

the international pilot study of schizophrenia*

The International Pilot Study of Schizophrenia (IPSS) began in 1966 as a large-scale cross-cultural collaborative project carried out simultaneously in nine countries that differ widely in their sociocultural and economic characteristics: Colombia, Czechoslovakia, Denmark, India, Nigeria, China, the Union of Soviet Socialist Republics, the United Kingdom, and the United States of America (World Health Organization 1973).

Designed as a pilot study, the IPSS set out to lay methodological groundwork for future international epidemiological and other research in schizophrenia as well as in other functional psychiatric disorders. The study was sponsored by the World Health Organization (WHO) and financed from three sources: the field research centers in each of the nine countries, the National Institute of Mental Health, and WHO. In 1968, 2 years after preparations for the study were initiated, the first patient was examined; in 1971, a 2-year followup of the initial cohort of 1,202 patients was completed; and a 5-year followup is now in progress.

Although there have been a large number of studies on the distribution, clinical picture, course, and outcome of schizophrenia, a number of problems have made it difficult to compare the results of such studies or to draw clear conclusions from individual studies about the nature of schizophrenia. These problems have included 1) the use of different diagnostic criteria in the different studies, 2) a lack of standardized and reliable methods of

psychiatric assessment, 3) imprecision and variation in the definition of outcome criteria, 4) variations in methods of followup, and 5) insufficient consideration of intervening variables on the course and outcome of schizophrenia. In particular, there has been a lack of prospective followup studies, using standardized, reliable methods of assessment, to compare the nature and course of schizophrenia and other functional psychoses in a variety of cultures.

In this context, it was the aim of the IPSS to answer certain basic methodological questions and to provide information about the nature and distribution of schizophrenia. The three major methodological questions were:

- Is it feasible to carry out a large-scale international psychiatric study that requires the coordination and collaboration of psychiatrists and mental health workers from different theoretical backgrounds and from widely separated countries with different cultures and socio-economic conditions?
- Is it possible to develop standardized research instruments and procedures for psychiatric assessment that can be reliably applied in a variety of cultural settings?
- Can teams of research workers be trained to use such instruments and procedures so that comparable observations can be made in developed and developing countries?

The major questions about the nature and distribution of schizophrenia that the study was intended to answer were:

- In what sense can it be said that schizophrenic disorders exist in different parts of the world?

*This paper has been prepared jointly by Dr. Norman Sartorius, Chief, Office of Mental Health, World Health Organization, Geneva, Switzerland (principal investigator); Dr. Robert Shapiro, Department of Psychiatry, University of Copenhagen, Copenhagen, Denmark (formerly collaborating investigator of the International Pilot Study of Schizophrenia at WHO Headquarters); and Dr. Assen Jablensky, Senior Lecturer, Department of Psychiatry, Sofia, Bulgaria (Consultant, Office of Mental Health, WHO Headquarters).

- Are there groups of schizophrenic patients with similar characteristics present in every one of the countries studied?

- Are there groups of schizophrenic patients whose symptoms differ in form or content from one country to another and, if so, are such differences the result of variations in diagnostic practice, or are they true cultural differences in the manner of presentation of the various types of schizophrenia?

- Does the clinical course of schizophrenia in one country differ from that in other countries?

- How do the characteristics of schizophrenic patients compare with those of other psychoses in various countries?

- Does the course of other psychoses differ from country to country?

With regard to these questions, three major conclusions stand out from the IPSS experience during the initial evaluation phase. First, it was demonstrated that it is possible to carry out effectively a large cross-cultural investigation if careful attention is paid to developing central coordination and providing frequent opportunities for face-to-face contact among investigators. Second, it was shown that it is possible to develop standardized and reliable research instruments and procedures for practical use in psychiatric studies and to train teams of research workers to use instruments and procedures so that comparable observations can be made both in developed and in developing countries. Third, symptomatologically similar groups of schizophrenic patients could be identified in every one of the nine centers involved in the study, and these groups of schizophrenic patients were symptomatologically different from patients in other diagnostic groups. Some groups of schizophrenic patients with center-specific characteristics were also found.

The IPSS methodology, main results, and progress are summarized below.

Method

Phasing

The IPSS was carried out in three phases: a preliminary phase, an initial evaluation phase, and a followup phase. During the preliminary phase, administrative,

operational, and organizational procedures were established and tested. In the main phase, approximately 135 patients were selected and examined in each center from among those patients contacting the centers during the 1-year period from 1 April 1968 to 1 April 1969, according to procedures and methods developed during the preliminary phase. Two years after the initial evaluation, the patients received a followup evaluation, which is being repeated 5 years after their inclusion in the study.

Selection of the Field Research Centers

The nine field research centers were selected on the basis of the following criteria: 1) the existence of a network of services capable of detecting a sufficient number of the likely cases of schizophrenia occurring in the population at risk; 2) the presence of several well-trained and motivated psychiatrists; 3) the possibility of setting up a reporting system so that potential cases would be known to the participating psychiatrists; 4) the availability of census data on the whole population of the area; 5) the absence of very high death or emigration rates that could make followup difficult, or the high prevalence of organic diseases that might mask or obscure the psychotic picture so as to make the diagnosis of schizophrenia difficult; and 6) the existence of a recognizable and distinct local culture or cultures.

The centers finally chosen on this basis were situated in Aarhus, Denmark; Agra, India; Cali, Colombia; Ibadan, Nigeria; London, England; Moscow, U.S.S.R.; Taipei, China; Washington, D.C., U.S.A.; and Prague, Czechoslovakia. The coordination of the research activities and a major part of the data analyses were carried out by headquarters at WHO, Geneva. The centers in Aarhus, London, and Washington contributed considerably to data analyses.

At the beginning of the study, the participating psychiatrists were brought together for joint training in the use of the instruments for assessment of patients. In the course of the study, regular meetings and exchanges of visits took place to review progress and plan future activities. It is to be emphasized that the frequent opportunities for face-to-face contact and discussion between investigators from different countries were major factors in the achievement of methodological consistency throughout the project.

Identification of Patients

To obtain a series of patients who might be sufficiently homogeneous for the purposes of the study, and who were likely to be available for long-term followup, all patients contacting each center were assessed with a demographic screen and a psychosis screen. The demographic screen was designed to exclude patients who fell outside the 15-44 age range and who had resided or slept regularly in the center's catchment area for less than 6 months prior to the index contact. The 15-44 age range was agreed upon in order to exclude psychopathology that might be outside the scope of the study; for example, childhood and juvenile psychoses, or presenile or senile psychoses. The residential criterion was designed to increase the likelihood of availability for followup.

The psychosis screen contained both exclusion and inclusion criteria. Among the former were 1) evidence of chronic psychotic illness (e.g., presence of psychotic symptoms for over 3 years prior to the current episode, or history of prolonged hospitalization), 2) abuse of alcohol or CNS-affecting drugs, 3) mental retardation, 4) evidence that the psychosis may have been caused or significantly influenced by an organic condition and 5) severe sensory, language, or speech difficulties which might impede the administration of the interview.

The inclusion criteria consisted of 10 categories representing areas of psychopathology usually regarded as indicative of psychosis. They were divided into symptoms whose presence automatically qualified the patient for inclusion, regardless of degree of symptomatology, and symptoms which qualified a patient for inclusion only if present to a severe degree. The symptoms included in the first category were delusions, hallucinations, gross psychomotor disorder, and definitely inappropriate and unusual behavior. The second category of symptoms comprised social withdrawal, disorders of thinking other than delusions, overwhelming fear, disorders of affect, self-neglect, and depersonalization.

Data Collection Instruments and Techniques

Eight instruments were used to assess patients in the initial evaluation phase of the study. The three basic

instruments were the Present State Examination (PSE), the Psychiatric History Schedule, and the Social Description Schedule.

Present State Examination

The PSE is a guide to structuring the clinical psychiatric interview aimed at obtaining a systematic, reliable, and valid description of the present mental state of patients suffering from one of the functional psychoses or neuroses. It provides items to be used in observing and questioning patients, which systematically cover all the areas of psychopathology usually explored in the course of a comprehensive clinical examination of a patient's current mental condition. There are instructions on how each item should be coded.

Although the PSE is a guide to structuring the clinical interview, it can be flexibly adjusted according to the clinical style of the interviewer and the necessities of the clinical situation. To facilitate the conduct of the interview, the items are grouped into sections, but the interviewer is not obliged to follow the order of sections in the schedule. For example, if the patient mentions particular psychiatric symptoms at the beginning of the interview, the interviewer can start his formal questioning at the appropriate section in the schedule. The clinical principle of cross-examination is followed throughout. The interviewer is instructed to ask questions until he is satisfied that a given symptom is present or absent or that no clear decision about the symptom can be made at that time. In this way the ratings do not merely reflect the patient's answers to questions but, rather, represent the psychiatrist's judgment about the presence of psychopathological phenomena.

The PSE schedule, developed over 11 years ago (Wing, Cooper, and Sartorius 1974), has gone through nine editions. It has been tested extensively and has been used in a number of studies, including the U.S.-U.K. diagnostic project (Cooper et al. 1972). The schedule used in the IPSS was a modified version of the seventh edition, which contained a total of 360 items. On the basis of their experience with a test of the eighth edition of the PSE made at each center, the collaborating investigators made suggestions about wording, ordering, additions, and deletions that would make the schedule more applicable to the particular circumstances of their centers. These suggestions, together with those based on

the experience of the team working on the U.S.-U.K. diagnostic study and of members of the Medical Research Council Social Psychiatry Unit in London, were incorporated into the version of the PSE schedule used in the study.

For the purposes of the study, the PSE was translated from English into seven languages: Danish, Hindi, Spanish, Yoruba, Russian, Chinese, and Czech. Repeated back translations¹ and many discussions were carried out to increase interlanguage equivalence. That this objective has been achieved can be seen from these factors: 1) The "target check," which includes a search for errors of meaning and which is considered important in the literature on translating research instruments, has been carried out frequently in all centers; 2) most of the members of the research team had at least some knowledge of the "source language" (English), and a considerable number of them were fluent in both languages; 3) in the statistical analyses, which were carried out to compose the units of analysis (see below), very similar patterns of correlations were found in the various centers; and 4) the psychiatrists and the other members of the research teams have been instructed about the meaning of the items in the schedules and the manner of using them.

To assess the reliability of the PSE, several procedures were undertaken at the levels of individual items ($N = 360$), groups of items combined to correspond more closely to clinical symptoms (units of analysis, $N = 124$), and groups of units of analysis representing major broad areas of psychopathology ($N = 27$). To assess intracenter reliability, an average of 21 interviews rated simultaneously by two psychiatrists were conducted in each center. The median value of the intraclass correlation coefficient (R) was found to be .77 for individual items, .81 for units, and .84 for groups of units of analysis. A total of 51 patients from the different centers were interviewed consecutively (within a week) by two psychiatrists to test the repeatability of the ratings. The median R for groups of units in this series of interviews was .57. In a test of the intercenter reliability, 21 interviews held in different centers were rated live or from videotape or films by an average of 10 psychiatrists

from different centers. The median R for units of analysis was .45, and for groups of units, it was .57. (These reliability analyses were carried out at the Washington center.) The reliability level was consistently higher for units of analysis based on patient-reported experiences than for units rated from direct observation.

Thus, the reliability of the study's principal research instrument, the PSE, was found to be reasonably high in spite of the obvious difficulties associated with the design of a multicenter, cross-cultural study.

Past History and Social Description Schedules

The task of constructing and standardizing the Psychiatric History and Social Description Schedules was in many ways a more complicated one because of the multitude and magnitude of the cultural and socioeconomic differences between the countries involved and because of the limited previous international experience with these types of instruments. The Psychiatric History Schedule was designed to cover areas such as previous illnesses and hospitalizations; history, symptomatology and course of the present episode; treatment; premorbid personality traits; and psychosexual adjustment, occupational history, use of alcohol and drugs, and overall satisfaction with the premorbid life situation. The Social Description Schedule contained items related to parents' and spouse's education and occupation; type of household; and patient's education, religion, marital status, work activities, and birth order, among others.

In view of the relative lack of previous work on the development of a standardized instrument for the transcultural collection of psychiatric history and social description data, the investigators felt that, during the initial evaluation phase of the study, the primary task with regard to the development of such schedules was to identify items that would be applicable and useful in a variety of cultures. This was viewed as the initial stage in the eventual evolution of useful instruments. Thus, much less emphasis was placed on testing the reliability of these instruments than on testing the reliability of the PSE, since it was felt that problems of applicability should be approached first. Some reliability assessments were carried out, however.

In a study of intercenter reliability of these two instruments, 15 raters from the different centers rated a

¹ After the PSE was translated from English into one of the seven target languages, it was given to a second translator who translated it *back* into English. The original and the back translations were then compared.

videotaped psychiatric history and social description interview. Difficulties arose, however, in reaching agreement on many items. This finding was not unexpected, and its significance was attenuated by the reasonable degree of agreement between raters when the two instruments were tested for intracenter reliability (36 paired interviews). The agreement was as low as 50 percent on only two items, and as high as 90 percent or complete on 70 items.

Diagnostic Reliability

The study of intracenter agreement by raters on diagnosis (190 simultaneous paired interviews) indicated a complete agreement in 165 cases out of 189 (87.3 percent). The agreement between interviewer and observer (see table 1) on the 3-digit International Classification of Diseases of the World Health Organization (1967) diagnostic category of schizophrenia was very high (91.3 percent). Agreement was acceptably high for all of the remaining diagnostic categories (some of them 4-digit categories), except for mania. Intercenter reliability in making diagnoses was tested on multiple rating of videotaped interviews. It was found to be satisfactory (ranging between 82 and 100 percent) for 3-digit diagnostic categories and less satisfactory when 4-digit diagnoses were made.

Training

An important element, which was regarded as essential to the uniform application of the research instru-

ments, was the training of the participating psychiatrists in the use of the PSE. To this end, two training seminars were held in London in 1967. Subsequently, the 26 paired interviews, which were carried out in each center during phase 1, served both for the assessment of the PSE and for further training of the investigators under "field" conditions. The degree of uniformity thus achieved was maintained or improved by simultaneous interviews carried out at regular intervals throughout the periods when patients were being taken into the study or re-interviewed for followup purposes. New psychiatrists joining the project had to do at least five simultaneous interviews with each of the other psychiatrists in their center.

Procedure for Data Collection

The collection of data during the initial evaluation phase proceeded in the following way: 1) Within 2 weeks after the patient's initial contact with the psychiatric facility at the field research center, his mental state was assessed, with the use of the PSE, by a project psychiatrist; 2) the past history of the patient and his illness was obtained, with the use of the Past History Schedule, through an interview of the patient or an informant by a psychiatrist, a psychologist, or a social worker; 3) social and demographic information concerning the patient and his family was obtained by a social worker, with the use of the Social Description Schedule, in an interview with the patient or an informant; and 4) on completion of all interviews, the psychiatrist recorded his diagnosis and prognosis of the case, as well as

Table 1. Agreement of diagnosis between interviewer and observer in simultaneous interviews in nine field research centers.

Diagnosis	Number of interviews	Agreement
Schizophrenia	127	116
Mania	11	6
Psychotic depression	16	13
Paranoid states	5	4
Other psychoses	15	12
Neurotic depression	7	7
Personality disorders	8	7
Total¹	189	165

¹ One interview was excluded because no diagnosis was made.

the reasons for his judgment, on a diagnostic assessment form.

Approaches to Data Analysis

The initial evaluation phase of the IPSS generated a very large amount of data. Over 2 million pieces of information were accumulated for the 1,202 patients in the study. In order to analyze these data so that conclusions could be reached about characteristics of different patient groups, four main methods of data analysis were carried out:

1) The psychopathology of the patient groups, as indicated by ratings on the PSE, was analyzed and described in the form of symptom profiles. These analyses were carried out primarily on two levels: the level of units of analysis and the level of groups of units of analysis. Starting with the 360 PSE items, those items that seemed to be facets of the same symptoms were grouped together on the basis of clinical judgment and statistical association indices. Such indices were calculated for items that were hypothesized as belonging together. The results of these analyses were then used to reform the groupings, which were then retested, and so on until units were established whose composition had stability, regardless of who examined the patient, what kind of patient was examined, and the center from which the patient came. Eventually, 129 units of analysis were established (e.g., "delusions of persecution," which includes items such as, Did you notice that some force was trying to act on you? to harm you? Did you notice that somebody was following you around, or spying on you?).

These 129 units were further condensed into 27 groups of units of analysis, which represent broader categories of symptoms (e.g., "delusions," made up of a number of units of analysis, each of which represents a specific type of delusion). Condensing the data in this way made it possible to describe a clinical profile for any patient or group of patients in the study on the basis of units or groups of units of analysis.

2) The PSE data were used to classify patients according to a computer diagnostic program (CATEGO), which was designed to provide a completely standard reference classification. These analyses were carried out in the field research center in London.

3) The PSE data were used as the basis for a cluster analysis of the IPSS patients. Cluster analyses were carried out in several centers; the analyses presented in volume 1 of the report of the IPSS (World Health Organization 1973) and referred to in this paper were carried out in the Washington center.

4) The three methods described above were combined to identify a concordant group of schizophrenic patients—a group of patients who are diagnosed as schizophrenic in a standardized clinical fashion according to clinical assumptions, who are allocated to the diagnostic class of schizophrenia by the computer program, and who belong to clusters that statistically select out schizophrenic patients, regardless of clinical assumptions.

Psychopathology of Patient Groups at Initial Evaluation

In all, 1,202 patients were examined. Of these, 811 received a clinical diagnosis of schizophrenia; 164, affective psychoses; and 102, other psychoses. The remainder had other diagnoses. The distribution of patients by diagnostic group and by center is presented in table 2.

In comparing the psychopathology of patient groups, it was decided to carry out the comparisons, in terms of the rank order of frequency of symptoms, on the level of groups of units of analysis, and in terms of the symptoms most frequently present among patients in each diagnostic group, on the level of units of analysis. Kendell's tau rank correlation coefficient was used to calculate the degree of concordance of rank order of average percentage scores on groups of units between pairs of symptom profiles. In addition, symptom profiles were compared using analysis of variance (ANOVA). When there was a high degree of concordance between the rank orders of two centers, the symptom profiles were referred to as similar.

Clinical Diagnosis of Schizophrenia

When the average percentage scores on the 27 groups of units of analysis for all schizophrenic patients within each center were examined, it was apparent that the rank order of the groups of units was very similar across

Table 2. Patients included in IPSS, by diagnostic group and by field research center.¹

Diagnosis ²	Field research center									All centers
	Aarhus	Agra	Cali	Ibadan	London	Moscow	Taipei	Washington	Prague	
Schizophrenia										
Simple	6	4	2	7	2	—	1	4	5	31
Hebephrenic	12	3	20	9	9	—	30	—	3	86
Catatonic	2	22	13	10	3	—	3	1	—	54
Paranoid	28	15	20	49	75	13	36	51	36	323
Acute ³	—	10	29	5	1	11	4	15	4	79
Latent ⁴	3	—	3	—	—	14	—	2	2	25
Residual ⁵	—	—	6	—	1	—	—	6	2	15
Schizo-affective	1	17	7	26	8	5	8	15	20	107
Other specified ⁶	1	3	—	3	—	34	—	—	3	44
Unspecified	—	27	1	11	1	—	3	3	1	47
Total	53	101	101	120	100	77	86	97	76	811
Affective psychosis										
Agitated depression	1	3	1	—	—	—	—	—	—	5
Manic-depressive, depressed	19	5	1	8	4	10	3	4	19	73
Manic-depressive, manic	20	20	3	4	6	1	2	2	8	66
Others	4	—	—	1	4	2	5	2	2	20
Total	44	28	5	13	14	13	10	8	29	164
Paranoid states⁷	10	—	—	1	—	—	9	—	9	29
Other psychoses⁸										
Reactive depression	1	—	1	5	1	6	4	2	1	21
Others	17	—	2	2	1	6	17	3	4	52
Total	18	—	3	7	2	12	21	5	5	73
All psychoses	125	129	109	141	116	102	126	110	119	1,077
Neurosis, personality disorders										
Depressive neurosis	2	9	6	3	11	10	10	14	6	71
Others	2	2	12	1	—	28	1	8	—	54
Total	4	11	18	4	11	38	11	22	6	125
All patients	129	140	127	145	127	140	137	132	125	1,202

¹ Adapted with permission from p. 161 of World Health Organization (1973).² Diagnoses are based on the International Classification of Diseases (World Health Organization 1967). Special diagnostic terms not found in the International Classification, but used at some centers, have been assigned as shown in footnotes 3 through 8.³ Periodic schizophrenia.⁴ Sluggish schizophrenia.⁵ Chronic undifferentiated schizophrenia.⁶ Shiftlike schizophrenia.⁷ Acute paranoid psychosis.⁸ Psychogenic paranoid psychosis.

centers. The schizophrenic groups of all centers had high scores on lack of insight, predelusional signs (such as delusional mood, ideas of reference, and perplexity), flatness of affect, auditory hallucinations (except for the Washington center), and experiences of control. Center scores were also high on delusions, derealization, and disturbances of mood, although these were not so uniformly high as for the first-mentioned groups of symptoms. Scores were relatively low across centers in the areas of qualitative psychomotor disorder (negativism, compliance, mannerisms, and similar abnormal behavior), pseudohallucinations, and affective changes other than incongruous affect.

Of the 36 possible comparisons between schizophrenic profiles of pairs of centers, there was a significant degree of concordance of rank order of frequency of groups of units of analysis in 32. Concordance was not significant for the Agra-London, Agra-Taipei, London-Washington, and London-Prague pairs. The analyses indicated that, with these few exceptions, the psychopathological characteristics of schizophrenic patient groups were similar in the different centers when the basis for comparison was rank order of frequency of groups of units of analysis.

Comparison of the 15 most frequently positive units of analysis supported this conclusion, since there was a high degree of similarity among the centers with regard to the symptoms that occurred most frequently in their schizophrenic groups of patients. The most frequently positive units of analysis for all centers combined, in decreasing order of frequency, were lack of insight, inadequate description of problems, suspiciousness, unwillingness to cooperate, ideas of reference, flatness of affect, delusions of persecution, delusions of reference, delusional mood, poor rapport, presence of auditory hallucinations, presence of verbal hallucinations, voices speaking to the patient, thought alienation, and gloomy thoughts.

Paranoid Schizophrenia

A similar analysis was done for the paranoid schizophrenia subgroup in the study, since the subgroup was sufficiently large ($N = 323$) for such an analysis. The average percentage scores on groups of units indicated that these patients were mainly characterized by lack of insight, experiences of control, predelusional signs, delusions, and flatness of affect. There were also high

ratings on auditory and "characteristic" hallucinations (such as voices discussing the patient and hallucinations from the body), although the ratings were lower in Washington than in the other field research centers. All centers had low scores on psychomotor disorders and disorders of form of thinking; all except London and Moscow rated low on pseudohallucinations, and all but Washington had low scores on affective change other than incongruity of affect.

When the symptom profiles of the groups of paranoid schizophrenic patients were compared center by center, in terms of concordance of rank order of groups of units of analysis, the profiles of eight of the nine centers showed a significant degree of concordance with one another. The profile of the other center, Washington, had a significant degree of concordance with four of the other centers (Aarhus, Cali, Ibadon, and Moscow) but not with the remaining four (Agra, London, Taipei, and Prague).² It can be concluded that, with these few exceptions, the psychopathological characteristics of the groups of paranoid schizophrenic patients were similar in the different centers when the basis of comparison was rank order of frequency of symptoms.

This conclusion was supported by comparison among the centers of the most frequently positive units of analysis, which indicated that there was a high degree of similarity among the centers with regard to the symptoms occurring most frequently in their paranoid schizophrenic groups of patients. The most frequently positive units of analysis for all centers combined, in decreasing order of frequency, were lack of insight, suspiciousness, delusions of persecution, delusions of reference, ideas of reference, unwillingness to cooperate, inadequate description of problems, delusional mood, flatness of affect, and presence of auditory hallucinations.

Psychotic Depression

A similar analysis was done for the IPSS patients diagnosed as having psychotic depression. When the profiles of these patients, expressed in average percentage scores on the 27 groups of units, were examined, they were found to show a high degree of similarity.

²Patients at the Washington Center were not rated on certain items, and the absence of information on those items—rather than the absence of particular symptoms—might have contributed to this lack of concordance.

Positive scores were high across all centers in the groups of affect-laden thoughts, neurasthenic complaints, lack of insight, depressed mood, and psychophysiological complaints. On the other hand, they were generally low on hallucinations, pseudohallucinations, and incongruity of affect—groups in which positive scoring would suggest schizophrenia.

When the symptom profiles of the groups of psychotically depressed patients were compared center by center in the four centers (Aarhus, Ibadan, Moscow, and Prague) with more than 10 such patients, there was a high degree of concordance for all comparisons, indicating that the psychopathological characteristics of this diagnostic group were similar in these four different centers.

Analysis of the frequency of positive scores on the units for all psychotically depressed patients revealed that there was a great similarity among the centers. It also indicated that the most frequently positive units apparently coincided with the generally recognized symptoms of psychotic depression. The most frequently positive units of analysis for all centers combined, in decreasing order of frequency, were depressed mood, gloomy thoughts, hopelessness, early waking, feeling worse in the morning, sleep disturbances, delusions of self-depreciation, anxiety, lack of insight, retardation, lack of concentration, inadequate description of problems, decreased energy, diminished appetite and weight, delusions of guilt, and tension.

Similarity and Dissimilarity between Clinical Conditions

The units of analysis and groups of units were analyzed further to examine the similarity and dissimilarity between the clinical conditions of those patients diagnosed as schizophrenic and those diagnosed as psychotically depressed. When the profiles of groups of units were compared for the two groups (figure 1), it was noted that, although there were some areas of similarity, as would be expected since both groups were composed of psychotic patients, there were major differences. Experiences of control, which ranked 6th in the schizophrenic group, ranked 21st in the depressive group; auditory hallucinations ranked 4th and 18th, respectively, in the two groups, while incongruity of affect ranked 10th among schizophrenics and 25th

among depressives. Psychophysiological disorder ranked 5th among the depressed patients and 17th among the schizophrenics. The degree of concordance between profiles, in terms of rank order of frequency of groups of units, was low. These findings suggest that the differences between patients diagnosed as schizophrenic and patients diagnosed as psychotically depressed justify classifying them in different categories.

Analysis of those units showing a significant difference in frequency of positive ratings between schizophrenia and depressive psychosis in each center suggests that the symptoms that differentiate between the two conditions may vary from center to center.

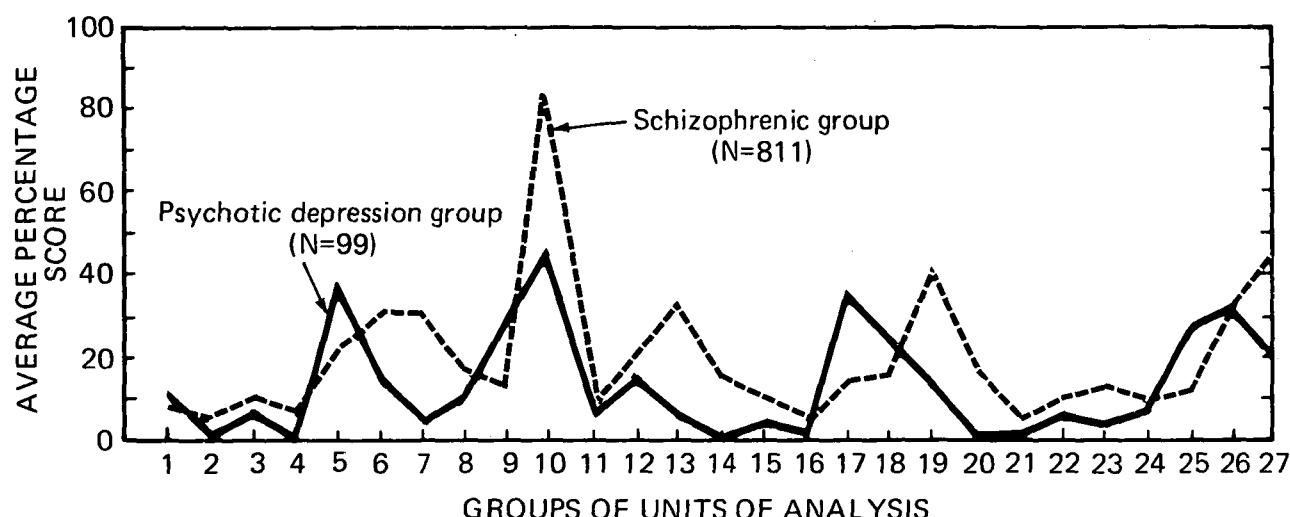
In addition to the comparisons of scores on groups of units among patient groups described above, analysis of variance and discriminant function analysis were performed, and preliminary results suggest similar conclusions.

Thus, when patients are grouped together according to clinical diagnosis, an analysis of psychopathology indicates that there is a high degree of similarity among the groups of schizophrenic patients in the different centers. However, until criteria specified in advance have actually been used to allocate patients successfully to the diagnostic groups used by the various clinicians taking part in the study, it cannot be claimed that their diagnostic rules have been fully examined. To further examine this question, two additional techniques were investigated. The first of these was a computer simulation of the diagnostic process and the second was a statistical clustering technique.

Computer Simulation of the Diagnostic Process

To standardize the classification of patients, the CATEGO computer program was developed (Wing, Cooper, and Sartorius 1974), which incorporated diagnostic rules and made possible a very reliable categorization. Designed to follow closely the diagnostic principles more or less commonly accepted in European psychiatry, the program involves a stepwise decision process in which different symptoms and syndromes receive a priori different diagnostic weights. For example, if certain key symptoms are present (e.g., hallucinatory voices discussing the patient in the third person, thought insertion or broadcast, and delusions of control—phenomena described by K. Schneider (1971)

Figure 1. Profiles of average percentage scores (groups of units of analysis) of schizophrenic patients and patients with psychotic depression.¹



- | | | |
|--|-------------------------------------|---|
| 1. Quantitative psychomotor disorder | 8. Delusions | 19. Flatness |
| 2. Qualitative psychomotor disorder | 9. Neurasthenic complaints | 20. Incongruity |
| 3. Quantitative disorder of form of thinking | 10. Lack of insight | 21. Other affective change |
| 4. Qualitative disorder of form of thinking | 11. Distortion of self-perception | 22. Indication of personality change |
| 5. Affect-laden thoughts | 12. Derealization | 23. Disregard for special norms |
| 6. Predelusional signs | 13. Auditory hallucinations | 24. Other behavioral change |
| 7. Experiences of control | 14. "Characteristic" hallucinations | 25. Psychophysical disorders |
| | 15. Other hallucinations | 26. Cooperation difficulties, circumstances-related |
| | 16. Pseudohallucinations | 27. Cooperation difficulties, patient-related |
| | 17. Depressed-related | |
| | 18. Anxiety, tension, irritability | |

¹ Adapted with permission from p. 404 of World Health Organization (1973).

as "first-rank" symptoms of schizophrenia),³ the program classifies this patient into class S (schizophrenic psychoses with delusions or auditory hallucinations of specified types). In the absence of the above-mentioned "key symptoms," a patient may also be assigned to class S if a pattern of other symptoms commonly encountered in schizophrenia is present. If class S symptoms are absent, the patient is assigned according to manifest symptomatology to one of the remaining eight classes (e.g., M = manic psychoses; D = depressive psychoses; and P = nonaffective paranoid psychoses) if schizo-

phrenic psychopathology can be excluded. In making the final classification, the program also takes into consideration data from the psychiatric history.

There was a high level of agreement (87 percent) between a diagnosis by a center psychiatrist of schizophrenia, mania, or depression, and the classification made by the computer. The results of the CATEGO classification indicated that there were schizophrenic patients with similar patterns of symptomatology in all nine centers. The first-rank symptoms seemed to be very highly discriminatory: if they were present, the chance that the clinical diagnosis would be schizophrenia was at least 95 percent. It also seems probable that, as far as the functional psychoses are concerned, the rules laid down

³ These symptoms seemed to be highly discriminatory: if they were present, the chance that the clinical diagnosis would be schizophrenia was at least 95 percent.

in the CATEGO program are similar to the rules that most of the clinicians in this study use in making a diagnosis.

Statistical Clustering Technique

A third method of classifying the IPSS patients was cluster analysis, which defines groups on purely empirical grounds, thus avoiding clinical assumptions and, as used in the IPSS, giving each piece of information the same weight. According to the underlying theory, the similar individuals in these groups will be the ones who have the greatest number of features in common and who are also most likely to be very similar in any other characteristics that might be evaluated.

Several methods of cluster analysis were applied to the IPSS data. McKeon's (1967) hierarchical clustering method grouped the 1,202 patients into 10 clusters which were then used in a number of further analyses. These analyses helped to identify particular groups of patients to which special emphasis will be given during the followup period.

Characteristics of a Concordant Group of Schizophrenics

Using three separate methods of classification—original clinical diagnosis, CATEGO diagnosis, and cluster analysis—it was possible to define a concordant group of 306 schizophrenics, which consisted of all patients whose original clinical diagnosis was schizophrenia, who fell into CATEGO class S, and who also belonged to one of those McKeon's clusters that contained a statistically significantly higher number of schizophrenics than would have been expected by chance (clusters 4, 5, and 7).

There are several important advantages to defining a concordant group that comprises patients who have been diagnosed in a standardized fashion according to clinical assumptions, and who also belong with clusters that statistically select out schizophrenic patients, regardless of clinical assumptions: First, it permits the identification and description of a group that excludes those patients who have been diagnosed as schizophrenic because of a lack of standardization of the diagnostic process, variation in clinical assumptions, or culture-bound factors. Second, it makes it possible to examine whether there are representatives of such a group in all centers. Third, it identifies a group of patients to whom

particular attention may be given during the followup phase of the study to determine whether their course of illness differs from that of other schizophrenic patients. If it can be shown that such a group of patients exists in all countries and has a specific clinical picture and differential course of illness, then the description of this group's characteristics may point the way toward a transculturally applicable definition of schizophrenia.

The psychopathological characteristics of the concordant group were examined in terms of the profiles of 27 groups of units and in terms of the frequency of units of analysis. Analysis of the 27 groups of units indicated that the most prominent symptom in this group was lack of insight. The other groups of units on which the concordant group had a high score were auditory hallucinations, flatness of affect, experiences (including delusions) of control, and pre-delusional signs. Poor rapport and other circumstances that might make it difficult to obtain information in the interview also have high scores.

When the frequency of positive ratings on units of analysis was examined, the following psychopathological characteristics were noted: 97 percent of the patients were characterized by lack of insight; 74 percent had auditory hallucinations; 70 percent, verbal hallucinations; 70 percent, ideas of reference; 67 percent, delusions of reference; 66 percent, suspiciousness; 66 percent, flatness of affect; 65 percent, voices speaking to the patient; 64 percent, delusional mood; 64 percent, delusions of persecution; 64 percent, inadequate description; 52 percent, thought alienation; and 50 percent, thoughts spoken aloud. There were no other symptoms that were present in 50 percent or more of the patients, although some symptoms were present almost as frequently—delusions of control, 48 percent; hearing voices speak full sentences, 44 percent; and poor rapport, 43 percent.

Comparison of the psychopathology of concordant schizophrenics and discrepant schizophrenics (patients who had a clinical diagnosis of schizophrenia, but who were neither in CATEGO class S nor in McKeon's clusters 4, 5, or 7) indicated that they differed markedly with regard to hallucinations, delusions, flatness of affect, and depressive symptomatology. The concordant schizophrenics scored much higher on delusions, hallucinations, and flatness of affect, while the discrepant schizophrenics scored higher on depressive symptomatology.

When the psychopathological characteristics of the concordant group were compared to those of the group of psychotically depressed patients, it was found that the concordant group of schizophrenics showed even less similarity to the psychotically depressed patients than did the group of all schizophrenics.

There were concordant schizophrenics in every one of the field research centers. When the average percentage scores on groups of units of analysis for concordant groups of individual centers were compared center by center, using an analysis of variance, there were no significant differences between any pair of centers. Thus it can be concluded that it is possible to identify a concordant group of schizophrenic patients that has a distinctive pattern of symptoms, that this pattern is consistent across centers, and that there are patients belonging to this group in every center in the study.

The Followup Studies

In the psychiatric literature, there is a scarcity of previous studies that might indicate how many psychotic patients in a variety of cultures can be reexamined a considerable length of time after an initial evaluation. There are very few truly cross-cultural studies of the course and outcome of schizophrenia and other functional psychoses, and those that exist have left many important methodological problems unsolved—a fact which limits considerably the value of their contribution.

The IPSS provided the opportunity to throw light on the feasibility of large-scale cross-cultural followup studies and to investigate prospectively, in a standardized and reliable way, the course and outcome of a great number of patients in different sociocultural environments. The patients included in the IPSS were followed up and examinations carried out 2 years and 5 years after the initial evaluation.

Followup instruments and procedures were tested in a 1-year followup of approximately 40 percent of the patients. The experiences gained during this trial were used to assess the feasibility of followups and to refine methodological procedures for the 2-year followup.

Apart from the PSE, the following instruments were used in the followup phase: the Followup Psychiatric History Schedule, the Followup Social Description Schedule, and the Followup Diagnostic Assessment Schedule. The information gathered with these four

instruments was supplemented in most centers by detailed narrative accounts of the clinical picture, course of illness, and social environment of the patient.

The followup phase included an assessment of the applicability and reliability of the followup instruments. As in the main phase of the study, some of the PSE interviews at the centers during the followup phase were simultaneous interviews, so that intracenter reliability could be determined. A number of multiple reliability exercises were done to determine intercenter reliability at the exchanges of visits by collaborating investigators.

The 2-year followup has been completed. It was possible to trace and re-interview a very high proportion of the patients: an average of 82.1 percent for all centers except London (where the limited staff resources during the followup phase did not permit the project psychiatrists to make home visits, although the great majority of the patients were traceable). In two of the centers the proportion of the re-interviewed patients was above 90 percent.

Only 2.9 percent of all patients could not be traced. Before a 2nd year examination could be carried out, 2.4 percent of all patients had died. In only 6.9 percent of the 1,202 patients was a followup interview refused by either the patient or his family. An additional 12.1 percent of the patients (the majority in this group were London patients) were not re-interviewed but could have been examined had more resources been available. Prominent among the difficulties encountered by the centers in attempting to trace and reexamine patients were the lack of population registers in some centers, large size of the catchment area (e.g., 52,076 sq km in Agra), changes in street names or numeration, and seasonal weather variations. The problems specific to each center will be described in detail in volume 2 of the report on the IPSS (World Health Organization, in press).

On the whole, the findings confirm the feasibility of studies of this type. Particularly striking is the fact that 97.1 percent of all the 1,202 patients in the nine centers could be traced for the 2-year followup study.

It should be emphasized, however, that to make the followup feasible, a considerable amount of work had to be carried out by the staff at the field research centers and the success of the followup could not have been attained without the enthusiasm of all of the staff in the study. The average time spent assessing a patient with the four major followup instruments was more than 2

hours. Over 2,100 hours were spent assessing IPSS patients with the 2-year followup schedules, and 4,300 man hours were spent carrying out home visits, 1,200 of these by psychiatrists. The number of home visits required to gather followup information about patients differed markedly from center to center. For example, Aarhus and Cali were quite similar with regard to the number of patients about whom sufficient information was gathered to be included in the followup group. In Cali, however, 81 home visits were required compared to 29 in Aarhus. Ibadan was the center which carried out the greatest number of home visits (104) to gather information.

One of the aims of the followup studies was to trace similarities and differences in the course and outcome of illness between patients in different centers, as well as between patients belonging to different diagnostic groups.

The data from the 2-year followup are being used for different analyses now in progress, which concentrate on five major areas:

- Comparisons of the symptomatology of the patients at the time of the 2nd-year followup (within and between diagnostic groups and within and between centers).
- Comparisons of the symptomatology of the patients at the time of the initial evaluation with the symptomatology of the same patients at the time of the 2nd-year followup.
- Assessment of the type of course of the disorder during the interval between the initial evaluation and the 2nd-year followup and at the end of the followup period.

In the above three analyses, particular emphasis is given to the concordant group of schizophrenics as compared to the discordant group and the other diagnostic groups.

- Analyses directed to the assessment of the predictive value of different characteristics of the patient (clinical and sociodemographic) for the 2-year course and outcome—these characteristics include factors related to premorbid personality, past history, initial symptomatology, evolution of episode of inclusion, and sociodemographic status.

- Analyses representing approaches toward the assessment of the validity of the three classification systems used in the study (clinical diagnosis, CATEGO, and clustering).

The 5-year followup is now in progress, and already more than half of the patients have been found and reexamined by psychiatrists.

Conclusions

The initial evaluation phase of the IPSS has demonstrated that it is possible, using standardized reliable methods of assessment, to identify schizophrenic patients in centers in nine countries of the world who are similar with regard to their clinical picture at the time of a psychotic episode. Furthermore, it demonstrated that the symptomatology of these patients was different from the symptomatology of patients with other functional psychoses. These findings have important implications for many types of research into schizophrenia, since they indicate that genetic, epidemiological, psychopharmacological, and other studies of schizophrenia can be carried out in a fashion that will allow comparability of results. The demonstration that it is possible to identify symptomatologically similar schizophrenic patients in a reliable fashion is an important step forward in attempts to understand more about the nature of schizophrenia.

There have been many previous followup studies of schizophrenia. The followup phase of the IPSS is unique, however, in allowing a description of the course and outcome of a large number of schizophrenic patients who were prospectively selected in several different parts of the world and reassessed and evaluated in standardized fashion. The analyses of the data of this study will thus allow previous hypotheses about the nature of the course and outcome of schizophrenia and about the factors which influence such course and outcome to be reconsidered in a new light.

In summary, the IPSS has developed standardized, reliable, and internationally applicable instruments for psychiatric assessment; it has demonstrated the feasibility of large-scale international transcultural psychiatric studies; it has provided managerial and operational methodology for carrying out such studies; it has provided methods of analyzing large amounts of data and procedures for training investigators in the use of

standard techniques; and it has enhanced basic knowledge about the nature of schizophrenia and other functional psychoses. The large patient sample that has been assessed in a standardized fashion in nine countries can serve as a reference group for future studies.

The IPSS has also resulted in the creation of a network of research centers in economically and socio-culturally very different countries. In these centers there are now research workers trained and experienced in transcultural psychiatry who have established working relationships with one another and who have indicated a desire to continue their study of transcultural and psychiatric problems.

References

Cooper, J. E.; Kendell, R. E.; Gurland, B. J.; Sharpe, L.; Copeland, J. R. M.; and Simon, R. *Psychiatric Diagnosis in New York and London: A Comparative Study of Mental Hospital Admissions*. New York: Oxford University Press, 1972.

McKeon, J. J. *Hierarchical Cluster Analysis*. George Washington University Biometric Laboratory, Washington, D.C., 1967.

Schneider, K. *Klinische Psychopathologie*. 9th ed. Stuttgart: Georg Thieme Verlag, 1971.

Wing, J. K.; Cooper, J. E.; and Sartorius, N. *The Measurement and Classification of Psychiatric Symptoms*. New York: Cambridge University Press, 1974.

World Health Organization. *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death*. Vol. I. 8th ed. Geneva: World Health Organization, 1967.

World Health Organization. *Report of the International Pilot Study of Schizophrenia*. Vol. I. Geneva: World Health Organization, 1973.

World Health Organization. *Report of the International Pilot Study of Schizophrenia*. Vol. II. Geneva: World Health Organization, in preparation.

Acknowledgment

This report is based on the data and experience obtained during the participation of the authors in the International Pilot Study of Schizophrenia (IPSS), a project sponsored by the World Health Organization, and funded by the World Health Organization, the U.S. National Institute of Mental Health, and the participating field research centers.

The collaborating investigators on this study have been the author, Dr. N. Sartorius, principal investigator, Dr. T.-Y. Lin, former principal investigator, and Ms. E. M. Brooke, Dr. F. Engelsmann, Dr. G. Ginsburg, Mr. M. Kimura, Dr. A. Richman, and Dr. R. Shapiro at the headquarters of the World Health Organization in Geneva, Switzerland; Dr. E. Strömberg, chief collaborating investigator, and Drs. A. Bertelsen, N. Engkilde, M. Fischer, C. Flack, and N. Juel-Nielsen at the field research center in Aarhus, Denmark; Dr. K. C. Dube, chief collaborating investigator, and Dr. B. S. Yadav at the field research center in Agra, India; Dr. C. Leon, chief collaborating investigator, and Drs. G. Calderon and E. Zambrano at the field research center in Cali, Colombia; Dr. T. A. Lambo, chief collaborating investigator, and Dr. T. Asuni at the field research center in Ibadan, Nigeria; Dr. J. K. Wing, chief collaborating investigator, and Drs. J. Birley and J. P. Leff at the field research center in London, England; Dr. R. A. Nadzharov, chief collaborating investigator, and Dr. N. M. Zharikov at the field research center in Moscow, U.S.S.R.; Dr. L. Hanzlíček, chief collaborating investigator, and Dr. C. Škoda at the field research center in Prague, Czechoslovakia; Dr. C. C. Chen at the field research center in Taipei, China; and Drs. L. Wynne and J. Strauss, chief collaborating investigators, and Drs. J. Bartko and W. Carpenter at the field research center in Washington, D.C.

A list of other staff contributing to the IPSS can be found in World Health Organization (1973).

dean research award

Dr. Arvid Carlsson, Professor and Chairman of the Department of Pharmacology at Sweden's University of Göteborg, was recently named recipient of the Stanley R. Dean Award for basic research in schizophrenia. This prestigious \$2,500 award has been given annually since 1962 to behavioral scientists who have made outstanding contributions to the understanding of schizophrenia. Accepting the award at the 12th Annual Meeting of the American College of Psychiatrists held February 5-9 in Phoenix, Ariz., Dr. Carlsson summarized his important research contributions in a paper entitled, "Pharmacological Approach to Schizophrenia."