 - \rho)} z_1$$
$$y_2 = \sqrt{\rho}\, x + \sqrt{(1 - \rho)} z_2 ,$$

(see Curnow and Dunnett [1962] for the general method and references). P_2 becomes

$$P_2 = \text{Prob}\left[z_1 < \frac{-(\Theta + \sqrt{\rho}\, x)}{\sqrt{(1 - \rho)}} , z_2 < \frac{-(\Theta + \sqrt{\rho}\, x)}{\sqrt{(1 - \rho)}} \right]$$

or

$$P_2 = \int_{-\infty}^{+\infty} f_1(x) \left\{ \Phi\left[\frac{-(\Theta + \sqrt{\rho}\, x)}{\sqrt{(1 - \rho)}} \right] \right\}^2 dx. \tag{2.3}$$

This single integral form is much more convenient for numerical evaluation than the previous bivariate form (2.2).

By a similar argument, the probability that three full-sibs, or one parent and two children, are all affected is

$$P_3 = \int_{-\infty}^{+\infty} f_1(x) \left\{ \Phi\left[\frac{-(\Theta + \sqrt{\rho}\, x)}{\sqrt{(1 - \rho)}} \right] \right\}^3 dx \tag{2.4}$$

The probability that a child will have the disease, given that a parent or a full-sib has it, is a conditional probability and is the ratio of the probability that both are affected to the probability that one is affected i.e. P_2/P_1 . The relative risk, i.e. the risk relative to the risk for a random individual in the population, is therefore P_2/P_1^2 .

The relative risk for a child with one parent unaffected or one full-sib unaffected, will be

$$\frac{(P_1 - P_2)}{(1 - P_1)P_1} .$$

The numerator is the probability that one is affected but not the other.

Similar formulae are given in Table 1 for the relative risks when information is available on two full-sibs or on one full-sib and a parent. If neither of the two full-sibs are affected, the relative risk will be

$$\frac{(P_1 - 2P_2 + P_3)}{(1 - 2P_1 + P_2)P_1} .$$

The numerator is the probability that two full-sibs are unaffected and another

TABLE 1

FAMILIAL PATTERNS OF DISEASE AND FORMULAE FOR RELATIVE RECURRENCE RISKS

Disease Status of Relatives				Relative Recurrence Risk
Parent 1	Parent 2	Full-sib 1	Full-sib 2	
1	-	-	-	P_2/P_1^2
-	-	1	-	P_2/P_1^2
0	-	-	-	$(P_1-P_2)/F_1(1-P_1)$
-	-	0	-	$(P_1-P_2)/P_1(1-P_1)$
1	1	-	-	I/P_1^3
1	0	-	-	$(P_2-I)/P_1^2(1-P_1)$
0	0	-	-	$(P_1-2P_2+I)/P_1(1-P_1)^2$
1	-	1	-	P_3/P_1P_2
-	-	1	1	P_3/P_1P_2
1	-	0	-	$(P_2-P_3)/P_1(P_1-P_2)$
0	-	1	-	$(P_2-P_3)/P_1(P_1-P_2)$
-	-	0	1	$(P_2-P_3)/P_1(P_1-P_2)$
0	-	0	-	$(P_1-2P_2+P_3)/P_1(1-2P_1+P_2)$
-	-	0	0	$(P_1-2P_2+P_3)/P_1(1-2P_1+P_2)$

0 ≡ Unaffected

1 ≡ Affected

- ≡ Unknown

For definitions of P_1, P_2, P_3 and I, see § 2.

is affected. This can be seen from straightforward probability arguments. Alternatively, from (1.1), the numerator is

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_3(x_1, x_2, x_3; \varrho)[1 - S(x_1)][1 - S(x_2)]S(x_3)\,dx_1\,dx_2\,dx_3$$

$$= \int_{-\infty}^{+\infty} f_1(x_3)S(x_3)\,dx_3 - 2\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_2(x_1, x_3; \varrho)S(x_1)S(x_3)\,dx_1\,dx_3$$

$$+ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_3(x_1, x_2, x_3; \varrho)S(x_1)S(x_2)S(x_3)\,dx_1\,dx_2\,dx_3$$

$$= P_1 - 2P_2 + P_3 .$$

The remaining situation where information is available on only one or two first-degree relatives is when information on both parents is available.

The probability that both parents and a child succumb is, from (1.2),

$$\int_{-\infty}^{-\theta}\int_{-\infty}^{-\theta}\int_{-\infty}^{-\theta} f_3(y_1, y_2, y_3; \varrho^*)\, dy_1\, dy_2\, dy_3,$$

where

$$\varrho^* = \begin{bmatrix} 1 & 0 & \rho \\ 0 & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix}.$$

After some manipulation this multiple integral can be written in single integral form as

$$I = \int_{\sqrt{2}\theta}^{\infty} f_1(x)\Phi\left[\frac{\sqrt{2}\,\rho x - \theta}{\sqrt{(1-2\rho^2)}}\right]\{2\Phi(x - \sqrt{2}\,\theta) - 1\}\, dx.$$

The formulae, involving I, for the relative risks for a child given information on two parents are given in Table 1.

3. TABULATION OF RELATIVE RECURRENCE RISKS

The four integrals P_1, P_2, P_3 and I depend on two parameters,

$$\theta = \frac{\mu}{\sqrt{(1+\sigma^2)}} \quad \text{and} \quad \rho = \frac{1}{2(1+\sigma^2)}.$$

The value of θ determines the frequency of the disease in the population

$$P_1 = \Phi(-\theta).$$

θ values have been chosen to give the following values of P_1 :

P_1	θ
0.5	0
0.05	1.6449
0.02	2.0537
0.01	2.3263
0.005	2.5758
0.001	3.0902

ρ can take values between 0 and 0.5. $\rho = 0$ corresponds to $\sigma^2 = \infty$ and the occurrence of the disease is then unrelated to the x-value and the relative recurrence risks are all 1. $\rho = 0.5$ corresponds to $\sigma^2 = 0$. This means that x has a threshold effect and so all individuals having x above μ succumb while all individuals with x below μ do not succumb. Table 2 shows the relative recurrence risks (relative to P_1, the risk for a random member of the population) for the various familial patterns, for the six disease frequencies, P_1, and for five values of ρ from 0.1 to 0.5.

4. FAMILIES CONTAINING MONOZYGOUS TWINS

Tables 3 and 4 give values of the relative recurrence risks in families

TABLE 2

RELATIVE RECURRENCE RISKS USING INFORMATION ON ONE OR TWO FIRST DEGREE RELATIVES

Disease Status of Relatives

Parent 1		1 –	0 –	1	1	0	1 –	1 0 –	0 –
Parent 2		– –	– –	1	0	0	– –	– – –	– –
Full-sib 1		– 1	– 0	–	–	–	1 1	0 1 0	0 0
Full-sib 2		– –	– –	–	–	–	– 1	– – 1	– 0
P_1	ρ								
.5	0.1	1.06	.94	1.13	1.00	0.87	1.12	1.00	0.88
.5	0.2	1.13	.87	1.26	1.00	0.74	1.23	1.00	0.77
.5	0.3	1.19	.81	1.39	1.00	0.61	1.32	1.00	0.68
.5	0.4	1.26	.74	1.52	1.00	0.48	1.42	1.00	0.58
.5	0.5	1.33	.67	1.67	1.00	0.33	1.50	1.00	0.50
.05	0.1	1.48	.97	2.14	1.45	0.95	2.02	1.44	0.95
.05	0.2	2.10	.94	3.96	2.00	0.89	3.33	1.95	0.89
.05	0.3	2.85	.90	6.56	2.66	0.81	4.83	2.52	0.83
.05	0.4	3.77	.85	10.0	3.44	0.72	6.44	3.15	0.75
.05	0.5	4.88	.80	14.2	4.39	0.61	8.14	3.82	0.67
.02	0.1	1.72	.98	2.83	1.70	0.97	2.62	1.69	0.97
.02	0.2	2.75	.96	6.50	2.67	0.93	5.12	2.61	0.93
.02	0.3	4.16	.94	12.8	3.99	0.87	8.37	3.78	0.88
.02	0.4	6.03	.90	22.2	5.70	0.80	12.2	5.19	0.82
.02	0.5	8.47	.85	34.4	7.94	0.70	16.5	6.84	0.74
.01	0.1	1.93	.99	3.52	1.91	0.98	3.20	1.90	0.98
.01	0.2	3.39	.98	9.53	3.33	0.95	7.13	3.26	0.95
.01	0.3	5.56	.95	21.3	5.41	0.91	12.8	5.14	0.91
.01	0.4	8.66	.92	40.8	8.34	0.85	19.9	7.59	0.86
.01	0.5	12.9	.88	67.5	12.4	0.76	28.3	10.7	0.79
.005	0.1	2.16	.99	4.38	2.15	0.99	3.91	2.14	0.99
.005	0.2	4.19	.98	14.04	4.14	0.97	9.96	4.06	0.97
.005	0.3	7.46	.97	35.8	7.32	0.94	19.6	6.99	0.94
.005	0.4	12.5	.94	75.4	12.2	0.89	32.6	11.1	0.89
.005	0.5	19.9	.90	133.0	19.3	0.81	48.6	16.7	0.83
.001	0.1	2.84	1.00	7.34	2.83	1.00	6.29	2.83	1.00
.001	0.2	6.89	.99	35.0	6.86	0.99	21.9	6.79	0.99
.001	0.3	14.9	.99	120	14.8	0.97	53.3	14.3	0.97
.001	0.4	29.5	.97	317	29.2	0.94	104	27.2	0.95
.001	0.5	54.3	.95	646	53.7	0.89	174	47.4	0.90

For definitions of P_1 and ρ see section 2. Equivalent information on relatives (see Table 1) is shown together above each column.

0 ≡ Unaffected
1 ≡ Affected
– ≡ Unknown

containing monozygous twins. The correlations, ρ_{ij} and ρ_{ij}^* for monozygous twins are assumed to be double that for ordinary full-sibs. This will be true

TABLE 3

RELATIVE RECURRENCE RISKS FOR A TWIN
DISEASE STATUS OF RELATIVES

Full-sib or Parent		-	-	1	1	0	0
Twin		1	0	1	0	1	0
P_1	ζ						
.5	0.1	1.13	0.87	1.18	0.93	1.07	0.82
.5	0.2	1.26	0.74	1.35	0.85	1.15	0.65
.5	0.3	1.41	0.59	1.51	0.73	1.27	0.49
.5	0.4	1.59	0.41	1.68	0.56	1.44	0.32
.5	0.5	2.00	0	2.00	0	2.00	0
.05	0.1	2.10	0.94	2.73	1.39	2.05	0.92
.05	0.2	3.77	0.85	5.24	1.73	3.60	0.81
.05	0.3	6.21	0.73	8.37	1.93	5.85	0.67
.05	0.4	9.90	0.53	12.2	1.81	9.37	0.47
.05	0.5	20	0	20	0	20	0
.02	0.1	2.75	0.96	3.92	1.64	2.71	0.95
.02	0.2	6.03	0.90	9.34	2.37	5.84	0.87
.02	0.3	11.6	0.78	17.2	2.98	11.1	0.74
.02	0.4	21.1	0.59	27.6	3.08	20.2	0.54
.02	0.5	50	0	50	0	50	0
.01	0.1	3.39	0.98	5.19	1.86	3.35	0.97
.01	0.2	8.66	0.92	14.6	3.00	8.45	0.90
.01	0.3	18.8	0.82	29.7	4.14	18.1	0.79
.01	0.4	37.7	0.63	51.3	4.62	36.4	0.59
.01	0.5	100	0	100	0	100	0
.005	0.1	4.19	0.98	6.89	2.11	4.16	0.98
.005	0.2	12.5	0.94	22.8	3.79	12.3	0.93
.005	0.3	30.5	0.85	48.0	5.89	29.8	0.83
.005	0.4	67.5	0.67	95.8	6.94	65.6	0.64
.005	0.5	200	0	200	0	200	0
.001	0.1	6.89	0.99	13.4	2.81	6.87	0.99
.001	0.2	29.5	0.97	65.3	6.49	29.2	0.97
.001	0.3	94.8	0.91	188	12.3	93.4	0.90
.001	0.4	264	0.74	407	18.0	259	0.72
.001	0.5	1000	0	1000	0	1000	0

For definitions of P_1 and ζ, see § 2.

0 ≡ Unaffected

1 ≡ Affected

- ≡ Unknown

TABLE 4

RELATIVE RECURRENCE RISKS FOR A FULL SIB (OR PARENT) OF MONOZYGOUS TWINS

		Disease Status of Twins		
Twin 1		1	1	0
Twin 2		1	0	0
P_1	ρ			
.5	0.1	1.11	1.00	0.89
.5	0.2	1.20	1.00	0.80
.5	0.3	1.28	1.00	0.72
.5	0.4	1.33	1.00	0.67
.5	0.5	1.33	-	0.67
.05	0.1	1.93	1.43	0.95
.05	0.2	2.91	1.91	0.90
.05	0.3	3.85	2.41	0.85
.05	0.4	4.65	2.91	0.80
.05	0.5	4.88	-	0.80
.02	0.1	2.45	1.68	0.97
.02	0.2	4.26	2.54	0.94
.02	0.3	6.15	3.56	0.89
.02	0.4	7.87	4.69	0.85
.02	0.5	8.47	-	0.85
.01	0.1	2.95	1.89	0.98
.01	0.2	5.70	3.17	0.96
.01	0.3	8.80	4.82	0.92
.01	0.4	11.8	6.77	0.89
.01	0.5	12.9	-	0.88
.005	0.1	3.56	2.13	0.99
.005	0.2	7.65	3.96	0.97
.005	0.3	11.8	6.69	0.94
.005	0.4	17.7	9.81	0.91
.005	0.5	19.9	-	0.90
.001	0.1	5.52	2.82	1.00
.001	0.2	15.3	6.64	0.99
.001	0.3	29.6	13.4	0.98
.001	0.3	29.6	13.4	0.98
.001	0.4	45.5	23.8	0.96
.001	0.5	54.3	-	0.95

For definitions of P_1 and ρ see section 2. 0 denotes unaffected, 1 denotes affected, — denotes impossible situations.

if the correlations are entirely genetic and the genetic variation entirely additive. Table 3 gives the relative recurrence risks for one of the twins given information about the other twin or about the other twin and a full-sib or parent. The appropriate formulae for relative recurrence risks, derived as before, are:

Monozygous Twin	Full-sib or Parent	
1	–	T/P_1^2
0	–	$(P_1 - T)/P_1(1 - P_1)$
1	1	A/P_1P_2
1	0	$(P_2 - A)/P_1(P_1 - P_2)$
0	1	$(T - A)/P_1(P_1 - P_2)$
0	0	$(P_1 - T - P_2 + A)/P_1(1 - 2P_1 + P_2)$.

In these formulae,

$$T = \int_{-\infty}^{+\infty} f_1(x)\left\{\Phi\left(\frac{-\Theta+\sqrt{(2\rho)}x}{\sqrt{(1-2\rho)}}\right)\right\}^2 dx$$

and

$$A = \int_{-\infty}^{+\infty} f_1(x)\left\{\Phi\left(\frac{-\Theta+\sqrt{(2\rho)}x}{\sqrt{(1-2\rho)}}\right)\right\}^2 \Phi\left\{\frac{-\Theta+\sqrt{(\rho/2)}x}{\sqrt{(1-\rho/2)}}\right\} dx.$$

T is the probability two monozygous twins both have the disease and A is the probability that both twins and a full-sib, or parent, have the disease. In Table 4, the relative recurrence risks are calculated for a full-sib or parent when information is available about both monozygous twins. The appropriate formulae are:

Monozygous Twins		
1	1	$A/(P_1T)$
1	0	$(P_2 - A)/P_1(P_1 - T)$
0	0	$(P_1 - 2P_2 + A)/P_1(1 - 2P_1 + T)$.

When $\rho = 0.5$, twins must both have the disease or both not have the disease since their genetic correlation is 1 and $\sigma^2 = 0$.

5. INFORMATION ON THREE FIRST-DEGREE RELATIVES

In Table 5, the relative recurrence risks are given when there is information on three full-sibs or two full-sibs and one parent. The requisite formulae are obvious extensions of those already derived for two full-sibs or one full-sib and a parent.

6. USE OF TABLES AND OTHER APPLICATIONS

The tables in this paper can be used to calculate the recurrence risks for the relatives when the frequency of the disease in the population at

TABLE 5

RELATIVE RECURRENCE RISKS FOR A FOURTH FIRST DEGREE RELATIVE

P_1	ρ	Number of First Degree Relatives with Disease			
		0	1	2	3
.5	0.1	0.83	0.94	1.06	1.17
.5	0.2	0.69	0.90	1.10	1.31
.5	0.3	0.58	0.86	1.14	1.42
.5	0.4	0.48	0.83	1.17	1.52
.5	0.5	0.40	0.80	1.20	1.60
.05	0.1	0.93	1.40	1.96	2.59
.05	0.2	0.85	1.83	3.09	4.55
.05	0.3	0.76	2.27	4.27	6.59
.05	0.4	0.68	2.72	5.44	8.57
.05	0.5	0.58	3.18	6.57	10.4
.02	0.1	0.96	1.66	2.56	3.66
.02	0.2	0.90	2.49	4.82	7.76
.02	0.3	0.84	3.47	7.53	12.6
.02	0.4	0.76	4.58	10.5	17.6
.02	0.5	0.67	5.79	13.5	22.6
.01	0.1	0.97	1.88	3.15	4.78
.01	0.2	0.93	3.14	6.78	11.7
.01	0.3	0.88	4.79	11.6	20.6
.01	0.4	0.81	6.80	17.3	30.6
.01	0.5	0.73	9.15	23.4	40.7
.005	0.1	0.98	2.12	3.87	6.26
.005	0.2	0.95	3.95	9.56	17.7
.005	0.3	0.91	6.60	18.0	34.0
.005	0.4	0.85	10.1	28.6	53.2
.005	0.5	0.78	14.5	40.6	73.7
.001	0.1	1.00	2.82	6.26	11.8
.001	0.2	0.98	6.69	21.3	46.6
.001	0.3	0.96	13.8	50.1	110
.001	0.4	0.92	25.4	93.1	195
.001	0.5	0.86	42.4	148	295

For definitions of P_1 and ρ see section 2.

risk, P_1 , and the correlation coefficient ρ are both known. For a particular disease, the tables may first have to be used to estimate ρ from known values of P_1 and the frequency of the disease among first-degree relatives of affected individuals, then used again to calculate the risks for other sets of information about relatives.

The tables are a little difficult to interpret in any general way. Figure 1 shows the recurrence risks plotted against the frequency of the disease in the population when $\rho = 0.3$. The scales on both axes are logarithmic. Each line corresponds to different information about the family. The particular types of family have been chosen to illustrate some fairly obvious conclusions that can be drawn from the tables. In Figure 1, the highest recurrence risks occur when two uncorrelated parents are both affected. Next comes families in which a parent and a full-sib, or two full-sibs, are

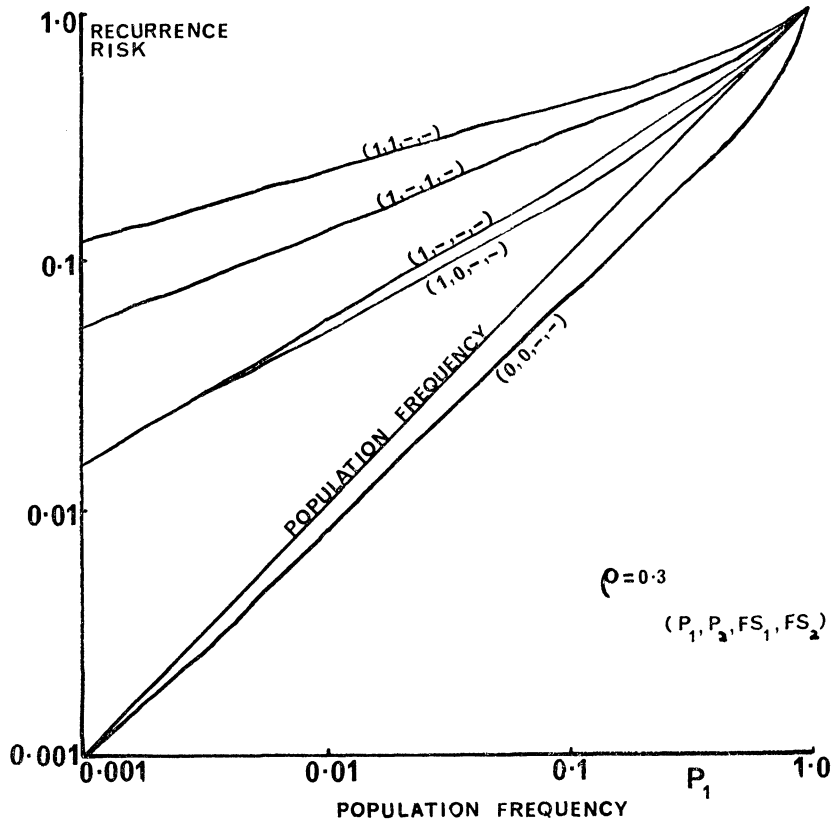

FIGURE 1

RECURRENCE RISKS WHEN $\rho = 0.3$ PLOTTED AGAINST THE POPULATION FREQUENCY OF THE DISEASE. BOTH SCALES ARE LOGARITHMIC. EACH CURVE CORRESPONDS TO DIFFERENT INFORMATION ABOUT THE FAMILY. THE ORDER OF PRESENTATION IS (PARENT 1, PARENT 2, FULL-SIB 1, FULL-SIB 2). 1 INDICATES AFFECTED; 0, UNAFFECTED; AND—, NOT KNOWN. SOME CURVES MAY REFER TO MORE THAN ONE SET OF INFORMATION ABOUT THE FAMILY—SEE TABLE 2.

affected. The line corresponding to one parent or one full-sib affected, is roughly half-way between the line for two parents affected and the line in which the recurrence risk equals the population frequency. The other two lines, (1, 0, –, –,) and (0, 0, –, –,) show that, whatever the population frequency, the knowledge that some relatives are unaffected does not appreciably reduce the recurrence risk when $\rho = 0.3$. The curves of Figure 1 can be compared with similar curves for $\rho = 0.4$ ($h^2 = 0.8$) given in Figures 3 and 4 of Smith [1971].

The results of this paper may be useful, not only in studies of disease, but in any situation where the attribute observed is the observable expression of an underlying continuous character containing a genetic component. The cause of the correlation between different observations of the attribute could also be in the nature of a repeatability rather than a heritability. This would apply, for example, when calculating the probabilities of a stillbirth; a male or a female; or twins at a future birth, given the corresponding information on previous births. Tabulations may then be needed for values of $\rho > \frac{1}{2}$.

Allowances could be made for the effects of age or sex on disease frequencies by allowing Θ to take different values for individuals of different age or sex.

ACKNOWLEDGMENTS

I am grateful to Professor D. S. Falconer, Dr. C. Smith and a referee for helpful comments on an earlier draft of this paper and to Dr. Smith for allowing me to see pre-publication copies of his two papers. I am grateful to my colleagues at Reading for helpful criticisms and for assistance with the computer programming.

LE MODELE MULTIFACTORIEL POUR LA TRANSMISSION GENETIQUE DE LA SENSIBILITE A LA MALADIE ET SES IMPLICATIONS SUR LE RISQUE A PARTIR DES PARENTS

RESUME

On discute le modèle multifactoriel pour l'étude de la transmission de la sensibilité aux maladies (Falconer [1965]). Dans ce modèle, la probabilité qu'un individu soit malade dépend d'une quantité continue sous-jacente x. x est supposé comporter une composante génétique conduisant à des corrélations entre parents. Pour certains groupes familiaux, les probabilités de tous les schémas possibles de réalisations de maladie sont calculables à partir d'intégrales comprenant seulement la densité normale à une variable et des fonctions de distribution cumulables. En utilisant ces probabilités, le risque répété pour un individu peut être calculé à partir de la connaissance de la présence de la maladie dans la famille. Des risques relatifs sont tabulés pour des individus appartenant à des familles sur lesquelles on possède de l'information concernant un (ou les deux) parent, ou deux ou trois pleins frères (ou de façon équivalente un parent et un ou deux pleins frères). Des risques récurrents dans les familles contenant une paire de jumeaux monozygotes sont donnés.

REFERENCES

Carter, C. O. [1969]. Genetics of common disorders. *Br. Med. Bull. 25*, 52–7.

Curnow, R. N. and Dunnett, C. W. [1962]. The numerical evaluation of certain multivariate normal integrals. *Ann. Math. Stat. 33*, 571–9.
Edwards, J. H. [1969]. Familial predisposition in man. *Br. Med. Bull. 25*, 58–63.
Falconer, D. S. [1965]. The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann. Hum. Genet. 29*, 51–76.
Falconer, D. S. [1967]. The inheritance of liability to diseases with variable age of onset, with particular reference to diabetes mellitus. *Ann. Hum. Genet. 31*, 1–20.
Smith, C. [1970]. Heritability of liability and concordance in monozygous twins. *Ann. Hum. Genet., 34*, 85–91.
Smith, C. [1971]. Recurrence risks for multifactorial inheritance. *Amer. J. Hum. Genet. 23*, 578–88.

Received May 1971, *Revised April* 1972

Key Words: Disease liability; Genetic counselling; Risk prediction; Human population genetics; Quantitative genetics.