Eur Arch Psychiatry Clin Neurosci (1991) 240:174–181

European Archives of Psychiatry and Clinical Neuroscience

The Genetic Epidemiology of Schizophrenia and the Design of Linkage Studies*

M. McGue[1] and I. I. Gottesman[2]

[1]Departments of Psychology and Genetics, University of Minnesota, Minneapolis, USA
[2]Department of Psychology, University of Virginia

Summary. There are three aspects of schizophrenia that are challenges to the design of linkage studies. First, analysis of twin and family data have consistently failed to identify a single major gene effect. Second, ascertainment of multiplex families does not guarantee the sampling of families in whom a major gene is segregating even if such a gene exists. Third, environmental influences appear to play an essential role in the etiology of at least some schizophrenia. The implications of these features for linkage strategies in schizophrenia are discussed.

Key words: Genetic epidemiology – Genetic models – Linkage – Schizophrenia

Introduction

Identifying the genetic mechanisms that underly individual differences in human behavior remains one of the most significant problems facing the human geneticist. Progress, impeded by the complexity of the behavioral phenotype, has been slow, however. It is little wonder then that the advent of the application of molecular genetic methods to human behavior (Gurling 1986) and the recent claims in linking behavioral conditions including affective psychoses (Egeland et al. 1987), reading disability (Smith et al. 1983), Alzheimer's disease (St. George-Hyslop et al. 1987), and schizophrenia (Sherrington et al. 1988) to specific regions of the genome has given rise to much optimism. Clearly the race has begun to apply molecular genetic techniques to the study of human behavior. As behavioral geneticists we have long been interested in the way knowledge of a genetic influence will bring about a better understanding of the development of psychopathology with an eye toward rational intervention and even prevention. It appears an especially appropriate time to consider whether behavioral pathology is best treated as other human genetic disorders or whether the study of genetic influences on behavior will require new and unique research approaches.

Throughout its development, behavioral genetics has relied heavily upon a Galtonian biometrical approach. This approach, although statistically sophisticated, has been widely criticized because of its inability to characterize the mechanisms of genetic influence (e.g., Vogel and Motulsky 1986). Over the past 22 years, we have engaged in biometrical analyses of schizophrenia twin and family data (Gottesman and Shields 1967; Rao et al. 1981; McGue et al. 1983, 1986). Although we agree that these analyses do not allow us to identify the underlying genetic processes, we believe they do help us to understand the genetics of schizophrenia. Indeed, as research embarks upon intensive molecular genetic studies of schizophrenia, the results of our biometrical analyses may have implications for the probable success of alternative strategies for identifying single gene effects on schizophrenia. We consider those results and some of their implications here.

Genetic Epidemiology of Schizophrenia

The most powerful and consistent predictor of schizophrenia risk is being an identical twin or first-degree relative of a schizophrenic (Gottesman et al. 1982; Eaton 1985). Table 1 gives estimated life time risks (i.e., age corrected values) for developing schizophrenia among the relatives of schizophrenics pooled from systematic studies undertaken in western Europe since 1920 (for a complete description of how the data were compiled see Gottesman et al. 1982; Slater and Cowie 1971). Values for a broader definition including "probable schizophrenia" are about 25% higher than those given in Table 1. With the exception of the twin data, where a large correlation in age at onset as well as continued follow-up obviates the need, all risks have been adjusted for variable age of onset using variations of the Weinberg method (Gottesman et al. 1982).

* A previous version of this paper was published in Schizophrenia Bulletin (1989, 15:453–464) and reprinted in V. Bulyzhenkov, Y. Chirsten, L. Prilipko (eds) (1990) Genetic approaches in the prevention of mental disorders, Berlin: Springer-Verlag, pp 24–38

Offprint requests to: M. McGue, Department of Psychology, 75 East River Road, Minneapolis, MN 55455, USA

Table 1. Rates of definite schizophrenia among the relatives of schizophrenics

Familial relationship	BZN[a]	% Affected
Offspring of two schizophrenics	134	36.6
MZ Twins	106	44.3
DZ Twins	149	12.1
Siblings	7523	7.3
Offspring of one schizophrenic	1678	9.4
Half-siblings	442	2.9
Nieces or nephews	3965	2.7
Grandchildren	739	2.8
First cousins	1600	1.6
Spouses	399	1.0

[a] BZN gives the age adjusted sample size (Gottesman et al. 1982)
DZ: dizygotic

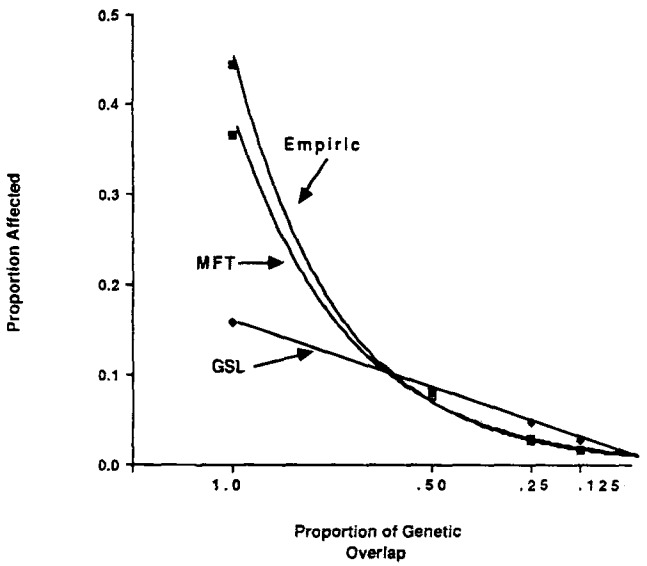

Fig. 1. Lifetime morbid risks among the relatives of schizophrenics observed (*Empiric*) and expected under a multifactorial threshold model with 80% heritability (MFT) and the generalized single locus model (*GSL*) proposed by Matthysse et al. (1986) as a function of proportion of genes shared with the proband. Proportion of genetic overlap is 1.0 for MZ twins, 0.50 for first-degree relatives, 0.25 for half-siblings, and 0.125 for third-degree relatives

There is a strong association between the magnitude of the risk in the relative and the degree of genetic relationship to the proband. The most distinctive features of this association are the monozygotic (MZ) twin concordance rate and the risk to the offspring of two schizophrenics, both of which are large relative to the risks among other family members. The MZ rate is approximately five times the rate among first-degree relatives, which in turn is approximately two and one half times the rate among second-degree-relatives. The association between familial risk and proportion of genes shared with the index case is approximated by an exponential decay function (Fig. 1). Alternative hypotheses about the genetic transmission of schizophrenia are distinguished by their ability to account for this essential epidemiological association.

Table 2. Parameterization of the generalized single-locus model

Genotype	*AA*	*Aa*	*aa*
Frequency	q^2	$2q(1-q)$	$(1-q)^2$
Penetrance	f_{AA}	f_{Aa}	f_{aa}

Twin, adoption and family studies are consistent in indicating that the familial aggregation of schizophrenia is accounted for largely, if not entirely, by genetic factors (Rosenthal 1972; McGue et al. 1986). Nonetheless, the MZ twin concordance rate is substantially less than 100%, and thus (non-familial) environmental factors, be they prenatal, perinatal or sociocultural, play a significant role in the etiology of at least some forms of schizophrenia. The existence of environmental influences serves to obscure the mechanism of genetic transmission resulting in continued debate as to whether schizophrenia is a single gene disorder, multifactorially transmitted or etiologically heterogeneous (Faraone and Tsuang 1985). This is not a purely academic concern as the likelihood of identifying single gene effects on schizophrenia risk depends upon the mode of transmission of the disorder.

Single Gene Transmission

The earliest hypothesis (Rosanoff and Orr 1911; Rüdin 1916) for the genetic basis of schizophrenia is that it is a single gene defect (Böök 1953; Slater 1958; Heston 1970; Slater and Cowie 1971; Matthysse et al. 1986). Under the so-called generalized single locus (GSL) model, the transmission of schizophrenia is attributed solely to the segregation of two alleles (a schizophrenia-promoting allele, e.g., *A*, and a non-schizophrenia promoting allele, e.g., *a*) at a single locus. Non-Mendelian familial risks are reconciled with single gene transmission by invoking reduced penetrance (i.e., individuals who inherit at least one copy of the disease-promoting gene do not necessarily develop the disorder) and allowing for the existence of sporadics (i.e, individuals who inherit no copies of the disease-promoting gene but who are still at risk). Table 2 gives the parameterization of the GSL model in terms of the frequency of the schizophrenia promoting allele (*a*) and the penetrances of the three genotypes (f_{AA}, f_{Aa}, and f_{aa}).

Both the parsimony as well as the assumption of a unitary genetic etiology for schizophrenia that promises success for molecular genetic strategies make the GSL model attractive. Nonetheless, the GSL model has repeatedly failed to account for the observed pattern of familial risk in schizophrenia (Elston and Campbell 1970; Kidd and Cavalli-Sforza 1973; Elston et al. 1978; O'Rourke et al. 1982; Tsuang et al. 1982; Baron 1986; McGue et al. 1986). This failure stems directly from the inability of the GSL model to predict an exponential relationship between familial risk and degree of the relationship with the proband.

James (1971) has shown how precise quantitative predictions of familial risks can be determined under a GSL model. The risk to relatives of a given class (K_r) can

be expressed as a function of three parameters derivable from the penetrances and gene frequency of the GSL model; the population prevalence of the disorder (K_p), the additive genetic variance (V_a), and the dominance genetic variance (V_d);

$$K_r = K_p + (\mu_1 V_a + \mu_2 V_d)/K_p \quad (1)$$

where μ_1 and μ_2 are the probabilities that the proband and the relative share, respectively, one or two alleles at a locus identical by descent.

In the absence of dominance (as appears to be the case for schizophrenia, where the risk to the siblings is not greater than the risk to the offspring of schizophrenics), predicted familial risk under the GSL model is given by a linear function of proportion of genetic overlap between the relative and the proband. That is, in the absence of dominance, Eq. 1 is reduced to

$$K_r = K_p + (V_a/K_p)\mu_1. \quad (2)$$

This linear prediction is in sharp contrast to the observed exponential relationship. Along with the empiric risks, Fig. 1 plots the familial risks predicted by a recently proposed, and illustrative, GSL model of schizophrenia (Matthysse et al. 1986). The degree of reduced penetrance required under a GSL model to account for the low familial recurrence of schizophrenia is clearly inconsistent with the relatively large MZ twin concordance rate. In those cases where the GSL model did fit schizophrenia family data, observations were made only on first-, second-, and third-degree relatives, but not twins or the offspring of dual matings (e.g., Holzman et al. 1988). Study of Fig. 1 shows that the linear relationship between risk and shared genetic background predicted by the GSL model holds approximately only for these classes of relatives.

Failure of the GSL model leads not only to rejection of simple single gene explanations for the transmission of schizophrenia but also to rejection of the hypothesis that schizophrenia represents a heterogenous mixture of single gene disorders only. Suppose that there were k different single gene defects, each giving rise, with possibly reduced penetrance, to schizophrenia. Neglecting those exceedingly rare cases who inherit more than one of these genes, the composite risk to a relative of a schizophrenic index case is a probability-weighted function of that relative's risk under each of the k separate single gene models. That is, the composite familial risk, being a weighted average of k separate linear functions, will also be a linear function of proportion of genes shared with the proband. The failure of simple single gene models clearly complicates the search for single gene effects.

Multiple Gene Models

Multiple gene models do not share the empirical shortcomings of the GSL model. Although many alternative multiple gene models have been proposed, we will focus upon the two most widely applied, the multifactorial threshold (MFT) and mixed models. Under a MFT model, genetic factors are assumed to be polygenic. That is, there is a large number of genes, each of small and equal effect, that combine additively with the effects of other genes and environmental factors to influence schizophrenia liability. The qualitative phenotype (here a diagnosis of schizophrenia) is assumed to arise when an individual's combined liability exceeds some threshold value along the unobserved liability continuum (Falconer 1965; Gottesman and Shields 1967).

The MFT model does an excellent job in accounting for the distinctive pattern of familial risk in schizophrenia (McGue et al. 1983). Figure 1 gives the predicted rates of schizophrenia under an MFT model that assumes all familial transmission is owing to polygenic factors with a multifactorial heritability of 80%, the remaining 20% of the liability variance being attributed to non-familial environmental effects (McGue et al. 1985). Although the predicted MZ concordance rate is somewhat low, this model is statistically consistent with the observed familial rates and, more importantly, produces a risk function with the characteristic exponential decline.

Despite its predictive adequacy, there is a general reluctance to accept the MFT model as an explanation for the transmission of schizophrenia. This reluctance stems, perhaps, from the failure to identify specific genetic and environmental contributors to the assumed multifactorial liability. Additionally, strict polygenic inheritance (i.e., many genes all of small effect) might preclude, for the near future, attempts at identifying single gene effects on schizophrenia risk through molecular genetic approaches. However, the fit of the MFT model does not preclude the existence of a single major gene whose effect upon schizophrenia risk is large relative to the effects of other (poly)genes, the possibility of rare single gene disorders that give rise to schizophrenia, or a tractable MFT model of a limited number of polygenes (3, 4, or 5) each with a "subcomponent effect" on schizophrenia (cf. Wright 1934; Thoday 1967). Molecular genetic strategies could, presumably, be tailored for each of the possibilities.

Meehl (1972a, b) was the first to suggest that both a major gene and polygenes play a role in the etiology of schizophrenia. Under Meehl's theory, inheritance of a single gene gives rise to a neural integrative deficit termed schizotaxia that is expressed at the personality level as schizotypy. Expression of clinical schizophrenia among individuals who inherit the single gene defect is postulated to be a function of status on a host of polygenical and environmentally influenced potentiators including anxiety, anhedonia, and social introversion.

The type of model proposed by Meehl has been termed the mixed (i.e., mixed major and polygenes) model by Morton and McClean (1974) who also developed analytical procedures for fitting the model to family data. Although the procedures outlined by Morton and McClean can be used, in theory, to identify major gene effects against a polygenic background, in practice the analysis of qualitative family data under the mixed model has yielded equivocal results. There have been three mixed model analyses of schizophrenia family data (Carter and Chung 1980; Risch and Baron 1984; Vogler et al. 1990).

In all three cases, the researchers failed to reject the multifactorial model that included polygenic effects only, in favor of a mixed model that included both a single major gene and polygenic effects. That is, they failed to find support for a single major gene effect upon schizophrenia risk. This failure may result from a lack of statistical power *or* to an absence of a single major gene effect on schizophrenia. With the given data it is difficult to choose between these two possibilities, although the repeated failure to identify single major gene effects on schizophrenia despite relatively large family studies suggests that if single gene effects exist they may be of modest magnitude only.

Simulation Studies

In order to identify the characteristics of mixed single gene/polygene models that generate accurate predictions of familial risk, we recently completed a simulation study (Gottesman and McGue 1990). In that study, familial risks under 275 different models were computed and compared with observed schizophrenia risks. Family data were generated under a general model that allowed for three additive contributors to schizophrenia liability; (1) a major gene component with two alleles (*A* and *a*) at a single locus in Hardy-Weinberg equilibrium (2) a polygenic component that was assumed to be normally distributed with a constant variance for the three genotypes (*AA*, *Aa*, and *aa*), and (3) a *non-familial* environmental component, also assumed to be normally distributed with a constant variance. *Familial* environmental effects, for which there is little empirical evidence (McGue et al. 1985), were not modeled in the simulation. The penetrances of the three genotypes were ordered according to $f_{AA} > f_{Aa} > f_{aa}$, with the heterozygote penetrance, f_{Aa}, constrained to equal the average of the penetrances of the two homozygotes as is expected when there is no dominance variance at the single locus. In all cases, the lifetime prevalence of schizophrenia was fixed at 1.0%.

Familial risks were functions of three (input) parameters: (1) f_{AA}, the penetrance of the most frequently affected genotype, (2) s, the percentage of schizophrenics who do not have any copies of the "schizophrenia gene" (i.e., the proportion of schizophrenics with the *aa* genotype termed the proportion of sporadics), and (3) h^2, the residual polygenic heritability (i.e, the proportion of liability variance due to polygenic factors after the major gene effect has been partialled out). Table 3 gives illustrative findings from these simulations. Three general conclusions were drawn;

1. When the penetrance of the most frequently affected genoytype was high (f_{AA} greater than or equal to 0.4), predicted familial risks were inconsistent with observed risks unless *both* the percentage of schizophrenics without the major gene was high (s greater than or equal to 0.60) and the residual heritability was large (h^2 greater than equal to 0.60). Put another way, if there exists a highly penetrant major gene for schizophrenia, simulations of family data suggest that few schizophrenics possess it. This is illustrated in Table 3 by the comparison of the three inconsistent models numbered 5, 6 and 7 with the consistent model numbered 2.

2. When the penetrance of the most frequently affected genotype was low (f_{AA} less than or equal to 0.2), predicted familial risks were inconsistent with observed risks unless the residual heritability was large (h^2 greater than equal to 0.60). Put another way, a low-penetrance gene is consistent with schizophrenia family data only when there is also a substantial polygenic effect. This is illustrated in Table 3 by the two consistent models numbered 1 and 3 and the two inconsistent models numbered 8 and 9.

3. A pure MFT model with large heritability (h^2 equal to 0.80) yields familial risk consistent with observed risks.

Table 3. Illustrative results from the mixed model simulation of schizophrenia

	Model						Predicted risks to relatives of schizophrenics (%)			
	Input parameters			Derived parameters – % total variance due to			MZ	1st	2nd	3rd
	f_{AA}	s	h^2	Major gene	Polygene	Env				
Consistent models:										
No. 1	0.10	0.80	0.80	2.2	78.3	19.5	41.3	9.3	3.6	2.0
No. 2	0.60	0.80	0.80	2.3	78.2	19.5	45.1	11.6	4.7	2.6
No. 3	0.10	0.025	0.60	32.9	40.3	27.8	43.6	10.8	4.1	2.2
No. 4	0.0	1.0	0.80	0.0	80.0	20.0	38.3	8.6	3.3	1.9
Inconsistent models:										
No. 5	0.60	0.60	0.20	4.9	19.0	76.1	17.1	8.1	4.3	2.6
No. 6	0.60	0.05	0.60	18.9	48.7	32.4	59.2	22.0	9.8	5.0
No. 7	0.60	0.05	0.20	18.9	16.2	64.9	37.8	17.2	8.6	4.7
No. 8	0.10	0.025	0.40	32.9	26.8	40.3	24.9	7.6	3.3	2.0
No. 9	0.10	0.80	0.40	2.2	39.1	58.7	10.5	4.0	2.1	1.5

Note: f_{AA} is the penetrance of the most frequently affected genotype, s is the proportion of schizophrenics who do not carry the major gene and h^2 is the residual multifactorial heretability. Table from Gottesman and McGue (1990)

Table 4. Distribution of number of affected nuclear family members under four alternative models for schizophrenia

No.	Model								Total[a]
	Input parameters			Number (%) of affected family members per family with at least one affected					
	f_{AA}	s	h^2	1	2	3	4 or more		
1	0.10	0.80	0.80	1156 (0.912)	97 (0.076)	14 (0.011)	1 (0.001)		1268
2	0.60	0.80	0.80	1133 (0.909)	99 (0.079)	12 (0.010)	3 (0.002)		1247
3	0.10	0.025	0.60	1172 (0.894)	112 (0.085)	24 (0.018)	3 (0.003)		1311
4	0.0	1.0	0.80	1214 (0.918)	88 (0.067)	17 (0.013)	4 (0.003)		1323

[a] Total gives the total number of nuclear families, out of the 50,000 generated, with at least one schizophrenic member

This is illustrated by the consistent model numbered 4 in Table 3.

These simulations suggest that three alternative classes of genetic models are consistent with the genetic epidemiology of schizophrenia. First, a heterogeneity-like model where the minority of schizophrenics inherit a highly penetrant but low-frequency gene (gene frequencies for models numbered 1 and 2 were, respectively, 0.018 and 0.003), while the majority are affected because of a high multifactorial loading. Second, a gene of modest-effect model where a low penetrance (about 10%) moderately prevalent (gene frequency for model 3 was 0.097) gene contributes along with a sizable multifactorial component to schizophrenia risk. Third, a pure MFT model. Any one of these models could account for the failure of mixed model analyses of schizophrenia family data to identify a single major gene effect.

Ascertainment Strategies

The results of our simulations suggest that the existence of a single major gene effect is not inconsistent with the observed schizophrenia family data, although the magnitude of this effect on overall risk may not be great. If a single major gene exists, the question remains as to how best to sample pedigrees of families for linkage studies who are informative with respect to the major gene. One popular, and seemingly sensible, strategy for identifying families for intensive molecular genetic study is to sample "loaded pedigrees" (i.e., so-called multiplex families with a large number of affected individuals). We were interested in determining, for the four models found to generate risk rates consistent with the observed familial rates, the extent to which (1) families with multiply affected members are expected to occur, and (2) multiplex ascertainment schemes succeed in enriching the sample with families who are segregating for the major gene.

For each of the four models, 50,000 nuclear families consisting of an index member (not necessarily affected) and possibly a spouse (probability of marriage was 0.864) and up to six children were generated according to the parameters of that model. Probability of marriage and number of offspring were chosen to reflect demographic features of the U.S. adult population. No adjustment was made for the reduction in fertility known to be associated with schizophrenia (Erlenmeyer-Kimling 1978; Vogel 1979, Ødegaard 1981). The average number of offspring was 2.22 (SD = 1.4, range = 0–6).

Table 5. Proportion of schizophrenics without major gene [$P(G^-|S^+)$] and proportion of normals with the major gene [$P(G^+|S^-)$] under three alternative ascertainment schemes

No. of affected family members	Model 1		Model 2		Model 3	
	$P(G^-\|S^+)$	$P(G^+\|S^-)$	$P(G^-\|S^+)$	$P(G^+\|S^-)$	$P(G^-\|S^+)$	$P(G^+\|S^-)$
1 or more	0.748	0.089	0.879	0.035	0.026	0.480
2 or more	0.658	0.089	0.789	0.054	0.000	0.427
3 or more	0.652	0.028	0.542	0.125	0.000	0.516

Table 4 gives the distribution of number of affected individuals among families with at least one affected member. Two features of Table 4 warrant comment. First, all four models predict that in a large percentage of cases schizophrenics will be the only affected members of their nuclear family. This prediction is in accord with the distribution of number of affected family members observed in large family studies (e.g., Lindelius 1970). Second, although rare, pedigrees with multiple affected members are expected under all four models of transmission. Although expected under highly penetrant single gene transmission, the observation, especially under uncertain ascertainment, of loaded pedigrees does not allow unequivocal inference of mode of transmission. Indeed, of the total of 200,000 nuclear families generated, only two contained as many as five affected members. One family was generated under model number 3, a mixed model with high multifactorial heritability, and the other under model number 4, a pure MFT model.

Table 5 gives the sample proportions of (1) affected individuals who carry no copies of the schizophrenia-promoting gene [$P(G^-|S^+)$], and (2) normal individuals who carry at least one copy of the schizophrenia-promoting gene [$P(G^+|S^-)$] under alternative sampling schemes for each of the three models for which there is a major gene effect. These proportions can be interpreted loosely as error rates under the alternative ascertainment schemes. That is, the chance that an affected individual in the sample does not have the gene and the chance that a normal individual does. As is evident from Table 5, multiplex sampling strategies are not expected necessar-

ily to enrich the sample for schizophrenics possessing the gene under any of the three models. Only for model three, where the gene frequency is high but the penetrance low, are most sampled families expected to be segregating the major gene. Nonetheless, in this case a significant proportion of unaffected individuals also possess the gene.

The Role of the Environment

The existence of environmentally induced schizophrenia-like conditions (Davison 1987) as well as MZ twin concordance rates substantially less than 100%, prompts the question: Is the influence of the environment on schizophrenia due primarily to environmentally induced phenocopies of the disorder, or is it that environmental effects combine multifactorially with an underlying genetic diathesis? We note that the popular distinction between sporadic and familial forms of psychopathology presumes that a major role of the environment is to induce nontransmitted forms of the disorder (see, however, Eaves et al. 1986 for a critical evaluation of this distinction as applied in genetic epidemiology).

Gottesman and Bertelsen (1989) evaluated schizophrenia risk among the offspring of Fischer's concordant and discordant twins. Relatively low rates of schizophrenia are expected among the offspring of discordant MZ twins if discordance among the genetically identical parents is due largely to environmentally induced phenocopies. Table 6 gives the essential findings from this study. Although the small sample allows only preliminary and cautiously drawn conclusions, the pattern of offspring risk is striking. For the discordant MZ twins, the rate of schizophrenia-like psychosis is similar among the offspring of the affected and the unaffected twins, and both rates are comparable to the overall risk to the offspring of a schizophrenic parent (Table 1). In contrast, for the discordant DZ twins, the rate of schizophrenia-like psychosis is significantly higher among the offspring of the affected twin than among the offspring of the unaffected twin. Furthermore, the rate among the offspring of the affected DZ twins is comparable to the rate among the offspring of schizophrenics, while the rate among the offspring of the unaffected co-twins is comparable to the risk among the second-degree relatives of schizophrenics (Table 1).

Table 6. Schizophrenia and schizophrenia-like psychosis in offspring of discordant twins

Parent status	Number	Affected	MR %
Monozygotic sample			
Affected twin	14	1	10.0 ± 9.0
Unaffected twin	24	4	17.4 ± 7.7
Dizygotic sample			
Affected twin	13	1	8.3 ± 7.6
Unaffected twin	52	1	2.1 ± 2.1

"Affected" column gives the total number of offspring affected with either schizophrenia or a schizophrenia-like psychosis, MR gives the estimated lifetime morbid risk using the Kaplan-Meier age correction procedures. (From Gottesman and Bertlesen 1989)

These data suggest that the discordance among the MZ twins cannot be attributed wholely to the existence of non-transmissible forms of schizophrenia. Apparently, the expression of schizophrenia depended upon both an inherited genetic diathesis, which was transmitted regardless of whether the diathesis was phenotypically expressed, and exposure to environmental stressors, to which the twins are differentially exposed. The possibility that the expression of schizophrenia might depend upon exposure to requisite environmental triggers would seem to complicate, but certainly not preclude, genetic linkage studies, which are highly sensitive to the existence of "false negatives" in the family.

Discussion and Conclusion

Alternative approaches to identifying single major gene effects on schizophrenia will be constrained, encouraged and/or defeated by salient features of the genetic epidemiology of this complex disorder. Here we have reviewed three such features, all of which have significant implications for designing linkage studies of schizophrenia.

First, nobody has ever been able to demonstrate statistically that a single major gene accounts for a large share of the overall risk for schizophrenia. We suggest that this reflects more about the nature of this complex disorder than it does about limitations of the statistical procedures. Approaches premised upon the hypothesis that schizophrenia is a unitary single gene disorder run counter to a vast amount of empirical data that suggests otherwise. At the least, such approaches need to be justified relative to their more empirically attractive alternatives.

Single gene effects on schizophrenia risk may exist. Our simulations suggest that if they do, they are likely to be the result of either a highly prevalent gene with a very low penetrance or a very low prevalent gene with a high penetrance. In either case, the contribution to overall schizophrenia risk is modest. For example, Huntington's disease (HD) is often misdiagnosed as paranoid schizophrenia; some 20% of HD cases present with paranoid schizophrenia phenotypes. HD is quite rare in the population with an incidence of 5 per 100,000. From the careful total population study in southern Sweden by Essen-Möller (Essen-Möller et al. 1956) that used in-depth interviews by psychiatrists with every inhabitant, we obtain a lifetime risk of 139 per 10,000 for schizophrenia. We can now answer the question: What proportion of schizophrenia-like psychoses are actually caused by what we now know to be a mutated dominant gene on chromosome 4 leading to HD? The answer is found by dividing the two population values or 5/100,000 by 139/10,000 and then taking 20% to get those HD cases who are, in this instance, "genocopies" of schizophrenia. The result of the calculations is that 7 in 10,000 schizophrenics have this "major gene for schizophrenia-like psychosis" and the gene would be identified by studying such special

pedigrees with RFLPs and a lucky choice of chromosome 4 as a starting place.

Second, ascertainment schemes aimed *only* at identifying "loaded pedigrees" may be useful, but then again they may not. Obviously, linkage studies require loaded pedigrees to provide powerful tests for single gene effects. Nonetheless, multiplex families are expected under GSL and MFT transmission. This may leave the molecular geneticist somewhat uneasy in claiming that the family he or she is expending great effort on happens to be a family segregating for the major gene. It is notable in this regard that the linkage studies for mental disorder that are currently available, whether for schizophrenia or affective psychoses, were conducted without defining a sampling framework. We have no way of knowing how many families in the general population were, in effect, screened to find the interesting multiplex families utilized to obtain the significant lod scores. Under such conditions, the linkage studies resemble the important data available from individual case histories provided by such pioneers as Freud, Kraepelin, and E. Bleuler. As such, pedigrees with impressive lod scores as well as seminal case histories can serve as hypothesis-generating sources but not as sufficient proof of etiology.

Third, environmental influences play an essential role in the etiology of schizophrenia. In the rush to molecular biology, it appears shortsighted for psychopathologists to consider no longer why it is that of two individuals, both of whom inherit a genetic diathesis, one will go on to develop the disorder while the other will not. It seems that progress will be maximized by tying molecular genetic approaches to further inquiry into environmental influence.

Research on the genetic epidemiolgoy of schizophrenia suggests that it is a highly complex genetic disorder whose precise mechanism of transmission remains uncertain. This will hardly come as a revelation to anyone. The pathway from gene product to behavioral expression is obviously long and presumably provides many opportunities for environmental modulation as well as moderation by other biological, physiological and behavioral systems. The transmission of schizophrenia appears wholly unlike the transmission of Huntington's disease and cystic fibrosis. Coronary heart disease and diabetes may represent more appropriate models for designing linkage studies of schizophrenia. The challenge that psychopathology presents to geneticist is not so much in finding multiplex families that can be intensively studied, but rather in adapting approaches that have proven useful with relatively simple medical genetic disorders to account for the complexities, heterogeneity, and environmental sensitivity of human behavior.

References

Baron M (1986) Genetics of schizophrenia: I. Familial patterns and mode of inheritance. Biol Psychiatry 21:1051–1066

Böök JA (1953) A genetic and neuropsychiatric investigation of a North Swedish population. Acta Genet Stat Med (Basel) 4:1–100

Carter CL, Chung CS (1980) Segregation analysis of schizophrenia under a mixed model. Hum Hered 30:350–356

Davison K (1987) Organic and toxic concomitents of schizophrenia: association or chance? In: Helmchen H, Henn FA (eds) Biological perspectives of schizophrenia. Wiley, New York, pp 139–160

Eaton WW (1985) Epidemiology of schizophrenia. Epidemiol Rev 7:105–126

Eaves LJ, Kendler KS, Schulz SC (1986) The familial sporadic classification: its power for the resolution of genetic and environmental etiologic factors. J Psychiatry Res 20:115–130

Egeland JA, Gerhard DS, Pauls DL, Sussex JN, Kidd KK, Allen CR, Hostetter AM, Housman DE (1987) Bipolar affective disorders linked to DNA markers on chromosome 11. Nature 325:783–787

Elston RC, Campbell MA (1970) Schizophrenia: evidence for a major gene hypothesis. Behav Genet 1:3–10

Elston RC, Namboodiri KK, Spence MA, Rainer JD (1978) A genetic study of schizophrenia pedigrees: II. One-locus hypotheses. Neuropsychobiology 4:193–206

Erlenmeyer-Kimling L (1978) Fertility of psychotics: demography. In: Cancro R (ed) Annual review of the schizophrenic syndrome, vol 5. Brunner Mazel, New York, pp 298:333

Essen-Moller E, Larson H, Uddenberg CE, White G (1956) Individual traits and morbidity in a Swedish rural population. Acta Psychiatr Neurol Scand Suppl 100

Falconer DS (1965) The inheritance of liability to certain diseases estimated from the incidence among relatives. Ann Hum Genet 29:51–76

Faraone SV, Tsuang MT (1985) Quantitative models of the genetic transmission of schizophrenia. Psychol Bull 98:41–66

Gottesman II, Bertelsen A (1989) Confirming unexpressed genotypes for schizophrenia: risks in the offspring of Fischer's Danish identical and fraternal discordant twins. Arch Gen Psychiatry 46:867–872

Gottesman II, McGue M (1990) Mixed and mixed-up models for the transmission of schizophrenia. In: Cichetti D, Grove WM (eds) Thinking clearly about psychology: essays in honor of Paul E. Meehl. Volume II: Essays on Individual Differences. University of Minnesota Press, Minneapolis

Gottesman II, Shields J (1967) A polygenic theory of schizophrenia. Proc Natl Acad Sci USA 58:199–205

Gottesman II, Shields J, Hanson DR (1982) Schizophrenia: the epigenetic puzzle. Cambridge University Press, Cambridge

Heston LL (1970) The genetics of schizophrenia and schizoid disease. Science 167:249–256

James JW (1971) Frequency in relatives for an all-or-none trait. Ann Hum Genet 35:47–49

Kidd KK, Cavalli-Sforza LL (1973) An analysis of the genetics of schizophrenia. Soc Biol 20:254–265

Lindelius R (1970) A study of schizophrenia: a clinical, prognostic, and family investigation. Acta Psychiatr Scand Suppl 216

Matthysse SW, Holzman PS, Lange K (1986) The genetic transmission of schizophrenia: application of Mendelian latent structure analysis to eye tracking dysfunctions in schizophrenia and affective disorder. J Psychiatr Res 20:57–65

McGue M, Gottesman II, Rao DC (1983) The transmission of schizophrenia under a multifactorial threshold model. Am J Hum Genet 35:1161–1178

McGue M, Gottesman II, Rao DC (1985) Resolving genetic models for the transmission of schizophrenia. Genet Epidemiol 2:99–110

McGue M, Gottesman II, Rao DC (1986) The analysis of schizophrenia family data. Behav Genet 16:75–87

Meehl PE (1972a) A critical afterword. In: Gottesman II, Shields J (eds) Schizophrenia and genetics. A twin study vantage point. Academic Press, New York, pp 367–415

Meehl PE (1972b) Specific genetic etiology, psychodynamics, and therapeutic nihilism. Int J Ment Health 1:10–27

Morton NE, MacLean CJ (1974) Analysis of family resemblance: III. Complex segregation analysis of quantitative traits. Am J Hum Genet 26:489–503

Ødegaard Ø (1980) Fertility of psychiatric first admissions in Norway 1936–1975. Acta Psychiatr Scand 62:212–220
O'Rourke DH, Gottesman II, Suarez BK, Rice J, Reich T (1982) Refutation of the general single-locus model for the etiology of schizophrenia. Am J Hum Genet 34:630–649
Rao DC, Morton NE, Gottesman II, Lew R (1981) Path analysis of qualitative data on pairs of relatives: application to schizophrenia. Hum Hered 33:325–333
Risch N, Baron M (1984) Segregation analysis of schizophrenia and related disorders. Am J Hum Genet 36:1039–1059
Rosanoff AJ, Orr FI (1911) A study in insanity in the light of the Mendelian theory. Am J Insanity 68:221–261
Rosenthal D (1972) Three adoption studies of heredity in the schizophrenic disorders. Int J Ment Health, 1:63–75
Rüdin E (1916) Zur Vererbung und Neuentstehung der Dementia Praecox. Springer, Berlin
Sherrington R, Brynjolfsson J, Petursson H, Potter M, Dudleston K, Barraclough B, Wasmuth J, Dobbs M, Gurling H (1988) Localization of a susceptibility locus for schizophrenia on chromosome 5. Nature 326:164–167
Slater E (1958) The monogenic theory of schizophrenia. Acta Genet Stat Med (Basel) 8:50–56
Slater E, Cowie V (1971) The genetics of mental disorders. Oxford Press, London
Smith SD, Kimberling WJ, Pennington BF, Lubs HA (1983) Specific reading disability: identification of an inherited form through linkage analysis: Science 219:1345–1347
St. George-Hyslop PH, Tanzi RE, Polinsky RJ, Haines JL, Nee L, Watkins PC, Myers RH, Feldman RG, Pollen D, Drachman D, Growdon J, Bruni A, Foncin JF, Salmon D, Frommelt P, Amaducci L, Sorbi S, Piacentini S, Stewart GD, Hobbs WJ, Conneally PM, Gusella JF (1987) The genetic defect causing familial Alzheimer's disease maps on chromosome 21. Science 235:885–890
Thoday JM (1967) New insights into continuous variation. In: Crow JF, Neel JV (eds) Proceedings of the Third International Congress of Human Genetics, Johns Hopkins Press, Baltimore, pp 339–350
Tsuang MT, Bucher KD, Fleming JA (1982) Testing the monogenetic theory of schizophrenia: an application of segregation analysis to blind family data. Br J Psychiatry 140:595–599
Vogel F, Motulsky AG (1986) Human genetics: problems and approaches. Springer, New York Berlin Heidelberg
Vogel HP (1979) Fertility and sibship size in a psychiatric patient population: a national comparison with census data. Acta Psychiatr Scand 60:483–503
Vogler GP, Gottesman II, McGue MK, Rao DC (1990) Mixed model segregation analysis in the Lindelius Swedish pedigrees. Behav Genetic 20:461–472
Wright S (1934) An analysis of variability in number of digits in an inbred strain of guinea pigs. Genetics 19:506–536