British Journal of Psychiatry (1990), **156**, 17–26

Lecture

Heterogeneity of Schizophrenia
Conceptual Models and Analytic Strategies*

MING T. TSUANG, MICHAEL J. LYONS and STEPHEN V. FARAONE

Schizophrenia is clinically heterogeneous but it is not known whether this is due to the existence of discrete subtypes. For the purpose of explication, 'indicators' of schizophrenia are divided into three levels: phenomenology, pathophysiology, and aetiology. Five heterogeneity models and a number of quantitative approaches are described. It is imperative to apply rigorous methods of study to the comparison of unitary models and competing heterogeneity models of schizophrenia.

The clinical heterogeneity of schizophrenia has long been recognised but the question remains as to whether this reflects heterogeneity at a more basic level. One school argues that schizophrenia is actually a group of aetiologically discrete disorders and hypothesises that research will eventually reveal the points at which subtypes may be cleaved. Supporting this view, Eliot Slater (1947) asserted: "In view of the fact that schizophrenia is a good deal more common than any single genetically determined disorder is otherwise known to be, heterogeneity is inherently probable." There are recent findings that, while requiring replication, strongly suggest genetic heterogeneity in schizophrenia. Sherrington *et al* (1988) found that schizophrenia is linked to a region of chromosome 5 in seven British and Icelandic families. Their work was motivated by a report (Bassett *et al*, 1988) of a family in which schizophrenia was associated with a translocation of a region of chromosome 5. Kennedy *et al* (1988) studied a large Swedish family and found no relationship between chromosome 5 and the occurrence of schizophrenia. As Lander (1988) observed, taken together, such findings are most consistent with the view that schizophrenia is heterogeneous with regard to aetiology.

The most influential and plausible alternative view to heterogeneity asserts that the observed heterogeneity of schizophrenia is due to a multifactorial polygenic (MFP) process whereby a single pool of aetiological factors produces a spectrum of disorders including schizophrenia. McGuffin *et al* (1987) recently reviewed the evidence for a discrete-subtype versus a unitary MFP view of schizophrenia and concluded that previous evidence suggestive of discrete subtypes is more compatible with a single multifactorial continuum of severity along which individuals may be ranked. However, there are good reasons for pursuing the investigation of discrete subtypes. An MFP approach does not readily suggest any specific physiological or environmental factor to target in treatment or prevention efforts. If subtypes with distinct aetiologies and pathophysiologies do exist, their identification would greatly facilitate progress in our understanding of schizophrenia. Biochemical or structural abnormalities that occur in some schizophrenics might not be apparent in a schizophrenic sample in which only a fraction had the subtype of schizophrenia associated with the abnormality. Moreover, the identification of subtypes associated with specific pathophysiologies and aetiologies offers the promise of subtype-specific interventions that might treat and prevent the disorder.

In the absence of compelling evidence to support either a unitary MFP or a discrete-subtype view of schizophrenia, it is necessary to compare these competing models. If the true taxonomic structure of the schizophrenic disorders is multifactorial, then attempts to specify discrete subtypes will produce ambiguous and misleading results. If discrete subtypes do exist, then multifactorial analyses will not be informative. Thus, a strategy to resolve heterogeneity must include a method for evaluating the various possibilities.

The MFP model suggests that schizophrenia is due to a large number of genetic and environmental variables that act in an additive fashion to produce a vulnerability to schizophrenia. If an individual's cumulative vulnerability exceeds a certain threshold, he or she will manifest the signs and symptoms of schizophrenia. According to the MFP model, no single factor is a necessary or sufficient cause for schizophrenia.

*This paper was presented, in part, as the Maudsley Hospital Bequest Lecture at the Royal College of Psychiatrists' Winter Quarterly Meeting on 27 January 1988 in London by Dr Ming T. Tsuang.

https://doi.org/10.1192/bjp.156.1.17 Published online by Cambridge University Press

The single major locus (SML) model suggests that schizophrenia is due to a single gene of major effect. However, the SML model is not incompatible with theories that assert that genetic factors interact with environmental factors. Of course, the problem may be even more complicated, with an SML process occurring against an MFP background.

Neither MFP nor SML models have provided a uniformly good fit to the data obtained in family, twin and adoption studies (Faraone & Tsuang, 1985; Faraone *et al*, 1988). For example, SML models that include reduced penetrance of the schizophrenia gene in heterozygotes seriously underpredict the concordance rates observed in identical twins (Elston & Campbell, 1970; Matthysse & Kidd, 1976). A number of investigators have concluded that the MFP model overpredicts the observed concordance rate of identical twins (Kidd & Cavalli-Sforza, 1973; Matthysse & Kidd, 1976). What accounts for the inability of any single genetic model to accommodate all empirical findings? The existence of substantial heterogeneity within schizophrenia is a clear possibility.

Conceptual models

The application of competing models holds the promise of allowing us to critically compare alternative views in the most productive manner. Lewis *et al* (1987) recently suggested that the sophistication of new research tools has led to the neglect of plausible hypotheses. They point out the temptation of using measures as an end in themselves rather than employing them in the context of testing hypotheses. Merely demonstrating differences between schizophrenics and some comparison group or among schizophrenics on a single measure is unlikely to advance our understanding of schizophrenia (Chapman & Chapman, 1977). What is required is a framework into which findings from different measures can be fitted. There are statistical data-reduction techniques, such as cluster analysis, that can mechanically combine data from a number of different measures. Such methods have their merits, although they have not been very productive in schizophrenia research. But because such methods do not allow us to incorporate some of the information that we do know, or at least about which we have reasonable hypotheses, they are not the optimal methods to employ.

Reasonable hypotheses can be formed about the causal ordering of some of the measurable phenomena associated with schizophrenia. These associated phenomena might be called 'indicators' following the usage suggested by Meehl (1973). In this context, 'indicator' refers to a measurable characteristic of an individual that is possibly related to schizophrenia. For example, auditory hallucinations would be considered a possible indicator, as would smooth-pursuit eye movement dysfunction or a positive family history of schizophrenia.

Level I indicators

For heuristic purposes, the indicators may be divided into different levels. Level I refers to putative indicators of aetiology. Some of the candidates for inclusion in level I include obstetrical complications and other types of perinatal insult (McNeil & Kaij, 1978). Birth during the winter season could be included here (Shur, 1982; Bradbury & Miller, 1985). Evidence for a viral aetiology would be a level I indicator (Crow, 1983; Knight *et al* 1987). Evidence for genetic aetiological factors would be included at level I. There are a number of genetic factors that could be considered individually. For example, a possible distinction between familial and sporadic forms of schizophrenia has been suggested (Reveley *et al*, 1984; Lewis *et al*, 1987) and might be employed as a level I indicator. Developments in the field of molecular genetics offer the possibility that the use of restriction fragment length polymorphisms may form the basis for identifying a gene or genes responsible for some or all cases of schizophrenia. The work of Sherrington *et al* (1988) and Kennedy *et al* (1988) mentioned above exemplifies this approach.

Another possibility is that schizophrenia might be like mental retardation, with one type of schizophrenia analogous to familial retardation and due to a multifactorial polygenic process, and another type or types analogous to pathological retardation due to an identifiable insult to the nervous system or a single gene of major effect.

An example of using level I indicators to investigate a putative subtype distinction in schizophrenia was reported by Kendler *et al* (1988). The subtypes included were paranoid, hebephrenic, and undetermined according to four diagnostic systems. The subjects were 332 patients who met DSM-III criteria (American Psychiatric Association, 1980) for schizophrenia. Because subtype diagnosis is only modestly stable over time (Kendler *et al*, 1985), the diagnosis of subtype was based on symptoms early in the course of the illness. The approximately 250 subjects who had relatives available for inclusion in the study were assigned a schizophrenia subtype diagnosis in each of the four diagnostic systems. The risk for schizophrenia (and all non-affective psychotic disorders) among relatives did not differ according

https://doi.org/10.1192/bjp.156.1.17 Published online by Cambridge University Press

to the subtype diagnosis of the proband. There was a non-significant trend for the relatives of paranoid schizophrenics to have a higher risk for schizophrenia (and all non-affective psychotic disorders) than the relatives of non-paranoid probands. An important question from the perspective of level I indicators was an evaluation of the degree of correspondence between the probands' subtype diagnoses and those of their relatives, i.e. did subtypes 'breed true'? While the agreement between the probands' and relatives' subtype diagnoses was modestly positive, in no case did it attain statistical significance. In terms of the proposed conceptual framework, these findings do not provide support for a distinction among the traditional subtypes of schizophrenia using familial morbid risk as a level I indicator.

Level II indicators

Level II includes indicators of possible pathophysiology. There are numerous findings that suggest changes in the brain associated with schizophrenia. Complicating the interpretation of such findings, however, is the fact that no finding is clearly specific to schizophrenia. Differences reported in the brains of schizophrenics thus far seem to be quantitative, not qualitative, and typically the distribution of results from a schizophrenic group shows a substantial amount of overlap with distributions from normals or from other psychiatric populations (Crow *et al*, 1987). Similarly, there are no brain abnormalities that seem to be present in all patients diagnosed as schizophrenic. The heterogeneity among putative indicators of pathophysiology might prove very helpful in investigating the possibility of meaningful subtypes within schizophrenia, especially if combined with data on clinical and aetiological indicators.

There are a number of anatomical findings that suggest level II indicators, such as ventricular enlargement and sulcal widening (Weinberger *et al*, 1980) and cytoarchitectural changes in the cortex (Benes *et al*, 1986). Advances in brain imaging technology such as magnetic resonance imaging will provide us with better data on anatomical and physiological changes that may be associated with schizophrenia. Another imaging technique, brain electrical activity mapping (BEAM), has begun to provide intriguing information about possible differences in the functioning of schizophrenics' brains (Faux *et al*, 1987). Regional blood flow studies have also provided findings, such as hypofrontality (De Lisi, 1986), that may serve as useful level II indicators.

Investigations of the biochemistry and pharmacology of schizophrenia also may provide useful indicators at this level. For example, abnormalities in dopamine receptors and the effects of chemical probes such as amphetamine challenge tests may demonstrate consistent differences among groups of schizophrenics (Angrist *et al*, 1980). The well known variability among schizophrenics in their response to neuroleptic treatment may prove informative (Bowers *et al*, 1987; Brown & Herz, 1988). Abnormalities in smooth-pursuit eye movements (Lipton *et al*, 1983), continuous performance tasks (Holzman, 1987), and neuropsychological functioning (Seidman, 1983) may be indicators of pathophysiology and should be included at level II. Level II indicators may be considered 'endophenotypes' (Cloninger *et al*, 1983), an intermediate level between aetiology and manifest clinical signs and symptoms.

Level III indicators

Indicators at level III represent clinical and phenomenological factors. Traditional signs and symptoms such as hallucinations, delusions, and formal thought disorder would be included here. Recent refinements in characterising the phenomenology of schizophrenia, such as positive and negative symptoms (Andreasen *et al*, 1982) and Type I and Type II schizophrenia (Crow, 1980), would be included here. Information about onset, pre-morbid personality, course, and outcome would also be level III indicators.

An example of the use of level III indicators in attempting to validate putative subtypes of schizophrenia is contained in a report by Kendler *et al* (1984). The study included 187 patients who met the Washington University criteria (Feighner *et al*, 1972) for schizophrenia. Based on extensive chart information, patients were assigned a subtype diagnosis according to each of four different diagnostic systems. The subtypes were then compared on long-term and short-term outcomes (level III indicators). Paranoid schizophrenics, according to all four diagnostic systems, had the best long-term outcomes in psychiatric, residential, and occupational domains. Although only one comparison was statistically significant, for each of the four diagnostic systems, the paranoids had the best short-term outcome. In general, the distinction between paranoid and non-paranoid schizophrenia was supported by differences in outcome, while the distinction between hebephrenic and undifferentiated schizophrenia was not.

Assignment of indicators to levels

Indicators are assigned to levels on the basis of hypothesised causal priority. Indicators at lower

https://doi.org/10.1192/bjp.156.1.17 Published online by Cambridge University Press

numbered levels are more likely to be causes of, or to be associated with the cause of, indicators at a higher numbered level than vice versa. For example, the hypothesis that genetic loading for schizophrenia at level I is a cause of observed hypofrontality at level II is more plausible than the idea that hypofrontality is a cause of genetic loading. Similarly, hypofrontality may be more likely to be a cause of formal thought disorder than the result. Of course, the alternative is that some unknown mechanism, 'X', causes formal thought disorder to occur and the occurrence of formal thought disorder then results in decreased activity in frontal regions. Inevitably, one has to make some tentative decisions about the assignment of some indicators to levels. However, such informed decisions are necessary to generate the hypotheses required for testing competing models, and the possibility of alternative orderings should also be evaluated.

Patterns of association

What is needed is a strategy for systematically examining indicators at various levels and determining patterns of association *across* the three levels. For heuristic purposes, six approaches to subtyping with differing patterns of indicators at different levels have been developed and are contained in Fig. 1.

Figure 1(a) illustrates the ***homogeneity*** model. In this approach it is hypothesised that there are no meaningful subtypes of schizophrenia, and that there is a single necessary and sufficient cause that produces a fixed pathophysiological process that leads to the range of symptoms associated with schizophrenia. The merits of this theoretical approach have been debated since the days of Kraepelin and Bleuler. This model could be considered the 'null hypothesis' in the investigation of heterogeneity in schizophrenia.

Parts (b) to (f) of Fig. 1 illustrate the heterogeneity models. For the sake of clarity of presentation, they are presented in the context of two alternative subtypes. If there are, in fact, subtypes of schizophrenia, there is no particular reason to believe that there are only two. However, the models may be generalised to include multiple subtypes in a relatively straightforward manner. For illustrative purposes, the two-subtype case will be presented.

Figure 1(b) depicts the *complete symptomatic specificity* model. In this model, there are two subtypes of schizophrenia and they are distinct at the levels of aetiology, pathophysiology, and symptomatology. Cause I leads to pathophysiology I, which in turn produces symptoms A, B, and C. Cause II leads to pathophysiology II, which produces symptoms X, Y, and Z. There is no overlap and the two subtypes of schizophrenia are clearly distinct at each level. The simplicity of this theoretical approach is appealing and this model seems to have been at least implicit in some traditional systems of subtyping. However, as McGuffin *et al* (1987) have pointed out, evidence adduced to sustain traditional approaches to subtyping has not been compelling. If complete symptomatic, pathophysiological, and aetiological specificity between subtypes were to be the case, previous efforts probably would have met with more success.

Figure 1(c) illustrates the *common final pathway* model. There are two distinct causes, but each results in the same pathophysiological process that produces the full range of symptoms associated with schizophrenia. In this model, there is only one type of schizophrenia at the level of pathophysiology and symptomatology (although there may be differences in severity), and no single cause is necessary. As the number of causes increases and with the assumption that none are sufficient, this approach increasingly resembles the multifactorial polygenic model.

Figure 1(d) depicts the *symptomatic non-specificity* model. There are two distinct causes that lead to two distinct pathophysiologies. However, each of the pathophysiological processes may produce the full range of symptoms of schizophrenia. In this case, the search for level III indicators useful for subtype identification is fruitless. The model does suggest, however, that indicators at levels I and II could have utility for identifying subtypes. While it has been suggested that the brain has only a finite number of ways that it may react to malfunctioning, it seems unlikely that two entirely distinct malfunctions would produce identical manifestations.

Figure 1(e), which illustrates *partial symptomatic specificity*, resembles model 1(d) but is more optimistic because it suggests that some level III indicators may be informative. According to this model, some symptoms of schizophrenia are non-specific with regard to underlying pathophysiology and aetiology, while others are specific. In this model, symptoms A and B are specific to cause I and pathophysiology I, symptoms Y and Z are specific to cause II and pathophysiology II, and symptoms C and X are non-specific indicators that may be produced by pathophysiologies I and II.

Figure 1(f) depicts *overlapping pathophysiologies*. It posits two distinct causes that produce somewhat different pathophysiological processes; however, there is some overlap in the pathophysiologies. Like model 1(e), it predicts that some level III indicators will be non-specific, but it also predicts that some indicators at level II will be non-specific as well. For

https://doi.org/10.1192/bjp.156.1.17 Published online by Cambridge University Press

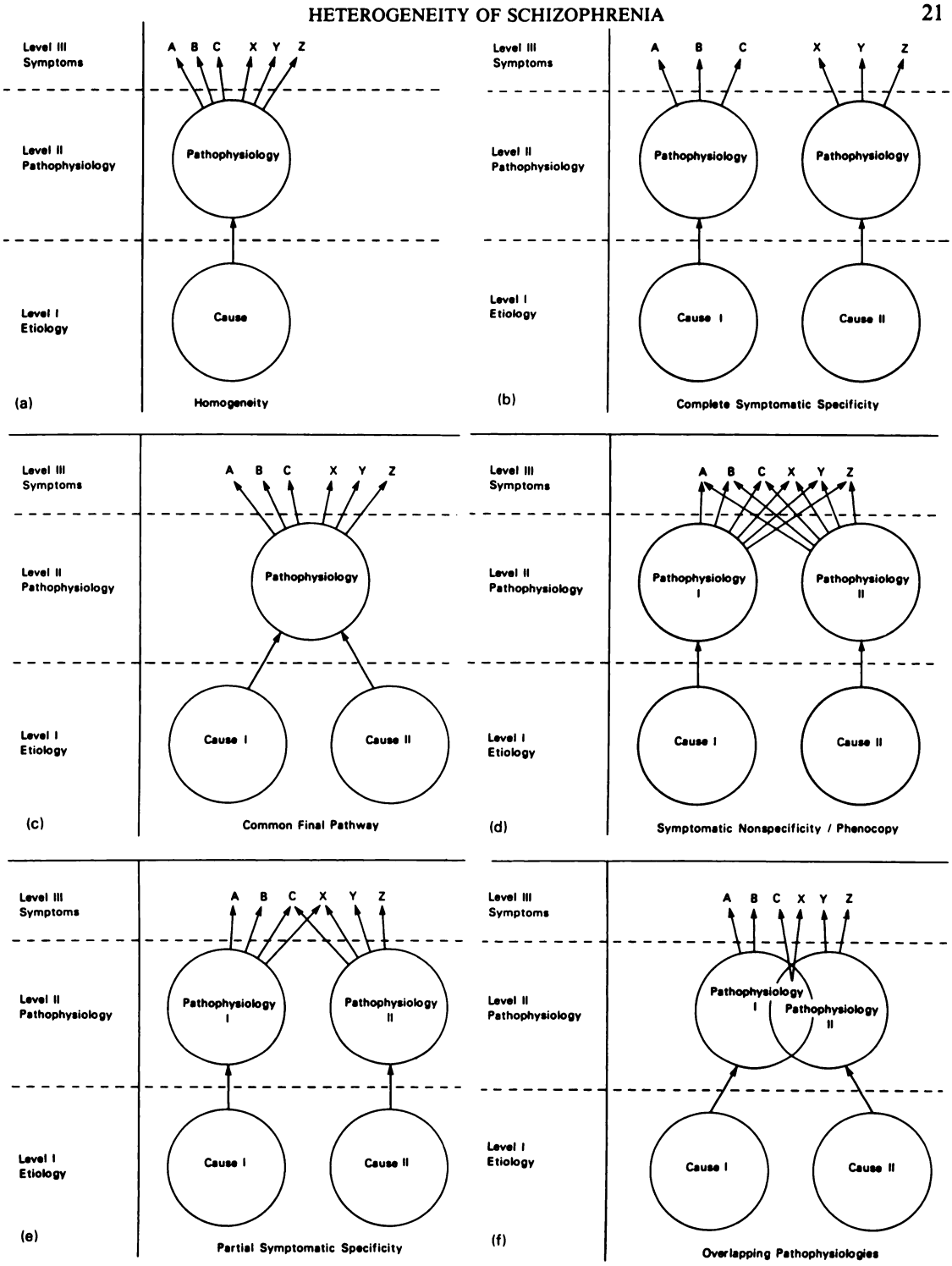


FIG. 1 Models for investigating the heterogeneity of schizophrenia.

https://doi.org/10.1192/bjp.156.1.17 Published online by Cambridge University Press

example, some phenotypic resemblance between early Huntington's disease and schizophrenia may be attributed to pathological changes in the brain common to both (Davison & Bagley, 1969). A key issue suggested by this model is the identification of specific versus non-specific level II indicators.

Example

Figure 2 illustrates an example of the partial symptomatic specificity model using mental retardation as an example. The three types of retardation included are PKU, Down's syndrome, and familial retardation. Each type has a distinct cause, resulting in differing pathophysiologies. However, at level III there are both specific and non-specific indicators. Low IQ and poor functioning are non-specific indicators that are associated with all three subtypes, while a mousy odour is uniquely associated with PKU, a thick fissured tongue is uniquely associated with Down's syndrome, and below-average intelligence among non-retarded first-degree relatives is uniquely associated with familial retardation. At level II, only PKU is associated with elevated levels of phenylalanine. At level I, only Down's syndrome is associated with chromosomal abnormalities. Some symptoms of schizophrenia may be analogous to low IQ and poor functioning in mental retardation, while some may be analogous to a mousy odour or a thick fissured tongue. At level II, the informative pathophysiological differences among schizophrenics are probably much more subtle than the elevated phenylalanine in PKU. This example of the partial symptomatic specificity model is provided for illustrative purposes; similar examples could be constructed for any of the models using other medical or psychiatric conditions.

Analytic strategies

It is possible to make predictions about the distribution of indicators under a unitary MFP model and a discrete-subtype model that may then be compared mathematically. In such analyses, three characteristics of indicators may be examined: bimodality, familial correlations, and independence of segregation.

Modality

Unlike the multifactorial model, the discrete-subtype model predicts that indicator distributions will be bimodal or, more generally, multimodal. With a large sample, bimodality should be detectable as long as the underlying distributions of subtypes are separated by two or more standard deviations and the base rate of the less prevalent subtype is greater than 20% (Harris & Smith, 1947). Tests for bimodality should examine probands and each relative class separately when, as is usually the case, probands are not randomly ascertained (Chakraborty & Hanis, 1987).

Bimodality is very difficult to detect if the base rate of the less prevalent subtype is very low (Harris & Smith, 1947). Methods for detecting rare subtypes are necessary because, if schizophrenia is like other heterogeneous familial disorders such as mental retardation, one or more rare subtypes are likely to be present. In the frequency distribution of an indicator that marks a rare subtype, members of the subtype should appear as outliers. Outlier visualisation may be facilitated by dynamic graphics programs that allow rotation of three-dimensional clouds of data points. Multivariate outliers involving more than three variables are more readily observed using the method of Fourier plots developed by Andrews (1972).

The observed distribution of patients will be a mixture consisting primarily of those who score in the normal range on the rare-subtype indicator, along with several 'contaminants' from the rare-subtype distribution. This is essentially the mixture alternative of the contamination model for outliers (Barnett & Lewis, 1984). Thus, the application of discordancy tests for outliers should help in the identification of rare subtypes. The exact test to be employed cannot be specified *a priori* because it will depend upon the distributional properties of the data. Ideally, an exact multivariate test would be applied to relevant indicators. If that is not possible, due to distributional constraints, informal multivariate methods can be applied. These can be further validated by exact univariate tests. Barnett & Lewis (1984) suggest a number of alternative methods for accomplishing this. Since outlying observations may be due to experimental error, it is necessary to find converging evidence to support a rare-subtype hypothesis. Thus, the assertion of a rare subtype will be more conclusive if the outlying individuals for one indicator distribution are also outliers on another indicator distribution, assuming that experimental error between the two distributions is uncorrelated. Further convergence can be found by demonstrating that first-degree relatives of outlier probands also have extreme values on the indicator variable.

Familial correlations

The examination of familial correlations has been useful in testing taxonomic models underlying genetic heterogeneity since Haldane (1941) demonstrated

https://doi.org/10.1192/bjp.156.1.17 Published online by Cambridge University Press

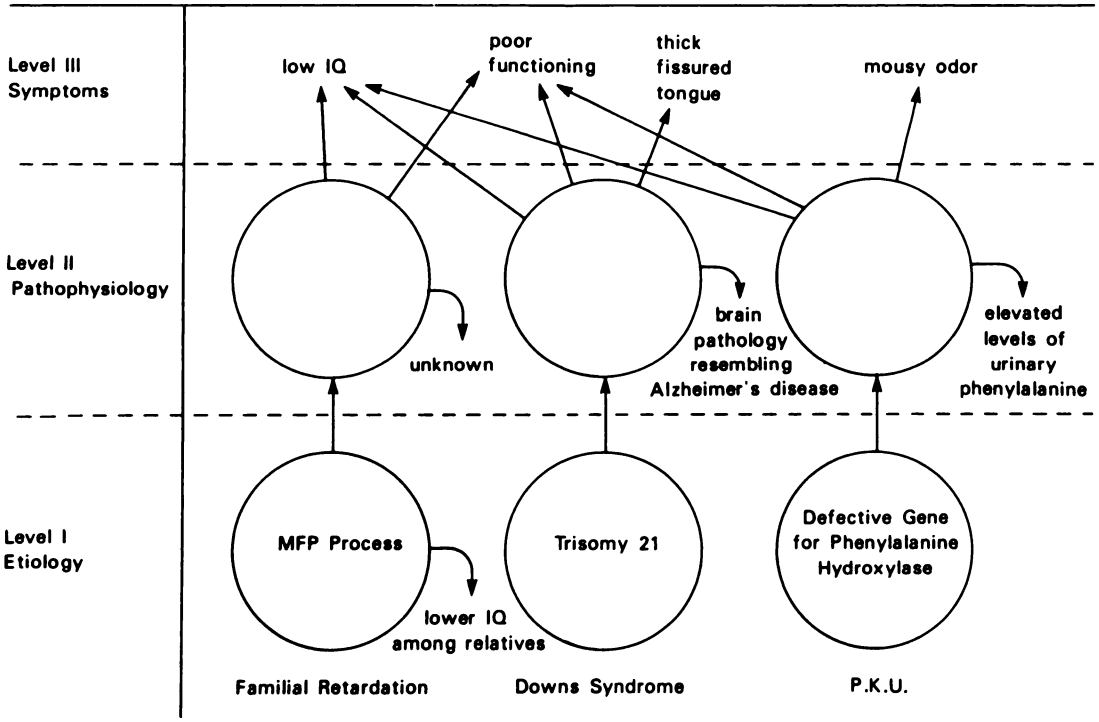


FIG. 2 Heterogeneity in mental retardation as an example of the partial symptomatic specificity model.

that such correlations can distinguish heterogeneity due to multifactorial processes from heterogeneity due to the presence of aetiology of more than one single major locus. Briefly, the correlation between ill first-degree relatives for an indicator variable should be about 0.5 if the process producing the indicator is multifactorial and the heritability is 1.0; it should approach 1.0 if aetiology of more than one single major locus is involved. However, as Harris & Smith (1947) demonstrated, the familial correlation can be less than or equal to 0.5 in the latter case when the indicator distributions of the subtypes differ by less than two standard deviations. Thus, familial correlations greater than 0.6 will provide relatively strong evidence for the discrete-subtype hypothesis; lower correlations would be more ambiguous without clear evidence of a bimodal or bitangential proband distribution. In a bitangential distribution, the second mode appears as a bulge on one side of the distribution. If at least one tangent other than the abscissa touches the distribution at two points, then the distribution is said to be bitangential (Harris & Smith, 1947). Fortunately, the analytic work of Harris & Smith indicates that correlations of 0.5 (the multifactorial expectation) should produce clear evidence for a bitangential proband distribution if the discrete-subtype hypothesis is correct. Familial correlations can be computed using the methods of Rao *et al* (1987). When multifactorial inheritance is assumed, the observed correlation can be corrected for the effects of non-random sampling according to the guidelines of Rao & Wette (1987).

Independence of segregation

The multifactorial and discrete-subtype hypotheses make qualitatively distinct predictions about segregation of subtypes within families. If a series of subtypes corresponds to the gradations of severity caused by a multifactorial process, then the risk for all subforms of this disorder should increase with the severity of the proband's illness. That is, one should *not* observe independent segregation of subtypes. In contrast, the hypothesis of discrete familial subtypes would predict independent segregation of disorders. That is, ill relatives of subtype A would be expected to have subtype A of the disorder; ill relatives of subtype B would be expected to exhibit subtype B and so forth. Similar predictions would hold for the segregation of biological markers within families.

https://doi.org/10.1192/bjp.156.1.17 Published online by Cambridge University Press

Taxonomic search procedures

Because we know so little about the putative taxonomic structure of schizophrenic subtypes, it may be helpful to employ the type of taxonomic search procedures advocated by Meehl (1973) to 'bootstrap' a latent taxonomy. Taxonomic search procedures usually make use of the axiom of conditional independence, which states that indicators of discrete subtypes are pairwise uncorrelated within a homogeneous population containing one subtype but are correlated within a heterogeneous population containing more than one subtype. That is, if there were two subtypes, I and II, then two putative indicators, A and B, would not be correlated within subtype I or within subtype II but would be correlated in a mixed group containing both subtypes. This occurs if the values of indicators A and B are associated with membership in subtype I or II. For example, if delusions are associated with a paranoid subtype and high ratings for negative symptoms associated with a non-paranoid subtype, then within a sample comprised only of paranoids they would not be correlated. In a sample made up of both paranoid and non-paranoid cases, they would be negatively correlated. Thus, the values of correlation coefficients estimated from various partitions of the sample can be used to develop subtypes.

Maxcov–hitmax

The 'maxcov–hitmax' taxonomic search procedure developed by Golden & Meehl (1979) begins with at least three quantitative variables. A group of variables is considered informative for the procedure if the variables are significantly intercorrelated within the total proband group. Sets of indicator variables can be chosen using a principal components analysis to summarise indicator intercorrelations. Consider three such variables, x, y, and z, that have extreme values for a subtype called A. The axiom of conditional independence assumes that the presence of intercorrelations among these variables increases with their ability to mark subtype A. That is, increasingly pathological values of each variable are associated with an increasing probability of belonging to subtype A. Since the procedure focuses on one subtype at a time, it can isolate subtype A from the total sample without making assumptions about the taxonomic structure of the rest of the sample.

The first step in the analysis is to find the value of each indicator that provides the maximum separation of probands with and without the subtype. This 'hitmax' cut-point is determined by computing the correlation between variables y and z from subsamples defined by successive intervals of the x (or input indicator) distribution. A hypothetical example may be useful. If there were a subtype of schizophrenia characterised by enlarged lateral ventricles, high scores on ratings of negative symptoms, and early age of onset, then one of these variables could be selected as the 'input indicator'. If age of onset were the input indicator, the sample would be divided into several groups at various values or cut-points of age at onset. The value of the age-at-onset cut-point at which the covariance within a subgroup on the ventricular size variable and negative symptom variable was maximised would be the 'best' cut-point for subdividing groups.

The covariance obtains a maximum value for the interval whose midpoint is the hitmax point because, based on the axiom of conditional independence, the observed covariance of two indicators will be at a maximum for samples containing equal portions of the two subtypes and will be at a minimum in the most homogeneous samples. Since each indicator distribution is assumed to estimate a latent bimodal distribution, successive intervals from the distribution should contain increasingly heterogeneous, then decreasingly heterogeneous, subsamples. The y, z covariance will show a corresponding increasing, then decreasing pattern. It is notable that such a pattern will not occur if the covariance between indicators is due to (a) a common correlation with a multifactorial liability, or (b) correlated measurement error. Thus, unlike factor or latent class analyses, the maxcov–hitmax method can discriminate between covariance due to heterogeneity and covariance due to other causes.

Using formulae provided by Meehl (1973), it is possible to obtain an estimate of the latent base rate of subtype A. The procedure can be repeated with different triads of indicators, letting the input indicator vary as well. Thus, one can obtain multiple estimates of the latent base rate of subtype A and the latent indicator means of the remaining subtype A patients. The observed agreement among the different estimates provides consistency tests of the results. A series of additional consistency tests is available to demonstrate that any emerging taxonomy is internally consistent.

A major advantage of the maxcov–hitmax method is that it is free of distributional assumptions. This makes it very suitable for the highly skewed data characteristic of putative measures of heterogeneity.

Conclusions

The MFP model is the most attractive – some might argue, seductive – alternative to aetiological heterogeneity. The idea that numerous, perhaps

https://doi.org/10.1192/bjp.156.1.17 Published online by Cambridge University Press

innumerable, factors are relevant to the phenomenon of schizophrenia is congenial to the impressions that we have formed in our clinical work with schizophrenic patients. A multifactorial approach seems needed to encompass the sometimes bewildering array of factors that previous research has suggested may be related to schizophrenia. However, as Dalen (1972) suggested, it may be helpful for us as researchers to take a more 'simple-minded' approach to meet the requirements of generating hypotheses that are capable of being empirically tested. It is important that we, as researchers, eschew vague notions of multicausality (or, worse yet, omnicausality). Dalen (1972) warned against the complacence that may result from embracing a generic and vague 'multifactorial' model that suggests that almost anything may be relevant to the aetiology of schizophrenia. Fortunately, investigators such as McGuffin *et al* (1987), and others, have provided us with rigorous versions of multifactorial polygenic models. If progress is to be made, it is important that we continue to investigate putative causal agents, whether under the guise of discovering which factors are among the relevant 'multifactors' or investigating heterogeneity among distinct 'level I indicators'.

At this time, evidence for genetic heterogeneity is neither clear-cut nor compelling. However, by default, through the inability of any single genetic model to accommodate all of the empirical evidence from family, twin, and adoption studies, the possibility of genetic heterogeneity remains viable, perhaps robust.

Acknowledgements

The preparation of this paper was supported, in part, by the Veterans Administration's Health Services Research and Development Program and National Institute of Mental Health Grants 1 RO1MH41879-01 and 1 R37MH43518-01.

References

AMERICAN PSYCHIATRIC ASSOCIATION (1980) *Diagnostic and Statistical Manual for Mental Disorders* (3rd edn). Washington DC: APA.

ANDREASEN, N. C., OLSEN, S. A., DENNERT, J. W., *et al* (1982) Ventricular enlargement in schizophrenia: relationship to positive and negative symptoms. *American Journal of Psychiatry*, **139**, 297–302.

ANDREWS, D. F. (1972) Plots of high-dimensional data. *Biometrics*, **28**, 125–136.

ANGRIST, B., ROTROSEN, J. & GERSHON, S. (1980) Responses to apomorphine, amphetamine, and neuroleptics in schizophrenic subjects. *Psychopharmacology*, **67**, 31–38.

BARNETT, V. & LEWIS, T. (1984) *Outliers in Statistical Data* (2nd edn). Chichester: Wiley.

BASSETT, A. S., MCGILLIVRAY, B. C., JONES, B. D., *et al* (1988) Partial trisomy chromosome 5 cosegregating with schizophrenia. *Lancet*, *i*, 799–801.

BENES, F. M., DAVIDSON, J. & BIRD, E. D. (1986) Cytoarchitectural studies of cerebral cortex of schizophrenics. *Archives of General Psychiatry*, **43**, 31–35.

BOWERS, M. B., SWIGAR, M. E., JATLOW, P. I., *et al* (1987) Early neuroleptic response: clinical profiles and plasma catecholamine metabolites. *Journal of Clinical Psychopharmacology*, **7**, 83–86.

BRADBURY, T. N. & MILLER, G. A. (1985) Season of birth in schizophrenia: a review of evidence, methodology, and etiology. *Psychological Bulletin*, **98**, 569–594.

BROWN, W. A. & HERZ, L. R. (1988) Neuroleptic response as a nosologic device. In *Handbook of Schizophrenia*, vol. 3 (eds M. T. Tsuang & J. C. Simpson). Amsterdam: Elsevier.

CHAKRABORTY, R. & HANIS, C. L. (1987) Nonrandom sampling in human genetics: skewness and kurtosis. *Genetic Epidemiology*, **4**, 87–101.

CHAPMAN, L. & CHAPMAN, J. (1977) Selection of subjects in studies of cognition. *Journal of Abnormal Psychology*, **86**, 10–15.

CLONINGER, C. R., REICH, T. & YOKOYAMA, S. (1983) Genetic diversity, genome organization, and investigation of the etiology of psychiatric diseases. *Psychiatric Developments*, **3**, 225–246.

CROW, T. J. (1980) Positive and negative schizophrenic symptoms and the role of dopamine. *British Journal of Psychiatry*, **139**, 379–386.

—— (1983) Is schizophrenia an infectious disease? *Lancet*, *i*, 173–175.

——, BERNDL, K., BOGERTS, B., *et al* (1987) Structural and functional deviations group report. In *Biological Perspectives of Schizophrenia* (eds H. Helmchen & F. A. Henn). Chichester: Wiley.

DALEN, P. (1972) One, two, or many? In *Genetic Factors in Schizophrenia* (ed. A. R. Kaplan). Springfield: Charles C. Thomas.

DAVISON, K. & BAGLEY, C. R. (1969) Schizophrenia-like psychoses associated with organic disorders of the central nervous system: a review of the literature. *British Journal of Psychiatry, Special Publication No. 4*, 113–184.

DE LISI, L. E. (1986) The use of positron emission tomography (PET) to image regional brain metabolism in schizophrenia and other psychiatric disorders: a review. In *Handbook of Schizophrenia*, Vol. 1 (eds H. A. Nasrallah & D. R. Weinberger). Amsterdam: Elsevier.

ELSTON, R. C. & CAMPBELL, M. A. (1970) Schizophrenia: the evidence for the major gene hypothesis. *Behavior Genetics*, **1**, 3.

FARAONE, S. & TSUANG, M. (1985) Quantitative models of the genetic transmission of schizophrenia. *Psychological Bulletin*, **98**, 41–66.

——, LYONS, M. & TSUANG, M. (1988) Mathematical models of genetic transmission. In *Handbook of Schizophrenia*, Vol. 3 (eds M. T. Tsuang & J. C. Simpson). Amsterdam: Elsevier.

FAUX, S. F., SHENTON, M. E., MCCARLEY, R. W., *et al* (1987) P200 topographic alterations in schizophrenia: evidence for left temporal–centroparietal amplitude deficits. *Electroencephalic Clinical Neurophysiology*, **40**, 681–687.

FEIGHNER, J. P., ROBINS, E., GUZE, S. B., *et al* (1972) Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry*, **26**, 57–63.

GOLDEN, R. & MEEHL, P. (1979) Detection of the schizoid taxon with MMPI indicators. *Journal of Abnormal Psychology*, **88**, 217–233.

HALDANE, J. B. S. (1941) The relative importance of principal and modifying genes in determining some human diseases. *Journal of Genetics*, **41**, 149–157.

HARRIS, H. & SMITH, C. A. B. (1947) The sib–sib age of onset correlation among individuals suffering from a hereditary syndrome produced by more than one gene. *Annual of Eugenics*, **14**, 309–318.

HOLZMAN, P. S. (1987) Recent studies of psychophysiology in schizophrenia. *Schizophrenia Bulletin*, **13**, 49–76.

KENDLER, K., GRUENBERG, A. & TSUANG, M. (1984) Outcome of schizophrenic subtypes defined by four diagnostic systems. *Archives of General Psychiatry*, **41**, 149–154.

https://doi.org/10.1192/bjp.156.1.17 Published online by Cambridge University Press

——, —— & —— (1985) Subtype stability in schizophrenia. *American Journal of Psychiatry*, **142**, 827–832.

——, —— & —— (1988) A family study of the subtypes of schizophrenia. *American Journal of Psychiatry*, **145**, 57–62.

KENNEDY, J. L., GIUFFRA, L. A., MOISES, H. W., *et al* (1988) Evidence against linkage of schizophrenia to markers on chromosome 5 in a northern Swedish pedigree. *Nature*, **336**, 167–169.

KIDD, K. K. & CAVALLI-SFORZA, L. L. (1973) An analysis of the genetics of schizophrenia. *Social Biology*, **20**, 254.

KNIGHT, J. G., KNIGHT, A. & PERT, C. B. (1987) Is schizophrenia a virally triggered antireceptor autoimmune disease? In *Biological Perspectives of Schizophrenia* (eds H. Helmchen & F. A. Henn). Chichester: Wiley.

LANDER, E. S. (1988) Splitting schizophrenia. *Nature*, **336**, 105–106.

LEWIS, S., REVELEY, A., REVELEY, M., *et al* (1987) The familial/sporadic distinction as a strategy in schizophrenia research. *British Journal of Psychiatry*, **151**, 306–313.

LIPTON, R. B., LEVY, D. L., HOLZMAN, P. S., *et al* (1983) Eye movement dysfunctions in psychiatric patients: a review. *Schizophrenia Bulletin*, **9**, 13–32.

MATTHYSSE, S. W. & KIDD, K. K. (1976) Estimating the genetic contribution to schizophrenia. *American Journal of Psychiatry*, **133**, 185.

MCGUFFIN, P., FARMER, A. & GOTTESMAN, I. (1987) Is there really a split in schizophrenia? The genetic evidence. *British Journal of Psychiatry*, **150**, 581–592.

MCNEIL, T. F. & KAIJ, L. (1978) Obstetric factors in the development of schizophrenia: complications in the births of preschizophrenics and in reproduction by schizophrenic parents. In *The Nature of Schizophrenia: New Approaches to Research and Treatment* (eds L. C. Wynne, R. L. Cromwell & S. Matthysse). New York: Wiley.

MEEHL, P. E. (1973) MAXCOV-HITMAX: A taxonomic search method for loose genetic syndromes. In *Psychodiagnosis: Selected Papers* (ed. P. E. Meehl). New York: Norton.

RAO, D. C., VOGLER, G. P., MCGUE, M., *et al* (1987) Maximum-likelihood estimation of familial correlations from multivariate quantitative data on pedigrees: a general method and examples. *American Journal of Human Genetics*, **41**, 1104–1116.

—— & WETTE, R. (1987) Nonrandom sampling in genetic epidemiology: maximum likelihood methods for multifactorial analysis of quantiative data ascertained through truncation. *Genetic Epidemiology*, **4**, 357–376.

REVELEY, A., REVELEY, M. & MURRAY, R. (1984) Cerebral ventricular enlargement in non genetic schizophrenia. *British Journal of Psychiatry*, **147**, 233–240.

SEIDMAN, L. J. (1983) Schizophrenia and brain dysfunction: an integration of recent neurodiagnostic findings. *Psychological Bulletin*, **94**, 195–238.

SHERRINGTON, R., BRYNJOLFSSON, J., PETURSSON, H., *et al* (1988) Localization of a susceptibility locus for schizophrenia on chromosome 5. *Nature*, **336**, 164–166.

SHUR, E. (1982) Season of birth in high and low genetic risk schizophrenics. *British Journal of Psychiatry*, **140**, 410–415.

SLATER, E. (1947) Genetical causes of schizophrenic symptoms. *Monatsschrift für Psychiatre und Neurologie*, **113**, 50–58. (Reprinted in *Man, Mind and Heredity* (eds J. Shields & I. I. Gottesman).)

WEINBERGER, D. R., BIGELOW, L. B., KLEINMAN, J. E., *et al* (1980) Cerebral ventricular enlargement in chronic schizophrenia: an association with poor response to treatment. *Archives of General Psychiatry*, **37**, 11–13.

*Ming T. Tsuang, MD, PhD, DSc, *Professor of Psychiatry and Director, Section of Psychiatric Epidemiology and Genetics, Harvard Medical School Department of Psychiatry, Massachusetts Mental Health Center; Director, Harvard Program in Psychiatric Epidemiology; Chief, Psychiatry Service, Brockton/West Roxbury Veterans Administration Medical Center*; Michael J. Lyons, PhD, *Assistant Professor, Department of Psychology, Boston University; Lecturer, Section of Psychiatric Epidemiology and Genetics, Harvard Medical School Department of Psychiatry, Massachusetts Mental Health Center; Research Psychologist, Psychiatry Service, Brockton/West Roxbury Veterans Administration Medical Center*; Stephen V. Faraone, PhD, *Instructor, Section of Psychiatric Epidemiology and Genetics, Harvard Medical School Department of Psychiatry, Massachusetts Mental Health Center; Research Psychologist, Psychiatry Service, Brockton/West Roxbury Veterans Administration Medical Center*

*Correspondence: *Brockton/West Roxbury Veterans Administration Medical Center, 940 Belmont Street, Brockton, Massachusetts, 02401, USA*

https://doi.org/10.1192/bjp.156.1.17 Published online by Cambridge University Press