Outcome in schizophrenia and related disorders compared between developing and developed countries

A recursive partitioning re-analysis of the WHO DOSMD data

THOMAS J. CRAIG, CAROLE SIEGEL, KIM HOPPER, SHANG LIN and NORMAN SARTORIUS

Background Data on the two-year pattern of course of illness have been collected in the WHO study of the Determinants of Outcomes of Severe Mental Disorder (DOSMD). These data are re-analysed using recursive partitioning, a method not yet applied to psychiatric data, to test the hypothesis that subjects from participating centres in developing countries had better outcomes than those in developed countries.

Method Subjects were those from the DOSMD study for whom two-year follow-up data were available (n=1056). The classification and regression trees recursive partitioning technique was used to examine the predictor variables associated with the outcome variable two-year pattern of course.

Results Pattern of course was best predicted by centre, but two developed centres (Prague and Nottingham) grouped with the developing country centres excluding Cali, having better outcomes than in the remaining developed country centres and Cali. Type of onset (insidious *v.* non-insidious) was the next strongest predictor, but its effect differed across these two centre groupings. Effects for some groups were modified by other predictor variables, including age, child and/or adolescent problems, and gender.

Conclusions The predominant predictor effects on two-year pattern of course continued to be centre and type of onset, but complex interactions between these variables and other predictor variables are seen in specific centre groupings not strictly defined by 'developing' and 'developed'.

The present paper presents a re-analysis of the data from the Determinants of Outcome of Severe Mental Disorder (DOSMD) study sponsored by the World Health Organization (WHO). That study was originally carried out to address outcome in first admissions with schizophrenia; in particular to test the unexpected finding of the earlier WHO International Pilot Study of Schizophrenia (IPSS; WHO, 1973) that outcome on six of seven measures was better for patients in developing countries than in the developed countries. The DOSMD study was developed to correct for the lack of homogeneity in the IPSS sample by using a 'treated incidence' sample. Over 1000 subjects from 12 centres in 10 countries were first assessed at study entry (close to the onset of psychotic illness) and were then followed for two years. As reported in a detailed WHO monograph, course and outcome were found, once again, to be more favourable for schizophrenic subjects living in centres classified as developing than for those in the developed grouping (Jablensky *et al*, 1992).

A subsequent report questioned whether the categories of developing and developed represented culturally homogeneous entities (Hopper, 1991). In addition, other studies of outcome in schizophrenics have produced conflicting results, some finding better outcomes among those in Third World countries, others finding them in industrialised settings (Warner, 1994). Edgerton & Cohen (1994) have argued that instead of viewing culture as relatively homogeneous in developing and developed countries, researchers should open this 'black box' through sustained and detailed examination of specific variables within local contexts. In view of this controversy and because of its strong epidemiological features, we re-analysed the DOSMD data using a statistical approach that allows a more detailed look at the patterns of prediction. This methodology, currently utilised in many areas in which classification or prediction are of concern, but not yet in psychiatry (Breiman *et al*, 1984; Goldman *et al*, 1988; Klein & Rolph, 1991; Clark & Pregibon, 1992), is presently used to identify predictors of course and outcome of schizophrenia. The method offers several advantages over earlier approaches that have used regression techniques: predictor variables do not have to be pre-categorised; linear as well as non-linear relationships of predictor variables to outcome can be developed; hierarchical orderings of the importance of predictor variables may be discerned. Specifically, the questions we sought to answer were:

(a) Is centre the most important predictor of course of schizophrenia over two years following onset?

(b) What is the relative importance of demographic and clinical variables previously shown to have a significant association with outcome in the prediction of short-term pattern of course?

(c) Are there interaction effects of the predictor variables on short-term pattern of course?

METHODS

Sample characteristics

For the re-analysis, we considered those subjects for whom two years of follow-up data were available (n=1078). From these, we eliminated those whose pattern of course could not be described (n=3) and those who were continuously non-psychotic over the follow-up (n=19). The centres included in the re-analysis and the sample size by centre are given in Table 1.

Outcome variables

The present analysis is based on pattern of course. This measure was chosen as the most clinically relevant from the six outcome measures found to be differentially better among developing centres by Jablensky *et al* (1992). Seven categories of pattern of course had been defined based on the number of psychotic episodes and degree of complete or incomplete remission during the two-year follow-up period (see Appendix 1). Jablensky *et al* (1992) considered two dichotomies of pattern of course: best *v.* all other, and worst *v.* all other. Similarly, our analysis looks at outcome in terms of pattern of course defined as 'good' (single or multiple episodes with complete remission between episodes) versus 'poor' (single

https://doi.org/10.1192/bjp.170.3.229 Published online by Cambridge University Press

or multiple episodes with incomplete remission between episodes, or continuous psychosis). (See Appendix 1.)

Prognostic variables

The re-analysis used the subset of demographic and clinical predictor variables that were significantly associated ($P \leqslant 0.05$) with the good and poor outcomes of the six measures considered by Jablensky *et al* (1992) in their logistic regression analyses. These include gender, marital status, setting (centre), type of onset (non-insidious *v.* insidious), ICD–9 clinical diagnosis, history of drug use, contact with close or casual friends, and adjustment problems in adolescence. Although they were not significantly associated with outcome, we also included history of alcohol use and adjustment problems in childhood, because these are of current clinical interest.

Diagnostic categorisation

Subjects in the DOSMD study had either received an ICD–9 clinical diagnosis of schizophrenia (295), a diagnosis of clinically related disorders (297; 298.3, .4, .8, .9; 291.3, .5; 292.1; 301.0, .2), or a diagnosis of other disorders with schizophrenia-suspected symptoms, or had been CATEGO classified on the basis of Present State Examination symptoms (Wing *et al*, 1974) into the S, P or O categories.

Table I Sample size and number in each pattern of course category by centre

Centre	All (*n*)	Pattern of course	
		Good[1] (*n*)	Poor[2](*n*)
Aarhus	80	22	58
Agra	76	59	17
Cali	139	63	76
Chandigarh (rural)	49	33	16
Chandigarh (urban)	105	62	43
Dublin	55	21	34
Honolulu	26	7	19
Ibadan	97	80	17
Moscow	156	30	126
Nagasaki	70	20	50
Nottingham	86	52	34
Prague	86	58	28
Rochester	31	12	19
All	1056	519	537

1. Single or multiple episodes of illness with complete remission between episodes.
2. Single or multiple episodes of illness with incomplete remission between episodes or continuous symptoms.

Statistical methods

In the first analysis presented by Jablensky *et al* (1992) the principal statistical method used to identify predictors of two-year outcome was log-linear modelling. The relationship between the log odds of outcome as measured by the various outcome variables and two predictor variables – centre (developed *v.* developing) and mode of onset (insidious *v.* non-insidious) – was first examined. The log-odds model posits that the odds (as a logarithm) that a subject will fall into one outcome category versus any other is a linear function of other independent or predictor variables. The analysis presented in the present paper uses recursive partioning (also known as tree-based modelling) both to identify groups differing in the pattern of course of illness and to understand the role of explanatory variables in predicting it (Breiman *et al*, 1984). The models that are produced are alternatives to additive models for regression problems and to additive logistic models for classification (Clark & Pregibon, 1992). Like the traditional approaches, tree-based models are used to relate classification or predictor variables to an outcome variable. The particular technique used here is known as classification and regression trees (CART; Breiman *et al*, 1984).

In its classification mode, one first prespecifies a set of classes into which the subjects fall, such as good and poor outcome. The technique aims to find category splits of the predictor variables that result in an accurate assignment of the subjects into the class in which they belong. For example, for the variable age, a split that produces a classification into outcome groups that is more accurate than any other split might be $\leqslant 25$ years and >25, because most young persons have good outcome and persons over 25 have poorer outcomes. The procedure is considerably more complicated than this since the algorithm considers all variables and all possible splits at each stage, and chooses the best variable split among all variables based on a misclassification error. It iteratively continues the process at each node of a split. (Because this approach is not yet well known, we provide greater details in Appendix 2.)

CART analyses were conducted in a classification mode for the pattern of course grouping. Variables found by Jablensky *et al* (1992) to be predictive of outcome were used as the predictor variables. Centres were entered ungrouped. The sample size for the pattern of course analysis is 1056. Figure 1 depicts the CART tree discussed in the text. It displays the splitting variables, the variable categories for each split, and at each node the percentage of subjects in each pattern of course grouping.

RESULTS

The resubstitution error for misclassification for the tree for pattern of course is 27%, and the cross-validation error is 37%.

Centre was the first splitting variable (Fig. 1). The developing centres Agra, Chandigarh (both rural and urban) and Ibadan grouped together with the developed centres Nottingham and Prague. The remaining developed centres formed the other node of the split along with a developing centre, Cali. Subjects in the grouping comprising predominantly developing centres were more than twice as likely as their counterparts to have better pattern of course. After the initial split on centre, at both nodes the next split was on the variable mode of onset. In the grouping consisting predominantly of developing centres, 77% of those with non-insidious onsets (a terminal node) experienced complete remissions, compared with only 44% of the same onset group in the predominantly developed centres. Eighty-three per cent of those with insidious onset in the group of predominantly developed centres had poor pattern of course (a terminal node), in contrast to only 50% of their counterparts in the predominantly developing group. Within the predominantly developing centres, additional contributors to poor pattern of course in the insidious onset group include both presence of childhood and adolescent adjustment problems and older age. The developed centre grouping of non-insidious onset was split a second time on centre, showing higher proportions of subjects in Honolulu and Moscow (73%) having poorer pattern of course than those in Aarhus, Cali, Dublin, Rochester and Nagasaki (49%). In addition, for the latter group of centres, gender entered the model, with proportionately more females than males having good outcome.

DISCUSSION

This analysis permits an assessment of the differential impact that predictor variables

https://doi.org/10.1192/bjp.170.3.229 Published online by Cambridge University Press

have on illness trajectory. Because the categories of a variable were not pre-grouped in advance, this statistical procedure was able to reveal a more detailed relationship between a variable and pattern of course. It also allowed a ranking of the hierarchical importance of variables.

Centre as a predictor variable

The aggregate finding of Jablensky *et al*, 1992, that the cohort of persons with an initial diagnosis of schizophrenia in the developing centres had a more benign pattern of course than those in developed centres, was generally upheld, but clear exceptions to this rule emerged. The tree-modelling approach revealed that in two developed centres (Prague and Nottingham) pattern of course was indistinguishable from that in the developing centres (and better than that in the other developed centres), while one developing centre (Cali) fell in the developed category. These exceptions may speak to the inconsistencies in the findings of several studies, in which clinical outcomes in developing countries were variously reported to be better, worse or similar to those for developed countries (Warner, 1994).

Relative importance of predictor variables

Several of the predictor variables found by Jablensky *et al* (1992) to be significantly associated with outcome entered the tree. Centre is the first splitting variable, followed by type of onset for both branches of the tree. Subsequently, the rank order importance of other variables differs within the two centre groupings formed by the first split. In the predominantly developing centres, further predictors include (in hierarchical order): adjustment problems of childhood and adolescence, and age (Fig. 1). In the predominantly developed centres, only gender entered the selected tree.

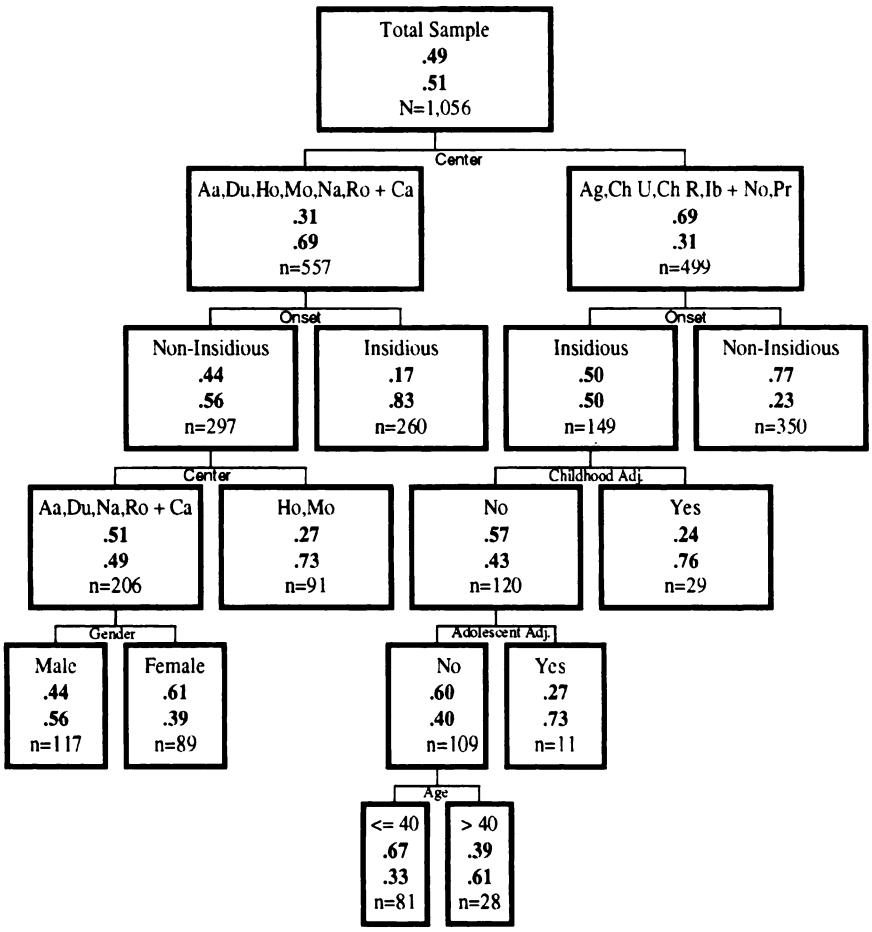


Fig. 1 Prediction tree for pattern of course. Proportion in 'good' and 'poor' outcome groups, respectively (see Appendix I), displayed in each node. Aa, Aarhus; Ag, Agra; Ca, Cali; Ch R, Chandigarh (rural); Ch U, Chandigarh (urban); Du, Dublin; Ho, Honolulu; Ib, Ibadan; Mo, Moscow; Na, Nagasaki; No, Nottingham; Pr, Prague; Ro, Rochester; Childhood/adolescent adj., childhood or adolescent adjustment problems.

Interactive effects of predictor variables

The tree reveals complex interactive patterns. For example, type of onset affects pattern of course but the relationship differs for each centre grouping. Those with non-insidious onset do very well in the predominantly developing centres, but so do half of those with insidious onset. In addition, if among these persons with insidious onset we further narrow the focus to those who had good premorbid adjustment (indicated by no childhood or adolescence adjustment problems) and who were younger, the percentage with a good pattern of course increases to 67%. On the other hand, in the developed centres (plus Cali), having non-insidious onset does not necessarily predict good pattern of course, as is seen for subjects in Honolulu and Moscow. Moreover, those with insidious onset tend to fare quite poorly, with no variables further modifying this effect.

The variations observed among predictor variables and outcome variables permit the posing of a number of specific questions that attempt to open the 'black box' of context and explore whether there is a cultural pattern (e.g. available extended kinship) or a mental health service system factor (e.g. supportive housing, availability of intensive drug treatment) that might provide an important ingredient crucial to a particular outcome. For example, there is an apparent advantage for those with insidious onset to be in a grouping composed of predominantly developing centres. However, this advantage is mitigated by the presence of premorbid childhood or adolescent adjustment problems. Does this finding suggest that such premorbid vulnerability factors neutralise whatever beneficial effect exists for those subjects in the predominantly developing centres? For non-insidious onset subjects in the developed centres, what are the factors shared by Honolulu and Moscow which negate the advantage of non-insidious onset and result in persons more likely to have patterns of course similar to subjects with insidious onset? Clearly, the data available do not permit a definitive exploration of these issues, but they do serve to highlight the complex role that culture may have in shaping findings.

https://doi.org/10.1192/bjp.170.3.229 Published online by Cambridge University Press

Limitations

Several limitations should be noted regarding our choices of predictor variables and pattern of course. First, we chose to include all predictor variables which had shown a significant association with any of the six outcome measures examined by Jablensky *et al* (1992) in order to examine the full range of possible interaction affects. However, it must be noted that some predictors (e.g. centre, age, gender, marital status) have clearly inherent validity and reliability, while others (e.g. adjustment problems, alcohol/substance use, contact with friends, etc.) are more subject to variability in retrospective recall by the subject and/or other informants. Second, our choice to categorise pattern of course into two levels based on quality of remission (e.g. complete *v.* other) was done to make it comparable with the categorisation used by Jablensky *et al* (1992) between the developing and developed countries. This categorisation does not address the possible effect of number of episodes on this comparison, although the overall proportion of the original study sample with two or more episodes was virtually identical for those with complete (16.6%) and incomplete (16.5%) remissions. Somewhat weakening both this and prior analyses are the problems inherent in retrospective assessment of quality of remission and end (or beginning) of episodes. If the technical problems of thoroughness of follow-up, clinical assessments, and quantification of outcome varied substantially across centres, despite the rigorous DOSMD study design, this too may have added to the heterogeneity of findings and been magnified by the use of a more discriminatory statistical procedure. Finally, it should be noted that the present study reports on findings only for pattern of course, whereas the monograph has considered several other outcome measures. The authors have carried out one other CART analysis on the outcome variable used by Jablensky *et al* (1992), percentage time with unimpaired functioning. Centre remained the first splitting variable and, as with the pattern of course, the centres did not strictly dichotomise into developed and developing centres.

CONCLUSIONS

Despite these potential limitations, our identification of different patterns of course among both developing and developed centres underscores the need for further study within the 'black box' of imputed cultural effects on outcome – in particular, of stigma, family support, availability of treatment and rehabilitation services, social relationships, and available employment for disabled individuals – as has been extensively discussed (Hopper, 1991; Warner, 1994; Edgerton & Cohen, 1994). In addressing these issues, closer study from a variety of disciplines (such as anthropology) is needed to further clarify actual day-to-day functioning of study subjects, practical management of burden and provision of support by family members and coping habits. These aspects need to be detailed within different cultural contexts where attribution of mental illness, duties of kin and the value and availability of non-market labour vary widely.

Perhaps the most crucial remaining question is the relationship of these observed short-term predictors and intervening variables with the long-term (e.g. 15+ years) course of psychotic illness. This is especially of interest in view of provocative evidence first, that predictors of long-term outcome may be quite different from those of short-term outcome (McGlashan, 1986); second, that long-term outcomes in schizophrenia may be considerably better than has been previously thought, even among relatively deteriorated patients (Huber *et al*, 1975; Bleuler, 1978; Ciompi, 1980; Harding *et al*, 1987; Jablensky, 1989); and third, that profound changes in the character of the cultural setting (e.g. via migration, industrialisation and changes in the health care system) have occurred over the decade and a half since the original DOSMD data were collected (Hopper, 1991). Under WHO auspices, a study is currently under way aimed at addressing these questions (Sartorius *et al*, 1996).

APPENDIX 1

Based on the number of psychotic episodes and degree of complete or incomplete remission during the two-year follow-up period, Jablensky *et al* (1992) defined the following seven categories of pattern of course:

(1) single psychotic episode, complete remission;

(2) single psychotic episode, incomplete remissions;

(3) single psychotic episode followed by one or more non-psychotic episodes, complete remission;

(4) single psychotic episode followed by one or more non-psychotic episodes, incomplete remissions;

(5) two or more psychotic episodes, complete remission;

(6) two or more psychotic episodes, incomplete remissions; and

(7) continuous psychotic illness (no remission).

For the present study, we defined two patterns of course based on the above classification: our 'good' category comprised classes 1, 3 and 5 above; our 'poor' category comprised classes 2, 4, 6 and 7.

APPENDIX 2

Classification and regression tree methodology

The classification and regression tree (CART) methodology, in its classification mode, begins by forming for a predictor variable the subgroups defined by either a range split of its values (if it is continuous), or a grouping of a variable's categories into two sets (if it is categorical). A subject in a subgroup is considered as classified into the outcome class into which the largest percentage in the subgroup falls. CART then calculates a misclassification error, the error representing the percentage in the subgroup that is incorrectly classified in this way. The computer-intensive technique runs through all possible splits or groupings for all predictor variables. At this first step, it chooses the variable and its split that provides an optimal classification in the sense that the error of misclassifying subjects into the designated outcome group is minimised. Two nodes of the tree are now established representing the subgroups formed by the optimal variable split. At each node, the process continues. CART takes account of the fact that different relationships may hold between variables in different parts of the data field. Thus, the best variable split at one node will generally differ from the best split at another. Additionally, the same predictor variable may be re-introduced in a sub-branch with a further split of its categories. A large tree is grown at whose extremes each terminal node contains only a few cases. The algorithm then selectively prunes branches off this large tree, yielding a sequence of smaller and smaller trees; each tree in the sequence is chosen as the one that has the smallest misclassification error of any other sub-tree of comparable size. The tree whose misclassification error is not significantly improved by the pruned branches is chosen as the best sub-tree.

Tree-based models are able to capture non-additive behaviours. Interactions between variables do not need to be pre-specified and are revealed through tree-growing. In addition, missing data can be accommodated, and the technique provides for cross-validation assessment of its results. In the pruning procedure, rather than using the apparent (resubstitution) misclassification error or apparent mean square error, a cross-validation method is used to obtain a more reliable estimate of error, representing the error that would be experienced in replication (Breiman *et al*, 1984).

https://doi.org/10.1192/bjp.170.3.229 Published online by Cambridge University Press

ACKNOWLEDGEMENT

This research was partly funded by the Laureate Foundation, Tulsa, Oklahoma, USA and NIMH grant MH51359. This paper used data obtained during the World Health Organization (WHO) study on Determinants of Outcome of Severe Mental Disorders coordinated at WHO Headquarters in Geneva by N. Sartorius and A. Jablensky. The chief collaborating investigators in the field research centres were: Aarhus: E. Stromgren; Agra: K. C. Dube; Cali: C. Leon; Chandigarh: N. N. Wig and V. Varma; Dublin: D. Walsh; Honolulu: A. Marsella and M. Katz; Ibadan: M. Olatawura; Moscow: R. A. Nadzharov and N. N. Zharikov; Nagasaki: R. Takahashi and Y. Nakane; Nottingham: J. E. Cooper; Prague: L. Hanzlicek and C. Skoda; Rochester: L. C. Wynne and T. Gift.

REFERENCES

Bleuler, M. (1978) *The Schizophrenic Disorder: Long-Term Patient and Family Studies*. New Haven, CT: Yale University Press.

Breiman, L., Friedman, J., Olshen, R., et al (1984) *Classification and Regression Trees*. New York: Chapman and Hall.

Ciompi, L. (1980) The natural history of schizophrenia in the long term. *British Journal of Psychiatry*, **136**, 413–420.

Clark, L. A. & Pregibon, D. (1992) Tree-based models. In *Statistical Models 5* (eds J. M. Chambers & T. J. Hastie), pp. 377–420. Pacific Grove, CA: Wadsworth & Brooks/Cole.

Edgerton, R. B. & Cohen, A. (1994) Culture and schizophrenia: The DOSMD challenge. *British Journal of Psychiatry*, **164**, 222–231.

Goldman, L., Cook, E. F., Brand, D. A., et al (1988) A computer protocol to predict myocardial infarction in emergency department patients with chest pain. *New England Journal of Medicine*, **318**, 797–803.

Harding, C. M., Brooks, G. W., Ashikaza, T., et al (1987) The Vermont Longitudinal Study of Persons with Severe Mental Illness: II: Long term outcomes of subjects who retrospectively met DSM–III criteria for schizophrenia. *American Journal of Psychiatry*, **144**, 727–735.

Hopper, K. (1991) Some old questions for the new cross-cultural psychiatry. *Medical Anthropology Quarterly*, **5**, 299–329.

Huber, G., Gross, G. & Schelttler, R. (1975) A long term follow up study of schizophrenia: Psychiatric course of illness and prognosis. *Acta Psychiatrica Scandinavica*, **52**, 49–57.

Jablensky, A. (1989) Epidemiology and cross cultural aspects of schizophrenia. *Psychiatric Annals*, **10**, 516–524.

—, Sartorius, N., Ernberg, G., et al (1992) Schizophrenia: Manifestations, incidence and course in different cultures. A World Health Organization ten-country study. *Psychological Medicine, Monograph Supplement*, **20**, 1–97.

Klein, S. P. & Rolph, J. E. (1991) Relationship of offender and victim race to death penalty sentences in California. *Jurimetrics*, **32**, 33–44.

McGlashan, T. H. (1986) Predictors of shorter-, medium-, and longer-term outcome in schizophrenia. *American Journal of Psychiatry*, **143**, 50–55.

Sartorius, N., Gulbinat, W., Harrison, G., et al (1996) Long term follow up of schizophrenia in 16 countries: A description of the International Study of Schizophrenia conducted by the World Health Organization. *Social Psychiatry and Psychiatric Epidemiology*, **31**, 249–258.

Warner, R. (1994) *Recovery from Schizophrenia*, (2nd edn), pp. 151–155. New York: Routledge.

World Health Organization (1973) *Report of the International Pilot Study of Schizophrenia*. Geneva: WHO.

— (1979) *Schizophrenia: An International Follow-Up Study*. New York: John Wiley.

Wing, J. K., Cooper, J. E. & Sartorius, N. (1974) *Measurement and Classification of Psychiatric Symptoms*. London: Cambridge University Press.

CLINICAL IMPLICATIONS

- While insidious onset is a strong predictor of poor pattern of course, its strength as a predictor varies both within and between the centre groupings, implying that a comprehensive examination of patients remains important.
- The finding that in the developing centres persons with insidious onset could have good outcome, coupled with the finding that prognosis was poorer for those with insidious onset and poor premorbid adjustment, suggests that interventions be intensified for this latter group.
- Clinical care might be improved if cultural effects could be more finely differentiated and used to tailor treatment to specific circumstances of patients.

LIMITATIONS

- Certain predictor variables may have been subject to unmeasured recall bias (e.g. adjustment problems in childhood and adolescence) in ways that are not uniformly distributed across sites.
- The categorisation of pattern of course into two levels based on quality of remission does not allow us to address other distinguishing features of course that might be predicted.
- The use of the new statistical technique with finer discriminatory power may have magnified the heterogeneity in findings that was due to variations in the conduct of the study across centres.

THOMAS J. CRAIG, MPH, Brooklyn VA Medical Center, Brooklyn, New York; CAROLE SIEGEL, PhD, KIM HOPPER, PhD, SHANG LIN, PhD, Nathan S. Kline Institute for Psychiatric Research, Orangeburg, New York; NORMAN SARTORIUS, MD, Department of Psychiatry, University of Geneva, Switzerland

Correspondence: Dr Carole Siegel, Head, Epidemiology and Health Services Research Laboratory, Statistical Sciences & Epidemiology Division, The Nathan S. Kline Institute for Psychiatric Research, 140 Old Orangeburg Road, Orangeburg, NY 10962, USA

(First received 19 June 1996, final revision, 23 September 1996, accepted 10 October 1996)