Am. J. Hum. Genet. 57:717–726, 1995

Latent-Class Analysis of Recurrence Risks for Complex Phenotypes with Selection and Measurement Error: A Twin and Family History Study of Autism

A. Pickles, P. Bolton, H. Macdonald, A. Bailey, A. Le Couteur, C.-H. Sim, and M. Rutter

MRC Child Psychiatry Unit, Institute of Psychiatry, London

Summary

The use of the family history method to examine the pattern of recurrence risks for complex disorders such as autism is not straightforward. Problems such as uncertain phenotypic definition, unreliable measurement with increased error rates for more distant relatives, and selection due to reduced fertility all complicate the estimation of risk ratios. Using data from a recent family history study of autism, and a similar study of twins, this paper shows how a latent-class approach can be used to tackle these problems. New findings are presented supporting a multiple-locus model of inheritance, with three loci giving the best fit.

Introduction

Autism is a rare disorder (prevalence 2–4/10,000) with an onset in early childhood, that is characterized by deviant communication, impaired reciprocal social interaction, and restrictive and repetitive behaviors. It is often, but by no means always, associated with low IQ. It is very rare for individuals with autism to marry and have children. Nonetheless, both twin and family studies have shown that genetic factors play a major etiological role. The 4% rate of autism among siblings of an autistic proband represents a risk ratio of ~100, when compared to the rate of autism in the general population (Smalley et al. 1988). However, in addition to those few relatives who show full autism, a much larger number show only some symptoms or show them only in mild form. Folstein and Rutter (1977*a*, 1977*b*) found that concordance rates rose from 36% to 82% for MZ twins and from 0% to 10% for DZ twins when the phenotype was extended from autism to include some form of cognitive impairment. A follow-up of this sample (Le Couteur et al. 1989) and a second twin sample (Bailey et al. 1995) suggested that the phenotype also included impairments in reciprocal social interaction. These data together strongly suggest a highly genetic disorder, but one in which there is substantial variability in phenotypic expression.

Using interview and observational measures intended for diagnosis of autism, Spiker et al. (1994) identified 7%–14% (8/117 and 20/142, depending on inclusion criteria) of the children in a study of families multiplex for autism, as meeting diagnostic criteria in some but not all three areas of behavioral impairment required for the diagnosis (social, communication, and restricted and repetitive behaviors). In addition, among the 22% of children who were clearly not autistic, the distribution of scores suggested several had significant sub-threshold impairment. The family study of Bolton et al. (1994) considered, in addition to autism and pervasive developmental disorder (PDD), two further categories of affected status; *narrow* (impairment in two areas) and *broad* (impairment in one). An odds-ratio of ~8 for sibs of autistic probands, as compared to sibs of Down syndrome probands was maintained even as the definition of affected status expanded to include, in addition to the 4% with autism or PDD, the further 7% with narrow and the still further 9% with broad (making an overall total of 20% affected first-degree relatives of autistic probands). Nonetheless, the cut-offs defining these categories remained somewhat arbitrary. There was no obvious bimodality in the distribution of total "symptom" scores, nor did there appear to be a clear threshold at which the autism/Down syndrome odds-ratio began to fall steeply.

For an "all-or-nothing" trait, the rate of decline in risk ratios with decreasing genetic relatedness can be informative about the number of loci involved (Hodge 1981; Risch 1990). However, the presence of measurement errors can have a complex effect on estimated risk ratios (Weiss et al. 1982; Majumder et al. 1983), with a general tendency for even quite low false-positive rates to substantially reduce uncorrected observed risk ratios. Extending the autism phenotype to include these lesser variants may reduce the rate of false negatives (failing to identify true affected cases), but it inevitably increases the rate of false positives or "phenocopies." A latent-class approach (Lazarsfeld and Henry 1968) that simultaneously defined the phenotype and estimated the risk

Received June 16, 1994; accepted for publication June 9, 1995.

Address for correspondence and reprints: Dr. Andrew Pickles, MRC Child Psychiatry Unit, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, United Kingdom. E-mail: spjwarp@pluto.ulcc.ac.uk

0002-9297/95/5703-0022$02.00

ratios would help to resolve this problem. The approach allows the affected status of relatives to be treated by probabilistic assignment to latent phenotypic classes. The data analyzed in such an analysis are not the single consensus or "best-estimate" diagnosis normally used, but a set of diagnostic ratings or, as here, items covering various specific areas of impairment. We present a model that includes parameters that estimate the rate of false positives and false negatives generated by each item, and these can be allowed to vary by item, informant, characteristics of the relative, or other relevant quantity. The model also includes parameters that estimate the rate or prevalence of each latent class, allowing a comparison to be made of the fit of models that constrain the rates of the latent phenotypic classes to be consistent with postulated single- and multiple-locus mechanisms, in the same way that, for example, Farrall and Holder (1992) have done for a simple phenotype that was assumed to be measured without error.

The autism family study data were collected using the family history method. Though it is often the only practical method of data collection beyond first-degree relatives, there are inevitable concerns about the quality of data. Andreasen et al. (1986) suggested relatively high specificity but quite low sensitivity for the family history approach. Although ratings of case vignettes from the autism family and twin study showed high reliability (Bolton et al. 1994), it would not be surprising if sensitivity declined with declining level of an informant's first-hand experience of the subject, for example with only the most severe cases being reported from among the most distant relatives (Thompson et al. 1982). Consequently, non-uniformity in measurement sensitivity could lead to an artifactually higher rate of decline in the recurrence risk with degree of relatedness, biasing conclusions against simple additive models with single or few loci, in favor of models with epistatic and dominance effects. The latent-class models proposed allow for systematic variation in the measurement performance of the diagnostic instrument.

Although a latent-class model can be restricted to two classes, representing the affected and the unaffected, it may also provide evidence for more than a single affected class. The autism data described above could be consistent with affected classes that vary in severity or that vary in type of expression. Such variation may arise through differences in environment or genes. While a high MZ concordance for the broad phenotype may point to the former, variation in the relative proportions of the affected classes with degree of relative—say, with the most distant relatives being only mildly affected—might point to the latter. Prevalence rates in relatives may also vary as a result of the influence of ascertainment method and reduced fertility (Risch 1983). Approximate adjustment for reduced fertility has been found to remove differences in morbid risk for sibs and parents of schizophrenic probands (Essen-Moller 1955). Clearly, any analysis of the pattern of risk ratios for autism should correct for such selection effects.

The Model

Basic Latent-Class Measurement Model

Consider a sample of I relatives subscripted $i = 1, \ldots, I$, with scores Y_{ij} on J symptomatic items, each having $K(j)$ possible categorical values from 1 to $K(j)$. These categorical scores are assumed to be ordered. The model accounts for the pattern of item scores by postulating M mutually exclusive classes of relative, subscripted m, $m = 1, \ldots, M$. The probability that a relative belongs to a particular latent class is given by π_{im}, where $\Sigma_m \pi_{im} = 1$ for all i, and all π_{im} are 0 or positive.

We consider first a model in which the latent classes are assumed to be ordered on a single underlying scale representing the extent of expression of the latent phenotype. Each class thus occupies a position δ_m on this scale. Without loss of generality, δ_1, the position of the first class, can be set to 0, with $0 < \delta_2 < \ldots < \delta_M$. The local independence assumption that underlies latent class models implies that the probability that individual i is scored on the jth measure as falling in category k depends solely on their latent class membership. Conditional on this latent class membership, an individual's scores are independent from one item to another. Conditional on latent class membership Z_i,

$$Pr(Y_{ij} = k \mid Z_i = m) = \int_{\alpha_{jk-1}}^{\alpha_{jk}} dH_m(s) ,$$

where $H_m(.)$ is some suitable distribution function on $-\infty$ to $+\infty$, and the α's are threshold parameters that, with the exceptions of α_{j0} and $\alpha_{jK(j)}$ that are set to $-\infty$ and $+\infty$ respectively, are to be estimated. Choosing $H_m(.)$ to be logistic distributions with common scale but with location that varies with m gives an ordered polychotomous logistic model (McCullagh 1980) with

$$\begin{aligned} &Pr(Y_{ij} = k \mid Z_i = m) \\ &= \frac{\exp(\alpha_{jk} + \delta_m)}{1 + \exp(\alpha_{jk} + \delta_m)} - \frac{\exp(\alpha_{jk-1} + \delta_m)}{1 + \exp(\alpha_{jk-1} + \delta_m)} \qquad (1) \\ &= p_{ijkm}(\alpha, \delta) . \end{aligned}$$

The likelihood of the response vector of the ith subject is then

$$l_i = \sum_{m=1}^{M} \pi_{im} \prod_{\substack{j=1 \\ k=Y_{ij}}}^{J} p_{ijkm}(\alpha, \delta) . \qquad (2)$$

The sample log-likelihood $\Sigma_i \log l_i$ can then be maximized

to obtain estimates of the parameter vectors π and δ, and the matrix of thresholds α.

Latent-class methods enjoy the advantages of other maximum-likelihood methods, in easily adapting to circumstances with "ignorable" missing data (Little and Rubin 1987). In the presence of missing item data, the product in equation (2) can be restricted to be over only those j for which data are available. This not only allows for inclusion in the analysis of relatives with incomplete phenotypic data (e.g., the children in our study who lacked the measures appropriate to adults), but also allows the weaker assumption that the missing data are missing at random (MAR) rather than missing completely at random.

Risk Ratios for Prevalences of Affected Latent Classes

In a sample of R types of relative, $r = 1, \ldots, R$, of affected probands, phenotypic prevalences would be expected to vary with the rates of the underlying genotypes. As Risch (1990) and others have shown, different genetic models imply different patterns of values for the risk ratios λ_r, and these impose constraints on the pattern of estimated prevalences. For V loci, subscripted $v = 1, \ldots, V$, with multiplicative (epistatic) effects

$$\lambda_r = 2^{-(r-1)} \prod_{v=1}^{V} (\lambda_{v1} + 2^{r-1} - 1), \qquad (3)$$

and the rate of decline in risk ratios with r increases with the number of loci V. In the absence of dominance, λ_1, the recurrence risk for parents or offspring, will also apply to sibs, and that for MZ twins can be obtained using the relationship $\lambda_{vMZ} = 2\lambda_{v1} - 1$.

It should be noted that in the absence of each locus having a distinctive phenotypic expression, the values of the locus-specific risk ratios, λ_{v1}, are not identified from phenotypic rates among relatives, only the combined effects of all loci. If variation in phenotypic expression is environmental or random in origin, such that within the set of affected latent classes the probability of membership of a particular latent class is independent of genotype, then the same risk ratios λ_r will apply to each affected latent class. Thus, if latent class M represents those without the phenotype and π_m ($m = 1, \ldots, M - 1$) is the population prevalence of each affected latent class, then the prior recurrence risk for latent class membership of individual i of relative type $r(i)$ required for the likelihood of equation (2), will be given by $\pi_{im} = \lambda_{r(i)}\pi_m$ for $m = 1, \ldots, M - 1$, for the affected latent classes, and by $\pi_{iM} = 1 - \Sigma_{m=1}^{M-1} \lambda_{r(i)}\pi_m$, for the unaffected class. Where required, risk ratios could be made a function of proband characteristics, such as sex.

Latent Classes for "Phenocopies"

The above argument assumes that membership of all but the Mth latent class requires an affected genotype. The measurement equations allow for the possibility that some individuals may appear positive on some items by chance but assumes that such errors are conditionally independent from item to item. Since data on all items are usually collected from the same informant, and also since there are likely to be other disorders (other than the focal disorder under investigation) that may be responsible for associations among at least some subsets of items, then correlated errors may occur. This suggests that although the latent-class model, so far defined, may reduce the problem of misclassification, the affected latent classes (1 to M − 1) may nonetheless contain a proportion of "phenocopies." The model can be expanded to account for phenocopies, by the inclusion of additional latent classes located on the underlying dimension of phenotypic expression close to or among the true affected classes. However, we would not expect these new latent classes to follow the pattern of recurrence risks of those for "true" cases. This difference allows the two types of latent class to be identified.

In the application that follows, we consider a simplified version of the above approach, in which allowance is made for the possibility that in the general population the mildest affected class, class M − 1, contains a proportion ω of cases that have an origin independent of their genetic relationship to the proband with the focal disorder. The prevalence by type of relative of this latent class is then given by a simple mixture of true cases and phenocopies (or "sporadics"), such that $\pi_{iM-1} = \lambda_{r(i)}\pi_{M-1}(1 - \omega) + \omega\pi_{M-1}$. Having a control group of subjects unrelated to an affected proband may be particularly helpful for the estimation of such models. The possibility that phenocopies might not be independent of the genetic system of the focal disorder, for example, representing somatic transformation of heterozygote genotypes (Majumder et al. 1983), has not been pursued here.

Extended Measurement Models for Variable Quality of Measures

In the measurement model of equation (1) all the items are equally sensitive to the latent phenotype. Systematic variation in item sensitivity can be accommodated by the inclusion of additional parameters θ_j that multiply the δ_m parameters. Model identification requires at least some of the θ parameters to be fixed, for example, being set equal to 1.

Family history data are usually obtained using informants, typically either the proband or another individual centrally located within the pedigree. We had two principal concerns: first, that knowledge of and thus sensitivity to the presence of disorder might be less for more distant relatives, and, second, that informants for case families might be more "sensitized" than controls to the possible presence of relevant symptoms among their relatives. Both concerns imply the need to consider variation over

a set of θ_{ij} parameters, that now allow for variation in item sensitivity by type of subject, as well as by type of measure. Another important example, but for disorders with a later age at onset than autism, would be variation in item sensitivity with subject's age, to account for differences in period-at-risk.

Ascertainment Bias due to Selection for Parenthood

Some types of relative enter into family history studies by virtue of parenthood. Such types of relative may be unlikely to include those with severe expression of a disorder. We can distinguish "observed" latent-class prevalences following selection, say, π^*_{im}, from those unselected, and estimate selection probability parameters σ_{im}, such that

$$\pi^*_{im} = \frac{\pi_{im}\sigma_{im}}{\sum_u \pi_{iu}\sigma_{iu}} .$$

These selected values, π^*_{im}, replace those for π_{im} in equation (2).

Samples

Details concerning the sampling of the co-twins included in this study can be found in Bailey et al. (1995). In brief, data on twins were obtained by a follow-up of the sample ascertained by Folstein and Rutter (1977*a*, 1977*b*), and a new sample ascertained in the same way. Both were epidemiological samples of same-sex pairs containing at least one autistic individual and were obtained by an exhaustive search throughout Britain. Of the total of 46 pairs and 1 triplet identified with a member meeting current ICD10 criteria, three pairs with a recognized medical condition were excluded.

Details about the process of sample selection for the family history study can be found in Bolton et al. (1994). In brief, 110 autistic probands were selected from a pool of clinic patients at the Maudsley Hospital, after preliminary exclusions to maintain ethnic homogeneity and to avoid concomitant medical disorder (Rutter et al. 1994) and probands with severe mental handicap, among whom accurate diagnosis is most difficult. Nine families refused to participate or could not be traced, and, on testing, a further two probands with fragile X were excluded. Data were thus available on 99 families.

Families with a Down syndrome proband were chosen as a comparison group because the disorder conferred no genetic risk for cognitive or social dysfunction in relatives but was likely to be similar to autism in terms of the possible impact on family functioning and the psychosocial development of relatives. The Down syndrome probands were selected from a pool of 199 (Gath and Gumley 1986), to broadly match the autistic probands with respect to age, sex, social class, birth order, and maternal age of the autism probands. Twelve families could not be traced or refused to participate. The final comparison sample consisted of 36 families.

Table 1

Distribution of Relatives by Type and Group

Relative	Autism	Down Syndrome
MZ co-twin	19	...
DZ co-twin	17	...
Parent	198	72
Sib	138	64
Grandparent	388	143
Uncle/aunt	365	174
Nephew/niece	8	16
Cousin	583	300
Total	1716	769

Measures

The diagnosis of probands was confirmed using the Autism Diagnostic Interview (Le Couteur et al. 1989), an investigator-based interview with the principal caretaker and by the Autism Diagnostic Observation Schedule (Lord et al. 1989), a structured observational schedule. Both instruments provide algorithms for determining diagnostic status and all autistic probands were confirmed as meeting ICD-10 (WHO 1992) and DSM-III-R (American Psychiatric Association 1987) criteria.

Data on relatives were obtained by interview with the proband's parents by using an investigator-based family history interview (Bolton et al. 1994). This provided data on a set of items that covered a wide range of possible impairments in communication, social relations, and stereotyped/repetitive behaviors during childhood, and for older relatives also in adulthood. In general, each item was coded as either absent, probably present, or definitely present. Fifteen of these items were chosen as indicators of the phenotype, nine referring to childhood and six referring to adulthood, and analyzed as three category ordinal variables.

Analysis and Results

Descriptive Data

We selected for analysis all sibs, parents, grandparents, aunts, uncles, nephews, nieces, and cousins aged 8 years or more at the time of data collection. Table 1 gives the number of such relatives by type and proband group. Of the 2,485 relatives (36 co-twins and 2,449 family-study relatives), we excluded 112 relatives (111 family study and 1 co-twin; 4.5% from the autistic samples and 4.7% from the Down syndrome comparison sample) because data were missing on more than two childhood measures, those measures being considered

Table 2

Distribution of Item Scores by Group

	Autism, with Score Frequency				Down Syndrome, with Score Frequency			
	0	1	2	Missing	0	1	2	Missing
Language delay	1590	15	30	5	729	1	2	1
Articulation disorder	1612	7	15	6	727	4	0	2
Reading retardation	1583	16	35	6	715	7	8	3
Childhood conversation	1606	14	19	1	731	1	1	0
Child social dysfunction	1611	8	21	0	732	1	0	0
Childhood friendships	1593	18	27	2	728	3	0	2
Affect	1613	12	14	1	730	2	0	1
Social play	1614	9	16	1	732	1	0	0
Childhood obsessive-compulsive disorder	1622	7	11	0	732	1	0	0
Adult conversation	1441	23	16	160	659	3	1	70
Adult social dysfunction	1451	10	22	157	662	1	0	70
Adult friendships	1411	30	33	166	657	4	1	71
Odd personality	1447	20	12	161	661	1	1	70
Circumscribed interests	1444	24	13	159	659	4	0	70
Adult obsessive-compulsive disorder	1471	8	4	157	663	0	0	70

relatively more important than those in adulthood. Otherwise, all subjects were included, making use of the ability of the latent-class approach to deal with MAR data in the remaining items. The distribution of scores by item and group are shown in table 2. Almost all missing data were due to the absence of measures on adult items for relatives who were still children at the time of interview. Since autism becomes evident within the first few years of life, correction for relatives not having completed the period of risk was thought unnecessary.

Some preliminary tabulations were undertaken using the definitions of the phenotype defined by Bolton et al. (1994). Being directly calculated from the manifest variables, these tabulations were more stringent in their "complete" data requirement than were the analyses using the latent-class method, and are thus based on a somewhat smaller sample. These tabulations were used to identify simple features in the pattern of prevalence that a latent-variable model would also need to account for. Table 3 presents the rates of autism/PDD, narrow, and broad, for parents, sibs, and second- and third-degree relatives, broken down by sex and proband type. There are a number of points of note. The only relatives with autism/PDD were male sibs of autistic probands and MZ co-twins. When this phenotypic measure alone was used, very little could be said about the pattern of risk ratios. The absence of cases within the Down syndrome control group, however, is of note, since the sibs, in particular, were in an environment (being cared for by parents and growing up with a handicapped child) that might be expected to generate an unusually high rate of false-positive "cases."

Moving on, to consider the phenotype defined by autism/PDD or narrow, consistently higher rates are seen in all types of relative of autistic probands than in controls. Only a very few relatives of Down syndrome probands, primarily male sibs, were classified as affected. On moving to the least stringent phenotypic definition (autism/PDD, narrow, or broad), rates in male sibs of autistic probands reach 30%, but a consistent small percentage of relatives of Down syndrome probands were also classified positively, suggesting that the measurement of this mildest level of expression may be prone to more error.

Two further points should be noted. First, the lower rates in females than in males is seen for all levels of expression. Second, the systematically higher rates in sibs than in parents, shown for both autism/PDD and the narrow phenotype, a potential consequence of reduced fertility, did not appear to extend into the broad phenotype. The rates of this mildest phenotype were similar or higher in parents as compared with sibs (15% − 2% = 13% for fathers, versus 30% − 19% = 11% for brothers, and corresponding rates of 7% for mothers and 3% for sisters).

The analyses that follow, although explicitly allowing for dependence between relatives and their respective probands, assumed that observations were independent between relatives. Of course, this is unlikely to be strictly true, since information on well over 2,000 relatives was obtained from only 135 families and informants. However, the phenotype appeared to be relatively evenly distributed across families, with the frequency distribution of the number of affected relatives in the 99 families of autistic probands being 40 families (none affected), 35

Table 3

Observed Rates for Varying Phenotypic Definitions

	Cumulative Percentage of Relatives Affected in Families of							
	Probands with Autism				Probands with Down Syndrome			
	A[a]	N[b]	B[c]	No. of Relatives	A[a]	N[b]	B[c]	No. of Relatives
MZ co-twin:								
Male	77%	85%	85%	13				
Female	75%	75%	100%	4				
DZ co-twin:								
Male	0%	0%	11%	9				
Female	0%	0%	17%	6				
Parents:								
Male	0%	2%	15%	97	0%	0%	8%	36
Female	0%	0%	7%	98	0%	0%	0%	33
Sibs:								
Male	5%	19%	30%	79	0%	3%	3%	30
Female	0%	4%	7%	58	0%	0%	6%	34
Second degree:								
Male	0%	1%	6%	358	0%	0%	3%	169
Female	0%	1%	4%	373	0%	0%	3%	152
Third degree:								
Male	0%	1%	6%	262	0%	1%	3%	144
Female	0%	0%	3%	265	0%	0%	1%	133

[a] A = autism or PDD.
[b] N = autism, PDD, or narrow.
[c] B = autism, PDD, narrow, or broad.

families (one affected), 11 families (two affected), 7 families (three affected), 1 family (four affected), 4 family (five affected), 0 families (six affected) and 1 family (seven affected). The corresponding frequencies for the families of Down syndrome probands were 23, 11, 0, 1, 0, 0, 1, and 0 families. These distributions did have slightly heavier tails than the standard binomial, but this was partly explained by the presence of some large families. When account was taken of family size, fitting a logistic regression model with family as a random effect suggested that lack of independence was not a serious problem, increasing the standard error (and thus confidence interval) for the difference in rates of the broad phenotype between the autism and Down syndrome relatives by $<10\%$. Since even for this parameter that was likely to be particularly sensitive to non-independence (being estimated from between family information only), this difference was small, the observations were assumed independent throughout the model fitting of the next section.

Model-Fitting Results

The latent class model was fitted by maximum likelihood, using a program written in Fortran-90 by one of the authors (A.P.). Optimization was performed using subroutine E04JBF from the NAG Numerical Library (NAG 1991). In preliminary analyses, two-item thresholds were estimated for each of the three-category symptom measures—making 30 parameters for the 15 measurement equations. These analyses were principally concerned with identifying the number of latent phenotypic classes. Little structure was imposed on the prevalence of the latent classes, these being allowed to vary by the sex of the relative and by proband group, and, in the case of the autism proband group, by type of relative. Within the affected classes, however, the relative proportions in the different classes of expression were assumed constant across group and across type and sex of relative. These preliminary analyses suggested that four latent classes (three affected and one unaffected), roughly corresponding to the classes shown in table 2, were sufficient to provide a full characterization of the phenotypic variation in these data.

Imposing the risk ratio relationship of equation (2) restricted the prevalence of disorder by type of relative, to follow a particular pattern that was dependent on the number of loci involved and the population prevalence of the disorder. In the case of autism, the latter clearly varies by sex, but we assumed that the risk ratios were the same for both sexes, simply applied to different population prevalences to give their respective recurrence risks. We then examined a number of model types that

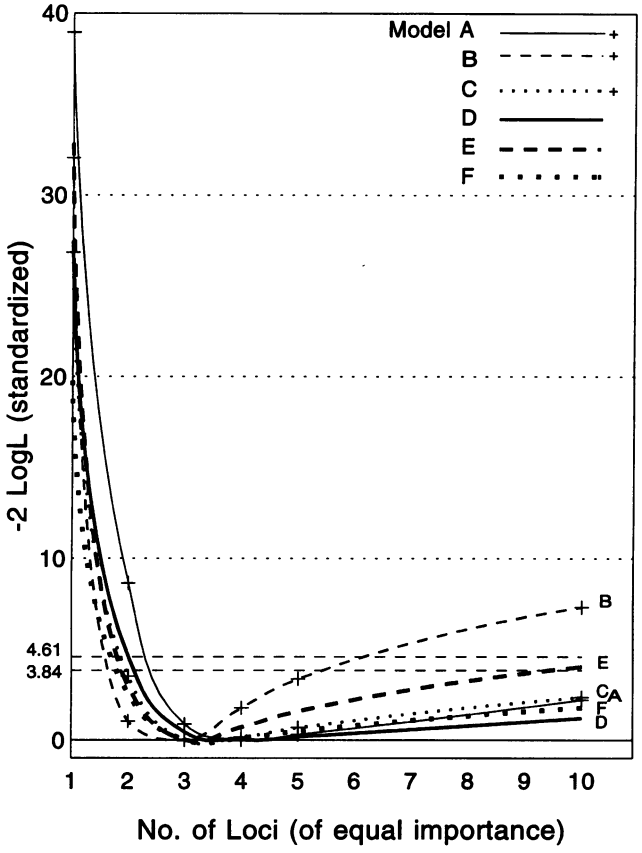

Figure 1 Variation of fit of latent-phenotype models with multiplicative loci (whole sample data n = 2373).

varied by the complications assumed to be present in our family history data, to determine the robustness of the conclusions reached relating to the likely structure of the genetic system.

For each type of model, a sequence of fits was obtained for models with 1, 2, 3, 4, 5, and 10 loci. The relative fit of models within each sequence, as estimated by minus twice the sample log-likelihood, is shown as a curve in figure 1. Different model types actually yielded very different absolute levels of fit, but the curves from different model types have been standardized such that the minimum of each curve is 0. Within each curve, therefore, the scale of the y-axis is the difference in −2logL (natural logarithm) between the fitted model and that with the optimal number of loci, corresponding to the scale of a likelihood ratio (LR) test statistic. We have also plotted horizontal lines at 4.61 and 3.84. Sections of curves below these horizontal lines fall, respectively, within the "lod minus 1" support interval (Conneally et al. 1985) of geneticists and the similar "95% likelihood interval" of statisticians.

Models of type A assumed three affected latent classes, a single unaffected class, no parental selection effects, and no variation in sensitivity (constant σ and θ matrices). Very substantial improvement in fit was found on moving from a one-locus model to a two-locus model; a further improvement was obtained on adding a third locus; and a very small improvement on adding a fourth. Additional loci yielded no further improvement; indeed, the fit became slowly but progressively worse. These results strongly suggested that a simple single-locus model was not appropriate for autism and that a two-locus model appeared unlikely. The model with four loci performed best, but the evidence against models either with three loci or involving a substantially greater number of loci was slight.

Model type B included a single additional parameter in the matrix σ to allow for parental selection effects in parents and grandparents among members of latent class 1 and 2 (those roughly corresponding to autism/PDD and narrow). Allowing for parental selection effects improved the fit markedly, giving a LR statistic of ~50 between type A and type B models each with the same number of loci. The curve for this model suggested a more clear-cut optimum at three loci, with models involving more than perhaps a further two or three loci being rejected.

Models of type C and D investigated the impact of variation in the sensitivity parameters θ. In models of type C, relatives and measures were classified into two groups, according to whether the informant might be expected to have good knowledge of the symptom because the informant lived with that relative during the relevant period. For example, parents' reports on adult measures for the proband's grandparents were in the first group, but the reports on the grandparents' childhood measures were in the second group. All reports for cousins, the most distant and only third-degree relatives of this study, were in the second group. A single parameter added to allow for variation in item sensitivity between these two groups resulted in a further significant improvement (LR ~25) with a lower estimated item sensitivity for the more distant relatives. Although the optimal number of loci remained three, the range of the likely number of loci increased substantially over that from model B. In model D a further parameter was added that extended possible reduced item sensitivity to all measures obtained on relatives in the Down syndrome group. This resulted in little improvement in fit (LR ~1). The minimum of the type D curve moved to the four-locus model, and the curve became still flatter for larger numbers of loci.

Models of type E explored the possible impact of phenocopies. In addition to the significant parameters for selection and variable sensitivity included in type C models, the type E models included a single additional parameter that divided the population prevalence of the mildest affected latent class into two, a proportion (1 − ω) to which the risk ratios were applied and a proportion ω of possible phenocopies to which risk ratios did not apply. This single parameter provided a significant

improvement in fit over model C (LR ~10). The estimated ω parameter suggested that, in the general population, the great majority of those in the mildly affected class would be false positives. Nonetheless, such false positives would form only a relatively small proportion of those identified among close relatives of autistic probands with the result that, as shown in figure 1, the optimum number of loci remained at three. Models involving large numbers of loci were excluded under this model.

Finally, since differences in rates by sex are a prominent feature of autism, a type C model was refitted to data that excluded female relatives. Shown as model F, the minimum of the curve remained in the three-locus position, but, as expected, the reduced sample gave rise to a slightly larger confidence interval.

For each relative, latent-class models can be used to derive a posterior probability of belonging to each latent class, on the basis of estimated class prevalences for that type of relative, but updated in an empirical Bayes's fashion by that relative's observed symptom profile. Figure 2 compares the sample rates of being in any of the three affected classes as "observed" using the standard cut-off as applied in Bolton et al. (1994), to the Bayes's estimates from the optimal three-locus models of type C and E. Both latent-class models, but especially model E, which formally removed phenocopies, estimated the rate in the Down syndrome sample as lower than the observed rate. By contrast, the rates in the relatives of autistic probands, particularly as given by model C, were sometimes higher than those observed. Such a sharpening of contrasts is a typical and expected consequence of a latent-class approach (see, for example, Zoccolillo et al. [1992] for further discussion). Although not shown, the Bayes's estimates for rates among MZ cotwins were 100%, suggesting that the slight incomplete penetrance shown in the raw data may have been due to error in the measurement of the broader phenotype.

The appropriateness of the overall model structure was also examined. As other authors have noted (e.g., Aitkin et al. 1981) the calculation of the expected frequency of all possible combinations of item scores is, for any reasonable number of items, computationally far more time consuming than is fitting the model and results in the need to compare sparse observed and expected tables. However, the local independence assumption allows easy calculation of the distribution of expected total scores over items. Three-locus models of type C and E gave plots of observed and expected cumulative distributions of item total scores (15 items giving a maximum score of 30) that fell almost exactly on the 45 degree line, with a maximum difference of 0.4%. There was thus no evidence of model misspecification.

A Mixed or Compound Phenotype Model

The previous models focused on variable phenotypic expression in terms of varying severity along a single dimension. An alternative approach is to consider the different phenotypic classes as made up of combinations of component elements. As a rather naive example, in the case of autism there could be components for impairment in cognitive (C), social (S), and repetitive (R) behavior. There would then be eight latent phenotypes representing all possible combinations from all three being present (CSR, notionally corresponding to autism), through three pairwise combinations (CS, CR, and SR, that when grouped together would correspond to those previously described as narrow), three isolated single components (C, S, and R, that together would correspond to broad), to the final single unaffected class. For such a model the δ_m parameters of equation (1) must also be subscripted *j*, to allow each latent class to be indicated by a relevant subset of the indicators only. We refer to this second model structure as a mixed or compound phenotype model.

Under independence of the component phenotypes, the prevalence of the compound latent classes would be

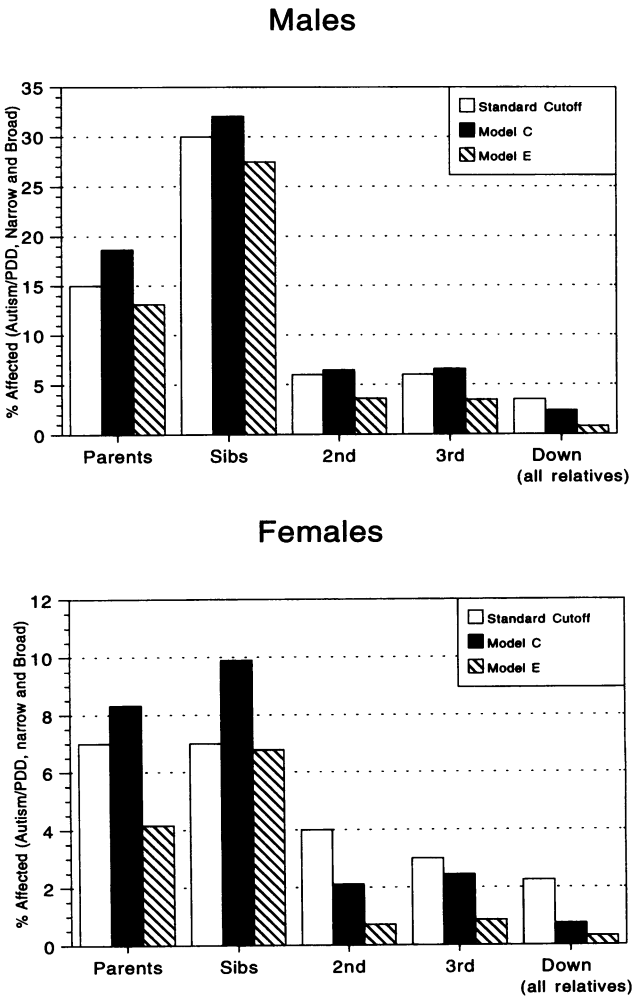

Figure 2 Posterior estimates of prevalences of all affected by relative type, sex, and group.

expected to be equal to the product of the marginal prevalences of each component. This is equivalent to considering the compound classes as co-morbid conditions arising out of chance association of component disorders. For each component disorder considered individually, a relationship among risk ratios of the form of equation (3) might apply. Thus for each component, risk ratios may be estimated that represent the effects of one or more component-specific loci. However, when considering the compound classes, we have a circumstance analogous to Risch's equation (10) (Risch 1990, p. 224), where the locus (loci) associated with each component contributes a "risk ratio factor," and the product of the corresponding risk ratio factors gives the expected risk ratio for the compound class. Since latent classes that represent compound phenotypes would be linked to more loci than would latent classes representing a single component, the decline of their risk ratios with degree of relative would be correspondingly steeper. Thus, in our example, the recurrence risks for autism/PDD, requiring all three components, would fall rapidly with degree of relative, whereas recurrence risks for the single-component broad phenotypes would fall much less rapidly. More distant relatives would therefore be relatively more likely to be identified with the single-component disorders than the (more severe) co-morbid disorders. Unlike the case of the variable-severity model, in this mixed phenotype model, having specific phenotypic associations with particular loci (or subsets of loci) would mean that the locus-specific risk ratio parameters, λ_{v1}, (or subsets of them) were identifiable and could be separately estimated.

We made only the most preliminary investigation of such a mixed phenotype model. For each of the three components a population prevalence was estimated, and the pattern of recurrence risks was restricted to that arising from estimating three risk ratio parameters and assuming a single independent locus for each component. This model did not fit well, overpredicting the number of relatives positive on just a single item and underpredicting the number with very high scores. Essentially, this model could not account for the fact that those affected in multiple areas were also more severely affected within each area than those affected in just one of the corresponding areas.

Discussion

This paper has illustrated how a latent-class approach can provide a flexible analytical framework within which to tackle a number of problems commonly associated with phenotypic measures used in genetic studies. Varying severity, varying misclassification rates, and ascertainment selection effects can all be dealt with at the same time as estimating risk ratio models.

The results from the application of this approach to autism strongly suggested the likely involvement of multiple loci. The ruling out of single-locus models accords with the conclusions of Jorde et al. (1991) using segregation analysis, but their support for a multifactorial threshold model involving additive genetic effects would not explain the divergent concordance rates found in MZ and DZ twins. Our analyses suggest three epistatic loci as plausible, though the bounds range from 2 to perhaps 10 loci (model E).

The exact relationship between the various forms of phenotype is less clear. The mixed or compound phenotype model fitted very much worse than the variable-severity phenotype model, but the structure of independent components considered may be too restrictive, and development of a more general model may be required. The mixed phenotype model sketched out above, and the variable-severity model that was more extensively investigated, may be thought of as the poles of a continuum, with independent genes for each component at one end and the same genes affecting all components at the other. Both are based on very simple assumptions about the relationship of the broader forms of the phenotype to full autism. However, the importance of distinguishing between these models is clear, since under the variable-severity model (types A–D), recurrence risks for relatives of probands with the broad phenotype would be expected to be the same as for those with full autism, whereas, under the mixed-phenotype model, the recurrence risks would be lower and likely to be limited to the area of impairment of the proband. Of course, under the type E variable severity model the great majority of those identified in the general population as having the broad phenotype would be phenocopies, implying no increased risk for their relatives.

The indicators of phenotypic status used in this analysis were selected because of their close conceptual correspondence to the impairments typically associated with autism, and because of their appropriateness for subjects of normal intelligence. All were found to be positively associated with the underlying putative dimension(s) of impairment. There is some evidence that relatives of autistic individuals may also be at risk of other symptomatology, notably major depression (De Long and Dwyer 1988; Piven et al. 1990, 1991). We did not include such symptomatology in the item pool analyzed, since the etiological significance of this association remains unclear. An extension to the latent class analyses presented here would provide a powerful framework in which to clarify this.

Acknowledgments

Thanks to John Gair for his assistance in programming and to Lindon Eaves, Emily Simonoff, and anonymous reviewers for valuable comment on an early draft of this paper. This work was supported by grants to Michael Rutter from the

Medical Research Council and the John D. and Catherine T. MacArthur Foundation. The study has relied on the work of many colleagues, and we would like to express particular appreciation to Eric Fombonne, Patricia Howlin, Ghazala Javaid, Heather Jordan, Mary Crowson, Olwen Morgan, Cathy Mulloy, Margaret Murphy, Jan Prior, Patti Rios, Christine Rutter, Doug Schmidt, Stephen Scott, Sharon Storoschuk, and Beth Whitlock. Most of all, we thank the families in both samples who gave us such generous cooperation at all stages of the research.

References

Aitkin M, Anderson D, Hinde J (1981) Statistical modelling of data on teaching styles. J R Stat Soc A 144:419–461

American Psychiatric Association (1987) Diagnostic and statistical manual of mental disorders (DSM-IIIR). American Psychiatric Association, Washington, DC

Andreasen NC, Rice JP, Endicott J, Reich T, Coryell W (1986) The family history approach to diagnosis: how useful is it? Arch Gen Psychiatry 43:421–429

Bailey A, Le Couteur A, Gottesman L, Bolton P, Simonoff E, Yuzda E, Rutter M (1995) Autism as a strongly genetic disorder: evidence from a British twin study. Psychol Med 25:63–77

Bolton P, Macdonald H, Pickles A, Rios P, Goode S, Crowson M, Bailey A, et al (1994) A case-control family history study of autism. J Child Psychol Psychiatry 35:877–900

Conneally PM, Edwards JH, Kidd KK, Lalouel JM, Morton NE, Ott J, White R (1985) Report of the Committee on Methods of Linkage Analysis and Reporting. Cytogenet Cell Genet 40:356–359

De Long GR, Dwyer JT (1988) Correlation of family history with specific autistic subgroups: Asperger's syndrome and bipolar affective disorder. J Autism Dev Disord 18:593–600

Essen-Moller E (1955) The calculation of morbid risk in parents of index cases, as applied to a family sample of schizophrenics. Acta Genet 5:334–342

Farrall M, Holder S (1992) Familial recurrence-pattern analysis of cleft lip with or without cleft palate. Am J Hum Genet 50:270–277

Folstein S, Rutter M (1977*a*) Genetic influences and infantile autism. Nature 285:726–728

——— (1977*b*) Infantile autism: a genetic study of 21 twin pairs. J Child Psychol Psychiatry 18:291–321

Gath A, Gumley D (1986) Behaviour problems in retarded children with special reference to Down's syndrome. Br J Psychiatry 149:161–171

Hodge SE (1981) Some epistatic two-locus models of disease. I. Relative risks and identity-by-descent distributions in affected sib pairs. Am J Hum Genet 33:381–395

Jorde LB, Hasstedt SJ, Ritvo ER, Mason-Brothers A, Freeman BJ, Pingree C, McMahon WM, et al (1991) Complex segregation analysis of autism. Am J Hum Genet 49:932–938

Lazarsfeld PF, Henry NW (1968) Latent structure analysis. Houghton Mifflin, Boston

Le Couteur A, Rutter M, Lord C, Rios P, Robertson S, Holdgrafer M, McLennan J (1989) Autism Diagnostic Interview: a standardized investigator-based instrument. J Autism Dev Disord 19:363–387

Little RJA, Rubin DB (1987) Statistical analysis with missing data. Wiley, New York

Lord C, Rutter M, Goode S, Heemsbergen J, Jordan H, Mawhood L, Schopler E (1989) Autism Diagnostic Observation Schedule: a standardized observation of communicative and social behavior. J Autism Dev Disord 19:185–212

McCullagh P (1980) Regression models for ordinal data (with discussion). J R Stat Soc B 42:109–142

Majumder PP, Chakraborty R, Weiss M (1983) Relative risks of diseases in the presence of incomplete penetrance and sporadics. Stat Med 2:13–24

NAG (1991) NAG Fortran Library Mark 15. Numerical Algorithms Group, Oxford

Piven J, Chase G, Landa R, Wzorek M, Gayle J, Cloud D, Folstein S (1991) Psychiatric disorders in the parents of autistic individuals. J Am Acad Child Adolesc Psychiatry 30:471–478

Piven J, Gayle J, Chase G, Fink B, Landa R, Wzorek M, Folstein S (1990) A family history study of neuropsychiatric disorders in the adult siblings of autistic individuals. J Am Acad Child Adolesc Psychiatry 29:177–183

Risch N (1983) Estimating morbidity risks in relatives: the effect of reduced fertility. Behav Genet 13:441–451

——— (1990) Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 46:222–228

Rutter M, Bailey A, Bolton P, Le Couteur A (1994) Autism and known medical conditions: myth and substance. J Child Psychol Psychiatry 35:311–322

Smalley SL, Asarnow RL, Spence MA (1988) Autism and genetics: a decade of research. Arch Gen Psychiatry 45:958–961

Spiker D, Lotspeich L, Kraemer HC, Hallmeyer J, McMahon W, Petersen PB, Nicholas P, et al (1994) Genetics of autism: characteristics of affected and unaffected children from 37 multiplex families. Am J Med Genet (Neuropsychiatric Genet) 54:27–35

Thompson WD, Orvascel H, Prussof BA, Kidd K (1982) An evaluation of the family history method for ascertaining psychiatric disorders. Arch Gen Psychiatry 39:53–58

Weiss KM, Chakraborty R, Majumder PP, Smouse PE (1982) Problems in the assessment of relative risk of chronic disease among biological relatives of affected individuals. J Chronic Dis 35:539–551

WHO (1992) The ICD-10 classification of mental and behavioural disorders: clinical description and diagnostic guidelines. World Health Organization, Geneva

Zoccolillo M, Pickles A, Quinton D, Rutter M (1992) The outcome of conduct disorder: implications for defining adult personality disorder Psychol Med 22:971–986