A Selective Review of Recent North American Long-Term Followup Studies of Schizophrenia

by Thomas H. McGlashan

Abstract

North American outcome studies of schizophrenia conducted within the past quarter century are reviewed if their minimum average followup is 10 years and they meet at least some modern design criteria. Ten such investigations are described and summarized. Taken as a whole, they demonstrate that schizophrenia can be a chronic disease whose outcome on the average is worse than that of other major mental illnesses. It is associated with an increased risk for suicide, physical illness, and mortality. The schizophrenic process, however, is not relentlessly progressive, as originally described, but appears to plateau after 5–10 years of manifest illness. Overall, outcome is heterogeneous, but much of the variance can be linked to sample characteristics, including expressions of psychopathology (broad vs. narrow diagnostic criteria, subtypes, and comorbidity), dimensions of chronicity (length of manifest illness, treatment resistance, age of onset, and institutionalization), and other predictor variables (gender, marital status, socioeconomic status, physical setting, and premorbid health). Long-term followup studies have yet to demonstrate clearly any effect of treatment on the natural history of schizophrenia. Finally, these studies support a broad definition of schizophrenia.

Followup Study Selection Criteria

The following criteria were used to select the followup studies for review. The studies are North American. They are recent, meaning within the past 25 years, when major methodological advances began to be introduced. They are followup studies of schizophrenia, especially schizophrenia by operationalized diagnostic criteria. Not included are followup studies of atypical schizophrenia (by Feighner et al. 1972 criteria), brief reactive psychosis or schizophreniform psychosis (by DSM-III criteria; American Psychiatric Association 1980), or schizoaffective disorder (by any criteria). Included are followup studies beginning with an identified schizophrenic patient. Not included are the so-called “follow-back” studies which attempt to identify premorbid clues marking a child who is vulnerable to developing schizophrenia. “Long-term” is defined as a minimum average followup of 10 years. Studies are included if they introduced any or incorporated many of the methodological advances elaborated in the article on followup design in this issue (McGlashan et al. 1988). Finally, studies are selected that focus on the long-term course and outcome of the schizophrenic proband, excluded are studies that focus on their families and first-degree relatives.

Ten studies or groups of studies met these criteria. These studies are summarized in table 1 and described in the text.

Reprint requests should be sent to Dr T H. McGlashan, Chestnut Lodge Research Institute, 500 W. Montgomery Ave., Rockville, MD 20850.
### Table 1. Selective review of North American followup studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Investigators</th>
<th>Schizophrenic n at baseline/ followup</th>
<th>Followup average length/-range</th>
<th>Sample drawn from</th>
<th>Diagnostic criteria</th>
<th>Comparison groups</th>
<th>Data source(s)</th>
<th>Followup</th>
</tr>
</thead>
<tbody>
<tr>
<td>Massachusetts Mental Health Center 1 (MMHC)</td>
<td>Vaillant (1964)</td>
<td>72%—2 yrs</td>
<td>No average</td>
<td>2–15 yrs</td>
<td>Schizophrenic patients admitted to MMHC 1947–1950</td>
<td>Clinical diagnosis based on Bleulerian criteria</td>
<td>None</td>
<td>Hospital records</td>
</tr>
<tr>
<td>Massachusetts Mental Health Center 2</td>
<td>Vaillant (1978)</td>
<td>66%—12 yrs</td>
<td>2–15 yrs</td>
<td>2–15 yrs</td>
<td>Schizophrenic patients admitted to MMHC and remitted between 1959 &amp; 1962</td>
<td>Clinical diagnosis based on Bleulerian criteria</td>
<td>None</td>
<td>Hospital records</td>
</tr>
<tr>
<td>Phipps Clinic</td>
<td>Stephens et al (1963)</td>
<td>56%—10 yrs</td>
<td>2–15 yrs</td>
<td>2–15 yrs</td>
<td>Schizophrenic patients admitted to Phipps Clinic between 1946 &amp; 1959</td>
<td>Clinical diagnosis based on Bleulerian criteria</td>
<td>None</td>
<td>Hospital records</td>
</tr>
<tr>
<td>Iowa 500</td>
<td>Teasdale et al (1979)</td>
<td>47%—10 yrs</td>
<td>4–16 yrs</td>
<td>4–16 yrs</td>
<td>Schizophrenic patients admitted to Iowa Psychopathic Hosp between 1934 &amp; 1944</td>
<td>Clinical diagnosis based on Bleulerian criteria</td>
<td>None</td>
<td>Hospital records</td>
</tr>
<tr>
<td>Alberta 1</td>
<td>Bland et al (1976)</td>
<td>67%—10 yrs</td>
<td>10 yrs</td>
<td>10 yrs</td>
<td>Schizophrenic patients admitted to MMHC and remitted between 1959 &amp; 1962</td>
<td>Clinical diagnosis based on Bleulerian criteria</td>
<td>None</td>
<td>Hospital records</td>
</tr>
<tr>
<td>Alberta 2</td>
<td>Bland et al (1976)</td>
<td>45%—10 yrs</td>
<td>10 yrs</td>
<td>10 yrs</td>
<td>Schizophrenic patients admitted to Alberta Schizophrenia Index.</td>
<td>Clinical diagnosis based on Bleulerian criteria</td>
<td>None</td>
<td>Hospital records</td>
</tr>
<tr>
<td>New York City Outpatient Clinic</td>
<td>Engelhardt et al (1962)</td>
<td>51%—10 yrs</td>
<td>12 yrs</td>
<td>12 yrs</td>
<td>Schizophrenic patients admitted to outpatient research clinic between 1958 &amp; 1962</td>
<td>Clinical diagnosis based on Bleulerian criteria</td>
<td>None</td>
<td>Hospital &amp; study records</td>
</tr>
<tr>
<td>Boston State Hospital</td>
<td>Gardner et al (1961)</td>
<td>60%—10 yrs</td>
<td>12 yrs</td>
<td>12 yrs</td>
<td>Schizophrenic patients admitted to hospital drug study in 1965</td>
<td>Clinical diagnosis based on Bleulerian criteria</td>
<td>None</td>
<td>Hospital &amp; study records</td>
</tr>
</tbody>
</table>

See footnotes at end of table.
### Sample Distinctions

<table>
<thead>
<tr>
<th>Outcome or natural history for schizophrenia</th>
</tr>
</thead>
</table>
| Poor  | 1 = 26%  
|       | (see text for scale details)  
| Good  | 2 = 6%  
|       | 3 = 25%  
|       | 4 = 15%  
| Remitted  | 5 = 20%  
| Remitted  | 61%  
| Chronic  | 59%  

### First Admissions

<table>
<thead>
<tr>
<th>First admissions</th>
<th>Recovered</th>
<th>Improved</th>
<th>Unimproved</th>
</tr>
</thead>
<tbody>
<tr>
<td>24%</td>
<td>46%</td>
<td>30%</td>
<td></td>
</tr>
</tbody>
</table>

### Remitted patterns

<table>
<thead>
<tr>
<th>Remitted patterns</th>
<th>28%</th>
<th>6%</th>
<th>25%</th>
<th>15%</th>
<th>26%</th>
<th>61%</th>
<th>59%</th>
</tr>
</thead>
</table>

### Schizophrenia worst

<table>
<thead>
<tr>
<th>Schizophrenia worst</th>
<th>Marks &amp; depression better</th>
<th>Surgical control best</th>
</tr>
</thead>
<tbody>
<tr>
<td>80%</td>
<td>10%</td>
<td></td>
</tr>
</tbody>
</table>

### First admissions

<table>
<thead>
<tr>
<th>First admissions</th>
<th>Recovered</th>
<th>Periodic mild</th>
<th>Periodic severe</th>
<th>Chronic mild</th>
<th>Chronic severe deficit</th>
<th>Chronic hospitalized</th>
<th>Suicidal</th>
</tr>
</thead>
<tbody>
<tr>
<td>58%</td>
<td>9%</td>
<td>9%</td>
<td>7%</td>
<td>9%</td>
<td>8%</td>
<td>2%</td>
<td>3%</td>
</tr>
</tbody>
</table>

### First admissions

<table>
<thead>
<tr>
<th>First admissions</th>
<th>Recovered</th>
<th>Periodic mild</th>
<th>Periodic severe</th>
<th>Chronic mild</th>
<th>Chronic severe deficit</th>
<th>Chronic hospitalized</th>
<th>Suicidal</th>
</tr>
</thead>
<tbody>
<tr>
<td>21%</td>
<td>30%</td>
<td>21%</td>
<td>12%</td>
<td>14%</td>
<td>2%</td>
<td>12%</td>
<td>2%</td>
</tr>
</tbody>
</table>

### First admissions

<table>
<thead>
<tr>
<th>First admissions</th>
<th>Recovered</th>
<th>Periodic mild</th>
<th>Periodic severe</th>
<th>Chronic mild</th>
<th>Chronic severe deficit</th>
<th>Chronic hospitalized</th>
<th>Suicidal</th>
</tr>
</thead>
<tbody>
<tr>
<td>21%</td>
<td>30%</td>
<td>21%</td>
<td>12%</td>
<td>14%</td>
<td>2%</td>
<td>12%</td>
<td>2%</td>
</tr>
</tbody>
</table>

### First admissions

<table>
<thead>
<tr>
<th>First admissions</th>
<th>Recovered</th>
<th>Periodic mild</th>
<th>Periodic severe</th>
<th>Chronic mild</th>
<th>Chronic severe deficit</th>
<th>Chronic hospitalized</th>
<th>Suicidal</th>
</tr>
</thead>
<tbody>
<tr>
<td>21%</td>
<td>30%</td>
<td>21%</td>
<td>12%</td>
<td>14%</td>
<td>2%</td>
<td>12%</td>
<td>2%</td>
</tr>
</tbody>
</table>

### Outpatients, largely institutionally chronic

<table>
<thead>
<tr>
<th>Outpatients, largely institutionally chronic</th>
<th>Eventually hospitalized</th>
<th>Number hospitalizations</th>
<th>Total length of hospitalization per patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>63%</td>
<td>2.7</td>
<td>44 mos</td>
<td></td>
</tr>
</tbody>
</table>

### Structured living situation

<table>
<thead>
<tr>
<th>Structured living situation</th>
<th>Nonsheltered employment</th>
<th>GAS Score</th>
<th>IMPA global severity score (markedly ill)</th>
</tr>
</thead>
<tbody>
<tr>
<td>97%</td>
<td>6%</td>
<td>40</td>
<td>4.8</td>
</tr>
</tbody>
</table>

(Continued on next page)
Table 1. Selective review of North American followup studies—Continued

<table>
<thead>
<tr>
<th>Study</th>
<th>Investigators</th>
<th>Schizophrenic sample at baseline/ followup</th>
<th>Followup average length/ range</th>
<th>Sample drawn from</th>
<th>Diagnostic criteria</th>
<th>Comparison groups</th>
<th>Data source(s)</th>
<th>Baseline</th>
<th>Followup</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chestnut Lodge</td>
<td>McGlashan (1984a, 1984b)</td>
<td>188/163</td>
<td>15 yrs 2-32 yrs</td>
<td>Patients discharged from the hospital between 1950 &amp; 1975</td>
<td>Operational criteria, several systems including DSM-III</td>
<td>Bi (n = 33)</td>
<td>UNI (n = 58)</td>
<td>Hospital record abstracts</td>
<td>Personal interview</td>
</tr>
<tr>
<td>Washington International Pilot Study of Schizophrenia</td>
<td>Carpenter et al (1987)</td>
<td>7/40</td>
<td>11 yrs (no range)</td>
<td>Patients admitted to psychiatric units in general hospitals, Prince Georges County, MD, in 1968-69</td>
<td>Clinical diagnoses based on DSM-II &amp; ICD-9</td>
<td>Non-schizophrenic patients (n = 13)</td>
<td>Hospital records</td>
<td>Structured personal interview</td>
<td></td>
</tr>
<tr>
<td>Columbia Psychiatric Institute</td>
<td>Stone (1986)</td>
<td>99/94</td>
<td>No average 10-23 yrs</td>
<td>Patients admitted to a long-term (avg 1yr) academic unit in NYC between 1963 &amp; 1976</td>
<td>DSM-III operational criteria</td>
<td>SA (n = 64)</td>
<td>M-D (n = 39)</td>
<td>Hospital records</td>
<td>Structured personal interview</td>
</tr>
</tbody>
</table>

Abbreviations: BL = bipolar affective disorder, UNI = unipolar affective disorder, SA = schizoaffective psychosis, BPD = borderline personality disorder, SPD = schizotypal personality disorder, M-D = manic-depressive psychosis, GAS = Global Assessment Scale, SES = socioeconomic status, IMPS = Inpatient Multidimensional Psychiatric Scale

1 = independence of assessment, 2a = data, missing information, 2b = data, missing subjects, 3a = diagnosis, operationalized criteria, 3b = diagnosis, multiple systems, 3c = diagnosis, comorbidity, 3d = diagnosis, reliability, 4a = sample description, demographics, 4b = sample description, predictors, 4c = sample description, reliability, 5a = outcome, multidimensional, 5b = outcome, veracity check, 5c = outcome, operationalized scales, 5d = outcome, case calibration, 5e = outcome, reliability
## Description of the Followup Studies

### Massachusetts Mental Health Center Followup Studies

Vaillant conducted two long-term followup studies out of the Massachusetts Mental Health Center, a State-supported university teaching hospital in Boston. The first (Vaillant 1964b), from the 1960’s, was retrospective in design. The hospital records of 72 schizophrenic patients consecutively admitted between 1947 and 1950 were assessed in 1962. The patients were given a clinical diagnosis based on Bleulerian concepts. They were included in the study if they retained this diagnosis during their entire hospital stay. The charts were rated by the author for the presence/absence of seven prognostic criteria. Items not mentioned in the chart were scored as absent.

Six to 12 months later, the author followed up the cohort. First he consulted the central files of the Massachusetts Department of Mental Health to determine whether any of the patients had required further hospitalization. Subsequent charts were then gathered for clinical information. Efforts were made to contact next of kin on those patients for whom definitive followup information could not be obtained from subsequent hospital admissions. For this, telephone contacts or personal interviews were used. Outcome was specified on the dimensions: hospitalization,
work, and independence. It was rated on a 5-point scale as follows: 1 = 10 or more years residence in a mental hospital; 2 = 3 or more years in a hospital and inability to work or manage a household more than 50 percent of the time spent in the community; 3 = less than 3 years hospitalized with inability to work more than 50 percent of the time spent in the community; 4 = ability to work more than 50 percent of the time, but unable to live independently of parents or siblings; 5 = able to work full time or to manage a household effectively and to live independently of family of origin. Patients in followup categories 4 and 5 were considered long-term social remissions.

Of the original 72 patients, 97 percent could be followed reliably for at least 2 years, 89 percent for at least 8 years, and 63 percent for a full 12–15 years after admission. The average followup period was not reported. The outcomes reported were as follows: 1—28 percent, 2—6 percent, 3—25 percent, 4—15 percent, 5—26 percent. Collapsing categories, 41 percent could be classed as social remissions (4 and 5), and 59 percent were “essentially invalid” (1 + 2 + 3). Outcome correlated with predicted prognosis based on the seven criteria. Finally, the author noted that reliable long-term followup proved more difficult among patients who appeared to do well. All patients with less than 8 years of followup fell into groups 4 and 5. This was perhaps the first suggestion that missing subjects in followup studies constituted a healthier group of patients.

The second followup study (Vaillant 1978) from the 1970’s was basically prospective in design. Between 1970 and 1975, Vaillant obtained followup information on 51 of 56 schizophrenic patients from the Massachusetts Mental Health Center who had achieved complete remission during 1959–62. Remission was judged either retrospectively by the patients’ doctors at the time (Vaillant 1962) or was ascertained prospectively in a short-term followup study conducted at that time (Vaillant 1964a). The diagnosis of schizophrenia was clinical and based on Bleuler’s primary symptoms (Bleuler 1911/1950). The patients had to be hospitalized at least 2 months and receive a discharge diagnosis of schizophrenia. Records were scored for prognostic factors as in the earlier followup study. Patients were followed up through hospital records or interview 4–16 years later (average 10 years).

Presence or absence of remission constituted the outcome measure. Remission was a multidimensional concept, and its rating required that five conditions be met: (1) the patient had been free from psychotic symptoms and (2) from Bleuler’s primary symptoms for 1 year, (3) the patient had reattained his/her best level of premorbid adjustment, (4) was not on phenothiazines, and (5) had at least one friend. According to these criteria, 61 percent of the patients sustained their remissions while 39 percent did not and went on to follow a “chronic” course. The latter was not characterized by number and length of subsequent hospitalizations. Scores on the prognostic variables did not discriminate outcome groups, leading Vaillant to conclude that remitted schizophrenia is not necessarily a discrete subtype requiring a fresh label or nosologic reassignment to the affective disorders.

Phipps Clinic Followup Study.

Stephens and coworkers conducted a large retrospective followup study of 472 schizophrenic patients discharged from the Phipps Clinic in Baltimore between 1948 and 1959 (Stephens and Astrup 1963, 1965; Stephens 1970, 1978). Patients were first admissions, hospitalized for at least 3 weeks, and received a discharge clinical diagnosis of schizophrenia (criteria not specified). The charts were independently classified according to specified process and nonprocess categories (Stephens and Astrup 1963), and later the charts were also independently scored for the presence or absence of 43 prognostic variables (Stephens et al. 1966, 1967).

Followups of from 5 to 16 years’ duration were obtained on 78 percent of the sample based on “letters, telephone conversations, and personal contacts with the patients, their relatives and hospitals” (Stephens and Astrup 1965). Missing subjects did not differ from included subjects vis-a-vis discharge status or process/reactive classification. Outcome was described in three categories: (1) “Recovered” meant complete recovery without evidence of residual pathology. It excluded patients with frequent exacerbations and remissions. (2) “Improved” included patients who may have appeared recovered at followup but who had repeated exacerbations and hospitalizations, as well as patients with residual symptoms. (3) “Unimproved” referred to active, chronic psychosis. Most of these patients remained hospitalized for most of the followup period.

Of the total followup cohort, 143 patients were assessed 10 years or more after baseline (average 12
years, range 10–16 years). Their outcomes were distributed as follows: recovered—24 percent, improved—46 percent, unimproved—30 percent. Outcome clearly correlated with the process/nonprocess category and with the prognostic variables in the directions expected.

The Iowa 500 Followup Study. The Iowa 500 followup study has been massively innovative as well as massive in scope (Morrison et al. 1972; Clancy et al. 1974; Tsuang and Winokur 1975; Tsuang et al. 1979; Tsuang et al. 1980). It introduced specified diagnostic criteria for the selection of subjects, psychiatric and nonpsychiatric comparison groups, detailed description of samples, and independence of diagnostic and outcome assessments, among other innovations.

The study was retrospective and based on patients admitted between 1934 and 1944 to the Iowa State Psychiatric Hospital, a 60-bed, short-term treatment facility serving the entire State of Iowa. Roughly 370 patients were admitted each year, and over 3,800 records accumulated. Among these patients, 13 percent received a chart diagnosis of schizophrenia, and 19 percent received a chart diagnosis of manic-depressive psychosis or involutional melancholia. Their records (874 in all) were selected for evaluation. While no missing data rate was reported, the records sound very complete as described (Morrison et al. 1972). Records were screened diagnostically using the Feighner criteria for schizophrenia and affective disorders (Feighner et al. 1972). These criteria were purposefully stringent to reduce heterogeneity—especially for schizophrenia. One-fourth of the affective-disordered patients' records and two-thirds of the schizophrenic patients' records were excluded. Schizophrenic patients' charts (874 in all) were excluded mostly for episodic course and/or short duration. This process resulted in the selection of three samples—200 patients with schizophrenia, 100 patients with mania, and 225 patients with unipolar depression. To this the investigators added a nonpsychiatric comparison group (n = 160) drawn from surgical patients (appendectomy and herniorrhaphy) hospitalized at the University of Iowa over the same time period. The study also evaluated the first-degree relatives of study patients (n = 2,055), but this will not be elaborated here.

Followup data were collected between 1972 and 1976, roughly 30–40 years after index admission. The mean age of the schizophrenic patients at followup was 64 years. Fieldwork was conducted blind to the original study diagnoses by trained personnel: psychiatric residents, medical students, and "other personnel" (Clancy et al. 1974). Patients alive and willing were given a face-to-face or telephone structured interview (Tsuang et al. 1980) designed to evaluate normality, psychopathology, and multiple domains of functioning with tested reliability. Outcome for deceased subjects was approximated by giving this structured interview to a first-degree relative. Other sources of information came from State hospital records, death certificates, and sundry medical records.

The team was able to trace and rate 95 percent of the original schizophrenic cohort. Followup diagnoses were assigned by consensus after review of the structured interview by up to four psychiatrists (Tsuang 1978a). Functional outcome was operationalized according to four dimensions: marital, residential, occupational, and symptomatic; each was rated on a 3-point scale: poor, fair, or good. Missing information was incorporated into the outcome ratings in specified ways (Tsuang and Winokur 1975; Tsuang et al. 1979).

This study sample of 200 Feighner criteria schizophrenic patients had the following characteristics: male—52 percent, married—20 percent, poor premorbid adjustment and work—50 percent, high school graduate—28 percent, precipitating factors—11 percent, age of onset (median)—26 years, age at admission (median)—27 years, ill more than 1 year before admission—85 percent. The population was predominantly rural and relatively nonmobile since they were hospitalized during the Depression years. They were also hospitalized in the era before electroconvulsive therapy, drugs, or outpatient treatment networks. Return to the community, therefore, depended entirely on improvement in mental status. Of the 200 schizophrenic patients from this sample, 25 percent were discharged to the community following index hospitalization. This sample comes the closest of all the North American followup studies to approximating natural history of schizophrenia unperturbed by modern treatments.

Outcome was dichotomized into good and fair/poor for each dimension. Results demonstrated unequivocally—on all dimensions—that schizophrenic patients had the poorest outcome. Surgical control patients did the best, and affectively disordered patients were in
between. The best scores among the schizophrenic cohort on the four dimensions of outcome were distributed as follows: (1) marital status—21 percent married; (2) residential status—34 percent at home or relatives’ residence; (3) occupational status—35 percent employed, retired, housewife, or student; and (4) symptomatic status—20 percent no symptoms.

Alberta Followup Studies. Bland and coworkers conducted two followup studies of schizophrenic patients in the 1970’s—the first of a broadly defined, first admission cohort (Bland et al. 1976) and the second of a “narrowly” defined subsample (Bland and Orn 1978). The samples were selected from patients hospitalized in 1963 at Alberta Hospital, an inpatient psychiatric unit serving the southern half of the Province of Alberta. Since this was virtually the only such facility in the area at the time, the admission rate approximated the incidence of schizophrenia over the entire Provincial population served.

In the first study (1976) the files of newly admitted patients with a chart diagnosis of schizophrenia were examined. Selected were 92 cases with adequate data for diagnosis that met DSM-II (American Psychiatric Association 1968) criteria for schizophrenia as agreed upon by two psychiatrists. Followup was conducted by interviewing patients and significant others (family and/or professional) in 1974 and 1975, 10 years after admission. It was successful in 88 (96 percent) of the cases. Deceased patients were included. Outcome was multidimensional and assessed social adjustment, marital stability, work productivity, institutional treatment, subsequent care/medicines, and a global estimate of psychiatric condition. The latter consisted of six levels (1) recovered, no social or intellectual deficit; (2) periodic mild social and/or intellectual deficit; (3) periodic severe social and/or intellectual deficit; (4) mild chronic social and/or intellectual deficit; (5) chronic severe social and/or intellectual deficit; (6) chronic, unremitting institutionalization.

Of the 88 patients followed up, 48 were male. Their average age at index admission (which was their first psychiatric hospital admission) was 34 years (range 14–66 years). Fifty-one percent of the cohort was single. Outcome proved to be better than the authors expected. The global rating broke down as follows: 1—58 percent, 2—9 percent, 3—9 percent, 4–7 percent, 5—9 percent, 6—8 percent. Regarding further care, the authors noted that of the 51 recovered patients, 45 percent had discontinued their medications within 10 months of discharge; only 39 percent continued medications for 5 years or more.

The second followup study (1978) began with the same baseline cohort from the first study. This time, however, the sample was constructed by the application of “stricter” criteria to the charts. Nearly one-half of the original sample was excluded for psychotic affective disorder, organic brain syndrome, alcoholism, mental deficiency, or another diagnosis sufficient to explain the clinical picture. Five additional subjects were dropped because it was learned that they were not first admissions. Included were 45 subjects meeting one or more of the following operationalized diagnostic criteria: (1) Schneider’s first rank symptoms (Schneider 1959) minus somatic passivity; (2) the New Haven Schizophrenia Index (Astrachan et al. 1972) using the recommended cutoff score; and (3) the Feighner et al. (1972) criteria, both definite and “probable” cases. The latter patients were defined as those meeting all criteria except the duration item which requires a chronic illness with at least 6 months of symptoms prior to the index evaluation, without return to the premorbid level of psychosocial adjustment. Followup was conducted as in the first study, but in 1977, 14 years after admission.

Forty-three patients were traced. Deceased subjects were included. Outcome was assessed using the same multidimensional ratings.

The final followup sample of 43 broke down by diagnostic criteria as follows: (1) first rank symptoms: positive in 38 subjects; (2) New Haven Schizophrenia Index: met by 42 subjects; (3) Feighner et al. criteria: 20 probable and 20 definite cases. Demographically the sample consisted of 22 males, and 53 percent of the sample was single at admission. The mean age at index admission (which was their first psychiatric hospitalization) was 33 years. While the sample was undoubtedly defined by more explicit criteria than the 1976 sample, these criteria were not necessarily more narrow as the authors contend. Schneider’s first rank symptoms and the New Haven Schizophrenia Index are often broadly encompassing systems (see section on nosology), and less than half of the sample met the strictly defined Feighner criteria. The resultant sample was more like the 1976 sample than different—largely acute, first admission schizophrenic patients.

Outcome, while not as good as in the 1976 sample, was not much worse. Global ratings broke down
as follows: 1—21 percent, 2—30 percent, 3—21 percent, 4—12 percent, 5—14 percent, 6—2 percent. Seventy-nine percent of patients had an average of 2.7 further hospitalizations averaging 15 percent of the followup period. Fifty-one percent of the sample remained on neuroleptic medications for a mean of 7 years.

New York City Outpatient Clinic Followup Study. Engelhardt et al. (1982) reported a 15-year followup of schizophrenic patients who had participated in a Brooklyn-based outpatient clinic drug study of maintenance phenothiazines. The label “outpatient clinic” is attached here to highlight its unique character as the only North American long-term followup starting with outpatients rather than with inpatients. As noted by the authors, “…less than one-half of the individuals carrying a diagnosis of schizophrenia are likely to be represented in ‘classical’ followup studies originating in-hospital” (p. 502). Furthermore, 21 percent of their sample had no prior history of hospitalization, and such a cohort is totally unrepresented in followup studies. The authors state that only by studying schizophrenic patients who have not yet experienced hospitalization can one obtain the maximum heterogeneity needed to develop an adequate data base of course and prognosis.

The study was retrospective in design. The sample consisted of 670 schizophrenic outpatients who participated in a federally funded study of ataractic drugs that began in 1958. Patients were between 18 and 45 years old, had a primary diagnosis of schizophrenia based on DSM-II criteria, and gave evidence of mental illness of at least 1 year’s duration. More than one-half of the sample showed signs of mental illness for 10 years or more. Hospitalization experience before the clinic admission varied. As mentioned, 21 percent had no previous history of psychiatric hospitalization. Twenty-two percent experienced exclusively “crisis” admissions in a municipal hospital, and the remainder (57 percent) had at least one admission to a long-term psychiatric facility.

Outcome was unidimensional and consisted of cumulative hospitalization rates for each followup year from the patient’s entry into the study (1958–62) through December 31, 1977. This made the minimum followup 15 years (range 15–20 years). Hospitalization data were collected from the records of the New York State Department of Mental Hygiene and the Kings County Psychiatric Hospital. Information on deceased patients and those hospitalized out of New York State was excluded.

At baseline, the sample had a mean age of 30 years. Fifty-four percent were male. Sixty-eight percent were white, 28 percent were black, and 4 percent were Hispanic. Three-fourths of the patients were in the lower two socioeconomic classes (Hollingshead and Redlich 1958). Forty-seven percent were single. Slightly more than one-third were either self-supporting or supported by a spouse. Forty-five percent were supported by relatives and 19 percent received public assistance. Only 27 percent had fairly regular employment, 21 percent worked sporadically, and 52 percent had rarely or never been employed.

Three hundred eighty patients were eventually hospitalized at least once, giving a 59 percent cumulative hospitalization rate by 15 years. The average number of hospitalizations was 2.7; the average length of a single hospitalization was 16 months. The average length of total hospitalization per patient was 44 months (range 1–174 months). Interestingly, the proportion of patients residing in hospitals per year went from a high of 35 percent at year 2 of the followup to a low of 12 percent by year 15. Such a trend could reflect a diminution in psychopathology; it could also be secondary to the deinstitutionalization movement which was in full swing at the time. The majority (63 percent) of patients eventually returning to hospitals did so within the first 2 years of followup. The authors cite this as evidence that schizophrenia is not a disease of slow, progressive deterioration; after approximately the first 5 years, the average patient’s course either plateaued or improved gradually.

Boston State Hospital Followup Study. Gardos et al. (1982a, 1982b) conducted a 12-year followup study of 124 chronic schizophrenic patients hospitalized at Boston State Hospital. The study was retrospective in design. Subjects had participated as the Boston State Hospital cohort of a federally funded multihospital collaborative study of chlorpromazine in 1965. The drug study had required a primary diagnosis of schizophrenia and at least 2 years of continuous hospitalization.

Because of their involvement in the drug study, subjects had been assessed repeatedly, and considerable data existed regarding their psychopathology and social adjustment. Following the drug study, this cohort continued to receive usual treatment consisting of neuroleptics, milieu therapy, and vocational rehabilitation. Additionally, in the late 1960’s and
1970's, they were part of the deinstitutionalization movement.

Followup was conducted in 1977, 12 years after admission to the drug study, by a psychiatrist and a psychiatric nurse. The hospital records of each patient were examined from 1965 until followup or until discharge. Then each patient was personally interviewed. Outcome was assessed multidimensionally. Complete followup data were gathered on 90 patients or 73 percent of the sample. Seventeen patients (14 percent) were deceased; analysis revealed them to have been a sicker cohort. Seventeen patients (14 percent) were not located or refused participation; analysis of existing data (e.g., poststudy hospital course) revealed them to be significantly healthier than the assessed cohort.

The sample fits the profile of the early onset, drug treatment resistant institutionalized patient. On the average, subjects were first hospitalized at age 24. They entered the study at age 41 after undergoing 13 years of hospitalization. They had all been exposed to neuroleptic medication, often in high doses. Eighty-one percent were single and 60 percent had not graduated from high school. Before hospitalization, 33 percent had been skilled workers or better, 26 percent semiskilled, and 41 percent unskilled. Their global severity of illness at study entry on a 7-point scale was 5.2, or markedly ill.

Followup 12 years later found nearly all of the 90 patients distributed among a variety of structured domiciles: 21 in State hospitals, 35 in nursing homes, 5 with their families of origin, 12 in family care homes, and 13 in cooperative apartments. Only four patients (3 percent) were living unsupervised in rooms. The mean Global Assessment Scale (Endicott et al. 1976) score was 40, corresponding to major impairment in work, relationships, communication, judgment, thinking, and mood. Their 7-point global severity of illness score at followup was 4.8, not much different from their score at baseline. Only 17 patients worked—12 in hospital-based industries or sheltered workshops and 5 in competitive but menial jobs. Summarizing, the authors stated, “The overall psychosocial adjustment of the interview cohort of 90 patients was rather dismal. The typical patient at followup can be characterized as markedly ill, receiving high doses of antipsychotics, not employed, and showing poor social skills” (Cardos et al. 1982a, p. 21).

**Chestnut Lodge Followup Study.** McGlashan reported a retrospective followup study on patients discharged over a 25-year period from Chestnut Lodge, a 90-bed private tertiary care residential treatment facility in Rockville, Maryland, near Washington, DC (McGlashan, 1984a, 1984b). Patients were largely young, chronically ill treatment failures from upper socioeconomic brackets referred from across the continental United States. They suffered from schizophrenia, affective disorders, schizotypal and borderline personality disorders, and “other” (undiagnosed). While patients received a primary diagnosis for the main study, their scores on all diagnostic systems were recorded to assess the presence/effects of comorbidity.

Followup was conducted between 1977 and 1983 by members of the research team without knowledge of the baseline data. Outcome was assessed an average of 15 years postdischarge (range 2–32 years) by personal interviews, in person and by telephone, with patients and/or significant others. The average age of the patients at followup was 47 years. Deceased subjects were included. Outcome was multidimensional in scope and included rating scales from existing followup studies. All measures were operationalized, tested for reliability, and case calibrated. Assessment was completed on 446 patients, or 72 percent of the total possible sample (n = 619). Eighty-six additional hospital records of nonlocatable and refusing subjects were abstracted and rated to test for sampling.
biases by comparing their baseline profiles with those of the patients' completing followup. Such biases proved to be minimal.

For the schizophrenic cohort, baseline data were available on 188 patients and followup data, on 163 patients. This cohort had the following baseline demographic and predictor profile: males—52 percent, married—23 percent, white—100 percent, socioeconomic status—1.6 (Hollingshead and Redlich 1958), age of onset—19 years, age of first hospitalization—23 years, age of index hospitalization—28 years, number of prior hospitalizations—3, length of prior hospitalizations—28 months, length of prior outpatient treatments—17 months, and admission psychopathology on a 7-point scale—5.5. The sample was severely and chronically ill; 90 percent had been ill for more than 2 years. They were also resistant to drug treatment insofar as most had been tried unsuccessfully on neuroleptic medications before index admission. In this way, they were similar to the schizophrenic patients in the Boston State Hospital sample.

By followup, this schizophrenic sample distributed as follows on a 5-point global outcome scale: recovered—6 percent, good—8 percent, moderate—23 percent, marginal—23 percent, continuously incapacitated—41 percent. The "marginal" anchor point meant that, on the average, the patient had spent about 25 percent of the followup period in sheltered situations, worked about 20 percent of the time, claimed some role-specific social contacts, and experienced symptomatic expressions of illness about 75 percent of the time. The entire cohort scored 37 on the Health-Sickness Rating Scale (Luborsky 1962), a 100-point global scale similar to the Global Assessment Scale, placing them at a level comparable to the Boston State Hospital sample. Outcome did not change significantly with followup time when compared across decades. Long-term course proved to be a stable plateau with no evidence of trends toward improvement or deterioration.

**Vermont State Hospital Followup Study.** Harding et al. (1987a, 1987b) recently reported a retrospective followup study based on a cohort of 268 DSM-I (American Psychiatric Association 1952) schizophrenic patients from Vermont State Hospital who were referred to a rehabilitation and community placement program at this institution between 1955 and 1960. Twenty to twenty-five years later, 97 percent of the patients, most still residing in Vermont, were located, and received comprehensive outcome assessment based on personal interviews in their own homes. Additionally, the followup assessors, who were unaware of the patient's baseline diagnostic and functional status, checked the veracity of each patient's outcome report with knowledgeable significant others. Outcome assessments were multidimensional and tested for reliability. Deceased patients were evaluated, although the reports to date focus only on the patients who were alive and personally interviewed (n = 168). Since virtually all patients were found and assessed, missing subjects were not an issue.

Baseline demographic, predictor, and psychopathological data from each patient's hospital record were rated using a standardized review form by investigators without knowledge of outcome. Ratings were tested for reliability. Kappa coefficients ranged between .40 and .95. Patients were then rediagnosed according to DSM-III criteria. The reliability of this judgment was tested twice, with variable results (Kappas ranging between .40 and .78). Of the original 268 patients, 118 received a DSM-III diagnosis of schizophrenia.

The final sample, after rediagnosis, consisted of 82 live and interviewed DSM-III schizophrenic patients. Their average age at followup between 1980 and 1982 was 61 years. At baseline, this sample had the following demographic and predictor characteristics: male—50 percent, single—62 percent, and completed high school education—45 percent. Before transfer to the rehabilitation program, 45 percent of the cohort had been hospitalized more than 6 years, 24 percent between 2 and 6 years, and 31 percent less than 2 years.

While undoubtedly chronic by length of hospitalization standards, this sample was unique in many respects. Fortunately, numerous additional sources of information exist about the Vermont State Hospital program and its subjects that provide clues. First, neuroleptic drugs (chlorpromazine and reserpine) were introduced to Vermont State Hospital for the first time in 1954 (Brooks 1956). Some patients improved enough to leave the hospital quickly. A larger number of patients improved but had no prospect of leaving the hospital for a variety of reasons: no family, lack of placement, absent financial resources, or "raw fear" of separating from the institution (Brooks, personal communication, March 29, 1986). It was this group of patients for which the rehabilitation/community placement program was devised, and this group from which the initial referrals to the program were made (Brooks and Deane).
1965). They were drug-responsive but deskilled secondary to institutionalization, not deterioration (Brooks 1960).

Second, study subjects were instrumentally functional; all patients referred to the rehabilitation program were already working at hospital jobs up to 30 hours per week (Chittick et al. 1961; Brooks, personal communication, March 29, 1986). In fact, many of the patients were selected for the rehabilitation program by asking hospital work supervisors which patients he/she could least afford to lose (Brooks 1959)! Up until 1954, two-thirds of all work at Vermont Hospital came from unpaid patient labor. Between 1955 and 1965, the success of the rehabilitation program cut this patient-labor force in the hospital by 85 percent (Brooks et al. 1970). Finally, the patients were highly motivated to work since they were poor and had no prospects of income from family or public welfare (Brooks, personal communication, March 29, 1986).

The best description of the selection process for the rehabilitation program comes from The Vermont Story (Chittick et al. 1961):

Patients are informally referred to the rehabilitation service by anyone in the hospital.... This usually insures that they are well enough to arouse the interest of at least one other person.... Obviously, the person referring and those accepting a patient must feel a certain optimism about him... [p. 24]

Possibly the patient is a good worker in the hospital industrial program or a cooperative patient on the ward. The patient may have undergone a sudden and dramatic change in behavior which has gained him the attention of others. [p. 27]

Clearly, patient selection was not a random or arbitrary matter. The process capitalized, in fact, upon emerging signs of health and hope.

The outcome of this patient cohort is also unique. A 5-year followup of the entire original 268 referrals to the program was reported in 1967 (Deane and Brooks 1967). Thirty percent of the patients were in the community without readmission and 40 percent were in the community but had undergone about two readmissions apiece. The remaining 30 percent were in the hospital, two-thirds as a result of readmission from the community. Most patients were single and used community care facilities for socializing; most replaced institutional employment with sheltered employment (e.g., cook in a nursing home with bed and board) and kept in close contact with their rehabilitation program counselors.

The picture was one of progress, but against strong regressive resistance with high utilization of mental health manpower and support systems.

Followup 20 years after the program in 1980–82 recorded that these patients had consolidated their gains of 5 years and surpassed them beyond everyone’s expectations. The DSM-III schizophrenic cohort scored as follows on several of the Strauss and Carpenter (1972) outcome dimensions: not in the hospital in the past year—82 percent, meets with friends every week or two—61 percent, employed in the last year—40 percent, and displays slight or no symptoms—68 percent. On the Global Assessment Scale, 60 percent of the patients scored over 61, designated as good functioning. No one scored in the poor functioning category (less than 31). Of further interest is that outcome was not significantly different on any dimension between the DSM-III schizophrenic cohort and either a DSM-I schizophrenic cohort (n = 149) or a heterogeneous DSM-III nonschizophrenic cohort (n = 71). As such, the Vermont study is singular in finding no differences in long-term outcome between schizophrenia defined by different criteria (DSM-III and DSM-I) and, more strikingly, between schizophrenia and nonschizophrenia (both DSM-III).

Washington-International Pilot Study of Schizophrenia Followup Study. Carpenter et al. (1987) recently completed an 11-year followup on a subgroup of the original Washington, DC, cohort of the International Pilot Study of Schizophrenia (IPSS; World Health Organization 1979). This followup is the only long-term North American followup study using a prospective design, that is, following the development of the illness with serial, independent, cross-sectional assessments. The first followup was conducted at 2 years (Strauss and Carpenter 1972, 1974) and the second at 5 years (Hawk et al. 1975; Strauss and Carpenter 1977).

The original sample consisted of 131 patients admitted in 1968–69 to the psychiatric units of general hospitals of Prince George’s County, Maryland, a largely middle- and lower-middle-class area around Washington, DC. Patients were referred for study if they displayed at least one psychotic symptom on admission and had no organic problems, drug, or alcohol abuse. Patients were assessed diagnostically and prognostically at baseline by structured interview. Ratings were tested for reliability.
The diagnosis of schizophrenia was made clinically using DSM-II and ICD-9 (World Health Organization 1978) criteria. This original cohort ranged between acute and sub-chronic. Subjects were screened out if they had been hospitalized for more than 2 of the previous 5 years, or if they gave evidence of continuous psychosis for longer than 3 years.

Followup concentrated on the 68 patients who were evaluated at the 5-year followup. The 11-year assessment was conducted blind to any prior information about the subjects. Fifty-three patients received complete assessment. The 15 missing subjects were compared with completed subjects on key baseline variables; no consistent biases were demonstrated. The followup cohort was assessed with a structured personal interview by telephone or in person. Outcome was multidimensional and operationalized (Strauss and Carpenter 1972, 1974). Forty of the followup cohort had an original diagnosis of schizophrenia; the remaining 13 nonschizophrenic patients carried diagnoses of manic-depressive disorder, personality disorder, or neurosis.

Followup found the schizophrenic patients to be functioning at a level inferior to the non-schizophrenic patients on all dimensions: hospitalization, work, social functioning, symptom severity, and global outcome. Outcome level did not change between the 5-year and 11-year followups, findings compatible with the idea that schizophrenia tends to plateau after 5 years.

Columbia-Psychiatric Institute Followup Study. In 1985–86, Stone conducted a retrospective followup of patients admitted between 1963 and 1976 to a long-term (average 1 year) treatment unit of the New York State Psychiatric Institute (PI) in New York City (Stone 1986; Stone et al. 1986). The unit is an academic training center in psychiatry for Columbia University and, at the time, specialized in providing intensive, psychoanalytically oriented psychotherapy. Five hundred fifty patients were selected who were less than 40 years old at admission, who registered an IQ greater than 90, and who were hospitalized on the unit for more than 3 months. Subjects were diagnosed according to DSM-III criteria applied to their hospital records. Ninety-nine patients met the criteria for schizophrenia. Comparison groups carried the following diagnoses: schizoaffective psychosis ($n = 64$), schizophreniform psychosis ($n = 64$), manic-depressive psychosis ($n = 39$), and borderline personality disorder ($n = 205$).

The author assessed outcome nonindependently by personal interview, telephone and face-to-face, between 10 and 23 years after admission. Five hundred four patients (92 percent of the sample) completed outcome assessment, which included ratings of work functioning during followup, current living situation, and global functioning by the Global Assessment Scale.

The overall cohort at admission was 22 years old, had an IQ of 119, and came from mid-level socioeconomic circumstances (2.7 on the Hollingshead-Redlich scale). They had experienced 1–3 months of hospitalization before admission.

At outcome, 13 percent of the schizophrenic sample had been or were married. Their Global Assessment Scale score was 39 (range 6–81), and only 8 percent of the sample reached a level of "good or recovered" (greater than 61). Outcome for the schizophrenic patients was inferior to that of all comparison groups.

The relevant studies have been described and summarized. Let us now review what the principal results have confirmed and/or taught us about the natural history, prognosis, and nosology of schizophrenia.

Natural History of Schizophrenia

1. Schizophrenia is a chronic disease, frequently disabling for a lifetime. This certainly comes as no surprise to anyone touched by the disorder, but its magnitude had never really been demonstrated against a normal control group until the Iowa 500. There, the outcome of schizophrenic patients proved to be significantly poorer than the outcome of nonpsychiatrically disordered surgical patients. This difference extended across all domains of functioning and across three to four decades of the adult lifespan.

2. The average outcome of schizophrenia is worse than that of other major mental illnesses. The Iowa 500 was, again, the first to demonstrate significant long-term differences between schizophrenia and other psychotic disorders, specifically mania and depression as defined by the criteria of Feighner et al. (1972). The Chestnut Lodge followup replicated this finding for (DSM-III) unipolar and bipolar affective disorders. Later reports detailed that the Chestnut Lodge schizophrenic cohort also had poorer long-term outcome compared to schizoaffective psychosis (Williams and McGlashan 1987; McGlashan and Williams 1987), schizotypal personality disorder (McGlashan 1986d), and borderline personality disorder.
(McGlashan 1986e). The Wash-
ington-IPSS investigation found that
schizophrenic patients were doing
worse than a mixed sample of
patients with affective, personality,
and neurotic disorders. The Colum-
bia-PI schizophrenic cohort had the
poorest outcome in comparison to
DSM-III-based schizoaffective psy-
chosis, manic-depressive psychosis,
and borderline personality disorder.
The only exception to this trend
was the Vermont State Hospital
study, which registered no dif-
ference between schizophrenic and
nonschizophrenic cohorts. Overall,
however, the preponderance of evi-
dence upholds Kraepelin’s original
hypothesis bifurcating the psy-
choses into the affective psychoses
and schizophrenia, with the latter
having a more pernicous long-term
course and outcome (Kraepelin
1896/1919).

3. Schizophrenia is associated with an
increased risk for suicide, physical ill-
ess, and mortality. When studied,
the rates of suicide proved to be
significantly higher for schizo-
phrenia than were the rates for
contrasting general populations. In
the first Alberta followup, the su-
cide rate was 2.3 percent. In the
Chesnut Lodge followup, it was 8
percent (Dingman and McGlashan
1986), and in both the Iowa 500
(Tsuang 1978b) and Columbia-PI
studies, it was 10 percent. Death
comes more quickly to schizo-
phrenic patients in general (Bland
et al. 1976, Tsuang and Woolson
1977), and this excess mortality can-
not be accounted for solely by
suicide and/or accidents (Tsuang
and Woolson 1978). From the Iowa
500 sample, infection and circula-
tory diseases also contributed
(Tsuang et al. 1980b). Shortened
survival was about 10 years for the
male schizophrenic patients and 9
years for the female schizophrenic
patients (Tsuang et al. 1980b).

4. The schizophrenic process, while
disabling and chronic, does not get pro-
gressively worse over the long-term.
The hypothesis that schizophrenia
follows a relentlessly downhill
course to dementia has finally been
put to rest. Deterioration of func-
tioning does characterize the
disease in its early stages and has,
in fact, become one of the DSM-III
diagnostic criteria. At some point,
however, loss of functioning
appears to “bottom out” or plateau.
This “point” varies widely between
individuals, but occurs roughly 5–10
years after the manifest illness
becomes unequivocally established.
As mentioned, the New York City
outpatient followup study found
that 63 percent of patients returning
to hospitals did so within the first 2
years of followup. More striking
were the differences in the Vermont
State Hospital cohort between their
5- and 20-year followup assess-
ments. As detailed above, the
5-year point found patients strug-
gling with disability and resisting
self-sufficiency. By 20 years,
however, they had consolidated
earlier gains, internalized rehabilita-
tive strategies, and progressed in a
steady, measurable fashion Some-
where between that 5 and 20 years,
the pressure of the disease
plateaued and/or relented
somewhat.

This “process plateau” in chronic
schizophrenia, once established,
appears to be stable. For example,
the Washington-IPSS study found
little to no change between schizo-
phrenic outcome functioning from 5
to 11 years. The Chesnut Lodge
followup study demonstrated no
significant differences in schizo-
phrenic outcome across three
decades after index hospital dis-
charge. Within any given sample,
both patients and controls, in fact, there is a remarkable steadi-
ness. This diminished variance in
the overall process, however, does
not appear to represent a rigidifica-
tion of disease and/or personality.
The semi-independent domains of
outcome remain semi-independent
(Carpenter et al. 1987), and poten-
tial still exists for progressive
rehabilitation (Harding et al. 1987a,
1987b).

5. Among patients with schizo-
phrenia, however defined, outcome is
heterogeneous. The above remarks
about long-term plateauing apply
mainly to patient samples in which
the schizophrenic process has
already become manifestly chronic.
For many patients, the disease
never gets this far. Among this
larger group, all of whom lay legiti-
mate claim to the diagnosis of
schizophrenia, heterogeneity of out-
come is the rule. The evidence
resides in the final column of table
1 where, depending on the sample,
outcome can vary between com-
plete recovery and continuous
incapacity. Furthermore, each level
on the outcome spectrum is repres-
ented by substantial numbers of
patients, we are not dealing with
spurious occurrences or false-posi-
tive diagnoses. As will be
discussed, a great deal of this het-
ergeneity can be attributed to
sample differences. Nevertheless,
an important fact remains: a lot of
patients recover from schizo-
phrenia. The certainty of negative
prognosis in schizophrenia is a
myth.

6. Much of the heterogeneity in the
long-term course of schizophrenia can be
linked to sample characteristics and/or
differences. How can we account for
this heterogeneity? Here we leave
the realm of knowledge and enter
the domain of speculation.
However, the long-term followup
perspective affords us a rich and
perhaps unique opportunity to gen-
erate some reasonable hypotheses. Close scrutiny of the outcomes in table 1, in conjunction with the characteristics of the samples studied, suggests that heterogeneity may be linked with levels of chronicity. For example, the best outcomes were recorded for patients from the two Massachusetts Mental Health Center studies, the Phipps study, and the two Alberta studies. These results arose from samples of acute, first admission, or remitted patients—that is, nonchronic patients who were either early into their illness course or who had demonstrated hegemony of health over psychopathology. This association of acuteness with good outcome was not absolute (e.g., the second Massachusetts Mental Health Center study), but it held on the average. Thus, the length of time over which the schizophrenic illness has been manifest—as prodrome, active positive/negative symptoms, or residual defect symptoms—may be crucial to ultimate long-term outcome. Furthermore, the chronicity threshold time period may be on the order of 6 months to 1 year. The outcome of the Iowa 500 schizophrenic cohort was decidedly worse than those of the cohorts just cited, and all of the Iowa 500 patients had been manifestly ill for more than 6 months, as required by the Feighner criteria, or for more than 1 year (85 percent of the sample), but not for much longer since the average time between onset and index hospitalization was short.

Studies of samples where the documented illness was even longer found still poorer outcomes (Boston State Hospital and Chestnut Lodge), thus supporting the validity of length of mental illness (LOMI) as an outcome-determining parameter. The Vermont State Hospital followup, however, stands as a key exception. Their cohort was decidedly chronic by the LOMI criterion, but their outcome was clearly better. If the LOMI criterion is valid, this discrepancy is curious. It may arise from methodological sources such as liberal outcome ratings or from hidden sample diagnostic heterogeneity (recall their variable diagnostic reliability). But it is hard to conceive that these could account for the total discrepancy.

A closer look, however, suggests that chronicity may contain more facets than just LOMI. Table 2 presents three long-term followup studies of chronic patients with sufficient baseline and outcome data for detailed comparison. The Boston State Hospital, Chestnut Lodge, and Vermont State Hospital studies used many comparable (and sometimes identical) outcome measures, as recorded in table 1. For comparison purposes, in the final column of Table 2, I have translated these into a single global outcome score based on a 5-point scale (0 = continuously incapacitated, 1 = marginal, 2 = moderate, 3 = good, and 4 = recovered). Table 2 also compares these study samples across key demographic, premorbid, and morbid characteristics. The demographic variables are gender, marital status, socioeconomic status, and physical setting. The premorbid variables are education and premorbid work/social functioning. The morbid variables are age of onset, age of first hospitalization, age of index hospitalization, length of prior hospitalizations/treatment (a measure of LOMI), and exposure/response to neuroleptic medications.

All samples were decidedly chronic by the LOMI criterion. They were also remarkably similar with respect to gender, marital status, and premorbid work/social functioning. The Boston State Hospital and Chestnut Lodge patients, furthermore, were alike in falling ill at an earlier age and in their exposure and poor response to prior trials of neuroleptic drugs. The differences in outcome between Boston State Hospital and Chestnut Lodge, therefore, may be linked to the marked discrepancy in socioeconomic status (which probably also accounts for the differences in education). This association is supported by the fact that Chestnut Lodge outcome was superior to that of the Boston State Hospital cohort in the followup dimensions of living situation and employment but not psychopathology—that is, in the domains most likely to be influenced by economic resources.

The Vermont State Hospital patients came from socioeconomic situations comparable to those of the Boston State Hospital patients. There, however, further similarities end. The Vermont followup patients were different from the Boston State Hospital and the Chestnut Lodge patients in three domains, all of which could be linked to their better outcomes. First, they resided in a rural setting. Second, they were exposed to drugs and responded, albeit incompletely. Finally, they appeared to have a considerably later age of illness onset. This again suggests that the sample may have been diagnostically heterogeneous and, for example, possibly contained affective disorders which have a later onset. If the sample was homogeneously schizophrenic, on the other hand, this difference suggests that age of onset may be a far more powerful predictor of outcome within schizophrenia than heretofore demonstrated.

Leaving physical setting and
socioeconomic status aside for the moment, the sampling and outcome differences among these studies suggest that chronicity is multidimensional. Four suggested dimensions are listed in the penultimate column of table 2. LOMI is a central dimension. Another well-known dimension is institutionalization, or the degree to which patients remain attached to or controlled by the treatment milieu. The New York City outpatient study found, for example, that the best predictor of subsequent hospitalization was amount of prior institutionalization—not degree of psychopathology. Some patients in that study were chronic by the LOMI criterion but had never been hospitalized, and they avoided later hospitalization significantly more often than patients referred to the study from long-stay institutions. This form of chronicity was probably ubiquitous among the three samples in table 2.

The introduction of powerful neuroleptic drugs into the treatment of schizophrenia over the last 30 years has established another dimension of chronicity: biological treatment resistance. This form of chronicity was definitely present in the Boston State Hospital and Chestnut Lodge samples. In the former, the followup cohort consisted of patients who had not responded to prior drug trials at that institution. In the latter, the followup cohort consisted of patients who had, by their index admission, failed to respond to prior trials of somatic treatments (insulin coma, electroconvulsive therapy, and/or neuroleptic medications) at other institutions. The Vermont State Hospital sample, however, was hospitalized in the predrug era. Patients there were referred to the rehabilitation program in conjunction with initial trials of neuroleptic medication being introduced to North America in the mid-1950's. While the program patients were unable to leave the institution without additional psychosocial

<table>
<thead>
<tr>
<th>Study</th>
<th>Male (%)</th>
<th>Single (%)</th>
<th>Socio-economic status¹</th>
<th>Physical setting</th>
<th>Education: % high school graduates</th>
<th>Premorbid work &amp; social functioning</th>
<th>Age of onset (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boston State Hospital</td>
<td>51</td>
<td>81</td>
<td>4 &amp; 5</td>
<td>Urban</td>
<td>40</td>
<td>Poor</td>
<td>—</td>
</tr>
<tr>
<td>Chestnut Lodge Hospital</td>
<td>52</td>
<td>77</td>
<td>1 &amp; 2</td>
<td>Suburban</td>
<td>82</td>
<td>Poor-moderate</td>
<td>19</td>
</tr>
<tr>
<td>Vermont State Hospital</td>
<td>50</td>
<td>62</td>
<td>4 &amp; 5</td>
<td>Rural</td>
<td>45</td>
<td>—</td>
<td>30²³</td>
</tr>
</tbody>
</table>

¹Hollingshead-Redlich (1958)
²Data available only for entire Vermont State Hospital cohort
³Calculated by subtracting average length of disability (10 years) from age at study admission (40 years)
interventions, they were, at least from the available anecdotal reports, drug responsive to varying degrees.

Age of onset also differentiated the Vermont State Hospital from the Boston State Hospital and Chestnut Lodge samples. Might this be another facet of chronicity, with earlier onset being linked to a more constitutionally determined pathogenesis? This interpretation is speculative, however, in the absence of further consistent empirical evidence linking age of onset with prognosis, familial pedigree, syndromal subtype, and so on. Accordingly, it is noted as a dimension of chronicity in table 2, but with a question mark.

Considering chronicity as a multivariate construct may help us at least propose reasonable hypotheses to account for the large differences in outcome among our comparative samples. It appears, for example, that the patients from Chestnut Lodge and Boston State Hospital were chronic in all four of the proposed dimensions, whereas the patients from Vermont State Hospital were chronic in only two of them. While the validity of these constructs requires further empirical study, the implications of this comparative exercise are clear: adequate sample description in followup studies is vital. Without it, the interpretation and generalization of findings becomes severely constricted. Minimally, descriptors should include the variables listed in table 2. Biological treatment resistance, in particular, is a new construct that needs further attention and operationalization.

Clearly, a great deal of long-term outcome variance can be accounted for by sample characteristics. As a spinoff of the North American followup studies, we also have a clearer notion of which characteristics may be the most important. Quantifying the amount of outcome variance accounted for by each characteristic remains a task for future followup investigations.

7. Long-term followup studies have yet to demonstrate clearly the effects of treatment on the natural history of schizophrenia. Valid and specific estimates of treatment effects require the presence of nontreatment and/or alternate-treatment control groups. Furthermore, all of the current treatments of schizophrenia, biological and psychosocial, are time limited. Following treatment termination, study patients are
released to the mixmaster of "doctor's choice," resulting in a complete loss of treatment homogeneity. Quantification of treatment influences, therefore, remains a matter for short-term followup. Long-term followup, in fact, may never be a proper strategy for this realm of inquiry.

Long-term followup can still provide useful information about treatment. It is instructive to know, for example, that many schizophrenic patients do well without maintenance medication. Of the 51 recovered patients in the first Alberta followup, 45 percent discontinued their medicine within the first 10 months of baseline. Fourteen percent of the Chestnut Lodge schizophrenic patients achieved drug-free remission for the entire followup (Fenton and McGlashan 1987b). Twenty-five percent of the Vermont cohort remained on medication regularly, 25 percent took medication only during symptomatic exacerbations, and the remaining 50 percent were noncompliant (34 percent) or required no prescriptions (16 percent). Such findings establish the existence and validity of important medication-relevant subgroups, the identification (or prediction) of which has implications for treatment and for identifying subtypes within the schizophrenic syndrome.

Such findings also suggest what kind of information about treatment can be gleaned validly from long-term followup. For example, uniformly good long-term outcome in a sample not receiving treatment "x" (e.g., neuroleptics) suggests that treatment "x" is not necessary for that sample. It says nothing about the efficacy of treatment "x" for that sample except that the benefit-to-risk ratio is unlikely to be advantageous. On the other hand, uniformly bad long-term outcome in a sample receiving treatment "y" strongly suggests that treatment "y" fails to alter the natural history of illness for that sample. Here, outcome does say something about the efficacy of treatment "y"; it demonstrates a low benefit-to-risk ratio of the treatment for this sample. Such were the conclusions drawn by McGlashan (1984c) and Stone (1986) about the utility of intensive individual psychoanalytically oriented psychotherapy as the primary treatment for young, chronic schizophrenic inpatients.

8. Long-term followup can be informative about how sociocultural factors may influence course. Sociocultural forces are in the nature of existential situation rather than controlled perturbations of treatment. Sociocultural status and physical setting, for example, often last a lifetime and can reasonably be expected to make a long-term difference. Here too, however, demonstrating an effect requires tracking the long-term course of at least two "captive" samples that are matched except for the characteristic under consideration. The Vermont State Hospital sample is instructive in this regard. It is a remarkable cohort, perhaps unique in U.S. psychiatry. As detailed in The Vermont Story (Chittick et al. 1961), these patients were provided with virtually everything that 20th century psychiatry has to offer, all within one decade between 1947 and 1957. Furthermore, these advances were introduced in proper sequence and applied with assiduous continuity of care to a target population that remained stable and local. The effects of this are undoubtedly reflected in the cohort's long-term outcome. Teasing apart which of the many interventions were primary, however, requires control groups that are probably impossible to find and/or construct.

These findings are still interesting because they challenge many long-standing notions, such as that a mutually negative interaction occurs between schizophrenia and lower socioeconomic status. The Vermont patients were unquestionably poor and uneducated. However, according to George Brooks (personal communication, March 29, 1986), the program's primary engineer, their poverty was a major source of motivation; they had to work. The socioeconomic advantage of the Chestnut Lodge schizophrenic patients, on the other hand, failed to protect them from the ravages of severe psychopathology. For many, money actually worked to their disadvantage; it allowed them to purchase isolation and freedom from the necessity for initiative. Together, the Vermont State Hospital and Chestnut Lodge studies serve as gadflies to time-honored sociological shibboleths touting a simple linear relationship between economics and illness.

Together, the Vermont State Hospital and Boston State Hospital studies suggest that physical setting may strongly influence course. Low income, inner city environments, like that around Boston State Hospital, may be particularly toxic, as witnessed also by the young, chronic, homeless patients that drift from one urban center to another. In contrast, small town, rural, and nontransient environments like that in Vermont may provide a stable foundation for needed continuity of care. These studies corroborate findings that the course of schizophrenia may be more benign in developing Third World countries than in technologically developed nations (World Health Organization 1979).
Prognosis or the Prediction of Outcome in Schizophrenia

Long-term followup studies have demonstrated the remarkable heterogeneity of course in schizophrenia. Many of these studies have also been instrumental in reducing this heterogeneity by identifying subgroups of schizophrenic patients with more homogeneous outcomes. The characteristics identifying these subgroups are called predictors of outcome or prognostic variables. We have already seen how outcome can be linked to sample characteristics, particularly various dimensions of chronicity. This section expands upon that discussion and reviews what the North American long-term followup studies have taught us about prognosis.

Prediction is the formal study of the association between measurable sample characteristics and outcome. Since, however, we do not know what schizophrenia is, confusion frequently arises as to whether a given variable is a sample characteristic or one of the criteria for diagnosis. What may be a prognostic variable in one study (e.g., LOMI) becomes a diagnostic criterion for schizophrenia in another (e.g., the Iowa 500). Clearly, then, the diagnostic criteria used to define schizophrenia in any sample have prognostic implications. This is the Heisenberg Uncertainty Principle of predicting outcome in schizophrenia: the entity you are measuring moves simply by virtue of how you define it. The implications of this are explored in more detail in the section under nosology. Here discussion focuses upon what we have learned about the long-term followup effects of the so-called "nondiagnostic" predictors of outcome. These include, in addition to the "classical" predictors, considerations of gender, subtyping, and comorbidity.

1. The schizophrenic process may be worse for men than for women. A recent review of the literature on sex differences and severe psychopathology (Bardenstein and McGlashan 1987) indicates that gender differences occur frequently in schizophrenic patients and that schizophrenic women exhibit a less deteriorated course of illness. Analysis of the Chestnut Lodge followup schizophrenic cohort by gender strongly endorsed this view (McGlashan and Bardenstein 1987). The women were superior to the men at baseline in their social and sexual/marital functioning. At long-term outcome, the women were significantly better than the men in these areas and, in time, in areas of symptomatology, substance abuse, and global functioning.

The other North American followup study to investigate gender effects (Loyd et al. 1985) found that sex made no significant contribution to the explanation of outcome differences. This lack of findings, however, may be a function of their defining schizophrenia by the Feighner et al. criteria. Lewine et al. (1984) found that the Feighner et al. system completely failed to diagnose women as schizophrenic. This may arise because of their more frequent affective symptoms. More obviously, one of the Feighner et al. criteria for schizophrenia requires that the patient be single. As the literature repeatedly documents, schizophrenic women are married more often than schizophrenic men. Thus, narrowly defined criteria for schizophrenia, like the Feighner et al. system, may select an "atypical" population of women, such that conclusions about sex differences may not be representative of the average female schizophrenic patient.

2. Long-term followup supports the validity of paranoid/nonparanoid, but not other, subtypes of schizophrenia. This conclusion comes from studies on the Iowa 500 sample. Using subtyping criteria of DSM-III, Research Diagnostic Criteria (Spitzer et al. 1977), ICD-9 (World Health Organization 1978), and Tsuang and Winokur (1974), Kendler et al. (1984) divided the Iowa 500 schizophrenic cohort into paranoid and nonparanoid subtypes. The latter were further divided into hebephrenic and undifferentiated subtypes. Long-term outcome was better for the paranoid subtype, and the paranoid/nonparanoid differences were greatest using the Tsuang-Winokur criteria, probably because they exclude patients with thought disorder and affective deterioration. Long-term outcome findings did not support the predictive validity of the hebephrenic/undifferentiated breakdown.

An earlier study found that the paranoid/nonparanoid subtypes were relatively stable diagnostically over the long-term, but not as stable as the diagnosis of schizophrenia itself (Tsuang et al. 1981). A change of subtype from paranoid to nonparanoid was more common than vice-versa.

3. Schizophrenia can coexist with other forms of psychopathology, and this comorbidity can have strong prognostic implications. These conclusions come from studies on the Chestnut Lodge followup sample of largely chronic schizophrenic patients. Severe obsessive-compulsive symptoms defined a subgroup (n = 21) of the study's schizophrenic cohort with a uniformly poorer long-term course and outcome (Fenton and McGlashan 1987b). Schizophrenia in patients with a concomitant schizo-
Typical personality disorder had a poorer prognosis at baseline, but a better outcome than "pure," non-comorbid schizophrenia (McGlashan 1986d). Schizophrenia with mixed borderline and schizotypal personality features also did better than schizophrenia without such comorbid Axis II psychopathology (McGlashan 1983).

These findings raise interesting questions. It is not surprising that "adding" obsessive-compulsive psychopathology to schizophrenia results in a worse combination, whether that combination be considered a subtype of schizophrenia or simply the coexistence of two separate illnesses (true comorbidity). It is surprising, however, that by the same model, "adding" Axis II schizotypal personality disorder and/or borderline personality disorder to schizophrenia results in a better combination vis-a-vis long-term course and outcome. Since the Chestnut Lodge followup study is the first and, to date, the only study investigating the prognostic significance of comorbidity, knowledge in this area must be considered tentative.

4. Many important predictors of outcome have been identified by the North American followup studies. The long-term followup investigations that tested for nondiagnostic predictors of outcome were the Massachusetts Mental Health Center, Phipps Clinic, Iowa 500, Chestnut Lodge, and Washington-IPSS studies. The major predictor dimensions to emerge from these efforts are summarized in table 3, along with the direction of each dimension associated with better outcome. Identified predictors came from many categories: genetics, premorbid functioning, illness onset, psychopathological signs and symptoms, and course of illness up to index admission. For the most complete assessment of prognosis, these predictors should be considered in conjunction with the additional sample characteristics discussed in the last section: socioeconomic status, physical setting, age of onset, length of manifest illness, biological treatment resistance, and institutional chronicity.

Building upon their identified predictors, four of the investigating teams developed prognostic scales. These consist of several key predictor variables, each scored as present or absent, and collected together to give a total score (Vaillant 1964b, Stephens 1978, Fenton and McGlashan 1987a; Carpenter et al. 1987). Such scales are in keeping with the notion that prognosis is a dimensional, not a categorical, phenomenon. There are no categorically "good" or "poor" prognostic schizophrenic patients versus categorically "good" or "poor" prognostic cases, but rather varying combinations along a spectrum. The scores achieved on these scales usually correlate better with outcome than the scores achieved on any single constituent predictor. The reader is referred to specific citations for details.

In addition to the idea that predictors are dimensional phenomena, the following "principles" of prediction have emerged from the North American followup studies. First, cross-sectional psychopathology alone and diagnoses based primarily upon cross-sectional symptoms have limited prognostic value. Longitudinal data have greater outcome predictive power (Carpenter et al. 1987). Second, predictions vary according to the outcome dimension being predicted (McGlashan 1986b; Carpenter et al. 1987). For example, premorbid social functioning may predict followup social functioning but not predict followup hospitalization. The latter, in turn, may be predicted by some other variable, like LOMI. Third, "like" predicts "like." In general, each domain of outcome functioning is best predicted by its premorbid dimensional counterpart (McGlashan 1986b; Carpenter et al. 1987). For example, marital status at outcome is best predicted by marital status at baseline. Fourth, predictors vary according to the length of followup. In the Chestnut Lodge schizophrenic sample, for example, premorbid functioning variables emerged as strong predictors of outcome for the first 10 years following index discharge. Family and manifest illness variables were important in the second followup decade. Genetic predisposition (family history of schizophrenia) emerged with manifest illness as central to predicting outcome at three decades and beyond (McGlashan 1986c). Finally, predictors of outcome exist even for chronic schizophrenia. In the Chestnut Lodge followup study, predictors accounted for approximately one-third of the outcome variance across six outcome dimensions (McGlashan 1986d).

**Schizophrenic Nosology**

Since we do not know its etiology(ies), the diagnosis of schizophrenia rests upon manifestations of the disorder. Which manifestations are chosen to define cases with the disease constitutes the discipline of diagnostics or nosology. Several factors are important in constructing any particular set or "system" of diagnostic criteria—for example, reliability, concordance with established clinical use, comprehensiveness, and specificity, or overlap and concord-
The factor most relevant to followup, however, is predictive validity, or the degree to which a set of diagnostic criteria can predict the long-term "behavior" of its identified constituents. An early example of long-term followup being used in this way comes from Vaillant (1963). He retrospectively identified 12 cases of "remitting schizophrenia"—that is, schizophrenic patients admitted to a State hospital in the early 1900's who remitted and were discharged. Followup more than 50 years later revealed that some patients remained well without further hospitalization. Seventy-five percent of them, however, were again hospitalized at some time for psychosis, and a portion of these remained chronic after relapse. Because of this, Vaillant contended that there exists no justification for separating the remitting schizophrenics from the broader class of schizophrenic patients. He used the evidence of long-term followup to contest the hypothesis that schizophrenia and remitting schizophrenia were different nosologic entities.

Predictive validity is usually estimated as the degree to which a given diagnostic system can predict long-term diagnostic stability, functional outcome, or specific functional states. In schizophrenia, the latter usually refer to chronically psychotic or defect end states, as per Kraepelin (1896/1919). Table 4 summarizes the various diagnostic systems that have been tested for predictive validity by data from the North American long-term followup studies. "System stringency" in column 3 refers to the inclusiveness of a given diagnostic system. Here, this means the number of patients defined as schizophrenic out of the given sample. Roughly defined, these systems are considered broad or inclusive if they identify three-fourths or more of any sample as schizophrenia and narrow or stringent if they identify one-third or less of the same sample as schizophrenia. In table 4, the degree of stringency of each system with the given followup sample is estimated roughly on a scale of 5.

### Table 3. Predictors of outcome in schizophrenia: North American followup studies

<table>
<thead>
<tr>
<th>Predictor category</th>
<th>Value/direction associated with better outcome</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genetics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family history affective disorder</td>
<td>Present</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Family history of schizophrenia</td>
<td>Absent</td>
<td>1, 2, 4, 5, 6</td>
</tr>
<tr>
<td>Premorbid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizoid personality</td>
<td>Absent</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Loner</td>
<td>Absent</td>
<td>5, 6</td>
</tr>
<tr>
<td>Social functioning</td>
<td>Better</td>
<td>7</td>
</tr>
<tr>
<td>Heterosexual functioning</td>
<td>Better</td>
<td>7</td>
</tr>
<tr>
<td>Married</td>
<td>Ever</td>
<td>1, 2</td>
</tr>
<tr>
<td>Skills &amp; interests</td>
<td>Present</td>
<td>5, 6</td>
</tr>
<tr>
<td>Work functioning</td>
<td>Better</td>
<td>1, 2</td>
</tr>
<tr>
<td>IQ</td>
<td>Higher</td>
<td>1, 2</td>
</tr>
<tr>
<td>Illness onset</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute (&lt; 6 months)</td>
<td>Present</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Precipitating factors</td>
<td>Present</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Psychopathology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressive features</td>
<td>Present</td>
<td>1, 2, 3, 5, 6</td>
</tr>
<tr>
<td>Hallucinations, delusions</td>
<td>Present</td>
<td>8</td>
</tr>
<tr>
<td>Absent</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Disorganized thought</td>
<td>Absent</td>
<td>7, 8</td>
</tr>
<tr>
<td>Assaultiveness</td>
<td>Absent</td>
<td>5, 6</td>
</tr>
<tr>
<td>Emotional blunting</td>
<td>Absent</td>
<td>1, 2</td>
</tr>
<tr>
<td>Confusion or perplexity</td>
<td>Present</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Concern with death</td>
<td>Present</td>
<td>7</td>
</tr>
<tr>
<td>Subjective distress</td>
<td>Present</td>
<td>7</td>
</tr>
<tr>
<td>Anxiety</td>
<td>Present</td>
<td>7</td>
</tr>
<tr>
<td>Course (before admission)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Process</td>
<td>Absent</td>
<td>1</td>
</tr>
<tr>
<td>Duration of prior hospitalizations</td>
<td>Shorter</td>
<td>7, 9</td>
</tr>
</tbody>
</table>

1 = Stephens et al (1966)
2 = Stephens (1978)
3 = Vaillant (1964a, 1964b)
4 = Stephens & Astrup (1963)
5 = McGlashan (1986c).
6 = Fenton & McGlashan (1987a)
7 = Carpenter et al (1987)
8 = Tsuang et al (1981)
Table 4. Diagnostic systems tested for predictive validity by long-term followup (FU)

<table>
<thead>
<tr>
<th>System</th>
<th>FU study or sample</th>
<th>System stringency in given sample: 1 = broad, 5 = narrow</th>
<th>Correlation with poor functional outcome</th>
<th>Correlation with defect state</th>
<th>Diagnostically stable over FU</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Haven Schizophrenia Index</td>
<td>Alberta 2 Phipps</td>
<td>1</td>
<td>Weak</td>
<td>Weak</td>
<td>No</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td></td>
<td>Chestnut Lodge</td>
<td>1</td>
<td>Weak</td>
<td>Weak</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Schneider's First Rank Symptoms</td>
<td>Alberta 2 Phipps</td>
<td>1</td>
<td>Weak</td>
<td></td>
<td></td>
<td>1, 2, 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>Weak</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Taylor &amp; Abrams (1978)</td>
<td>Phipps</td>
<td>2</td>
<td>Weak</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Feighner definite</td>
<td>Alberta 2 Phipps</td>
<td>5</td>
<td>Strong</td>
<td>Strong</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Iowa 500 Phipps</td>
<td>5</td>
<td>Strong</td>
<td>Strong</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Iowa 500 Phipps</td>
<td>5</td>
<td>Strong</td>
<td>Strong</td>
<td>Yes</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Chestnut Lodge</td>
<td>5</td>
<td>Strong</td>
<td>Strong</td>
<td>Yes</td>
<td>4</td>
</tr>
<tr>
<td>Feighner probable (without duration criterion)</td>
<td>Alberta 2 Phipps</td>
<td>1</td>
<td>Moderate</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Non-Feighner or atypical schizophrenia</td>
<td>Iowa 500 Phipps</td>
<td>2</td>
<td>Moderate</td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Research Diagnostic Criteria</td>
<td>Phipps</td>
<td>4</td>
<td>Strong</td>
<td>Moderate</td>
<td>Intermediate</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Chestnut Lodge</td>
<td>3</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Intermediate</td>
<td>4</td>
</tr>
<tr>
<td><em>DSM-III</em></td>
<td>Phipps</td>
<td>5</td>
<td>Strong</td>
<td>Moderate</td>
<td>Intermediate</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Chestnut Lodge</td>
<td>3</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Intermediate</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Vermont State</td>
<td>3</td>
<td>Weak</td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Hospital Phipps</td>
<td>3</td>
<td>Moderate</td>
<td></td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

Notes:
- System stringency: 1 = broad, 5 = narrow
- Correlation with poor functional outcome: Weak, Strong
- Correlation with defect state: Weak, Strong
- Diagnostically stable over FU: No, Yes
- Citation: 1, 2, 3, 4, 5, 6, 7
Table 4. Diagnostic systems tested for predictive validity by long-term followup (FU)—Continued

<table>
<thead>
<tr>
<th>System</th>
<th>Predictive validity tested</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FU study or sample</td>
</tr>
<tr>
<td>12-point Flexible System</td>
<td>Phipps Vermont State Hospital</td>
</tr>
<tr>
<td>Hospital Diagnosis</td>
<td>DSM-I schizophrenia</td>
</tr>
<tr>
<td>2 = Bland &amp; Om (1980a, 1980b)</td>
<td></td>
</tr>
<tr>
<td>4 = McGlashan (1984a)</td>
<td></td>
</tr>
<tr>
<td>5 = Tsuang et al (1979)</td>
<td></td>
</tr>
<tr>
<td>6 = Tsuang et al (1981)</td>
<td></td>
</tr>
<tr>
<td>7 = Tsuang &amp; Fleming (1986)</td>
<td></td>
</tr>
<tr>
<td>8 = Harding et al (1987a, 1987b)</td>
<td></td>
</tr>
</tbody>
</table>

As demonstrated in table 4, there exists a rough positive correlation between diagnostic stringency and predictive power. The Feighner et al. definite system is clearly both the most narrow and most predictive of long-term diagnostic stability and poor outcome. The Feighner et al. system, like the Research Diagnostic Criteria, DSM-III, and Feighner probable systems, excludes patients with prominent affective symptoms. Like DSM-III, it excludes patients with a LOMI of less than 6 months. Like no other system, its criteria include items usually regarded as poor prognostic predictors like family history of schizophrenia, single marital status, or poor premorbid work/social functioning. Other diagnostic systems, including DSM-III, are less stringent and, accordingly, demonstrate a weaker correlation with long-term outcome and diagnostic stability.

Overall, the studies in table 4 do not dispute the nosologic principle that the correlation of any diagnostic system to long-term outcome/diagnostic stability is proportional to the number of prognostic variables included as criteria (Stephens 1978; McGlashan 1984c). These prognostic variables may be defined as anything other than nonaffective cross-sectional signs and symptoms. Often they consist of factors that can be regarded as postbaseline—that is, as outcomes of the illness process itself like functional inferiority or well-established psychopathology. Such criteria maximize the homogeneity of diagnostic samples. This may be an advantage for certain kinds of research, but it also renders the test of predictive validity less applicable and more an exercise in tautology.

The Vermont State Hospital findings are rather unique vis-a-vis schizophrenic nosology (Harding et al. 1987a, 1987b). The finding that long-term followup failed to distinguish between schizophrenic and nonschizophrenic patients challenges the assumption that outcome can specify or differentiate major nosologic entities. Furthermore, the finding that long-term outcome was no different for the DSM-III and the DSM-I schizophrenic patients challenges the utility of predictive validity for testing between different definitions of the same disorder. That is, it questions the assumed advantage of operationalized diagnosis. These findings may reveal that the forces of chronicity and institutionalization, when extant with sufficient length and strength, erode nosologic differences and transform distinctive
psychopathologies into a homogeneous mass that defies differentiation with existing assessment methodologies. At the least, these interesting findings question many of our cherished assumptions, like the two-psychoses hypothesis of Kraepelin, and invite us to think again.

The long-term followup studies reviewed here support a broad definition of schizophrenia. Stringent systems, and/or those that incorporate prognostic dimensions as diagnostic criteria, may select narrower and smaller samples with fewer false positives. But they also exclude many who probably have the disorder. Schizophrenia is heterogeneous. Some patients with the disorder do well at followup and should not, for this reason alone, be regarded as misdiagnosed. Good outcome schizophrenia is not, "in reality," an affective disorder, nor should it even be recategorized as such without compelling evidence from several domains in addition to long-term course. Other patients look good at baseline—the acute, remitting, good prognosis cases. Long-term followup, however, demonstrates repeatedly that a distinct proportion of these develop a chronic remitting or chronic unremitting course. For this reason, separating these patients off into diagnostic categories other than schizophrenia and its subtypes is also premature. Overall, until significant progress is made in reducing the heterogeneity of schizophrenia, its criteria should err in the direction of inclusiveness.

Implications

In summary, the North American followup studies have confirmed that schizophrenia is a chronic disease, frequently disabling for a lifetime and with an outcome generally worse than that of other major functional mental illnesses. While schizophrenia can be lethal, especially vis-a-vis increased suicide risk, the process does not appear to be progressively dementing as originally described. On the average, functional deterioration appears to plateau and even to relent somewhat after 5-10 years of manifest illness. Outcome overall is heterogeneous, but much of this bewildering variance can be accounted for by sample characteristics. Many of these characteristics are linked with or are expressions of psychopathology, such as diagnostic type of schizophrenia (broad vs. narrow criteria), subtype, comorbidity, or dimensions of chronicity (LOMI, biological treatment resistance, age of onset, institutionalization). Other characteristics are more orthogonal to psychopathology such as demography (gender, marital status, socioeconomic status, physical setting) and the sample's ratings on multiple "nondiagnostic" predictor variables, especially premorbid health. The careful assessment and reporting of such sample characteristics is vital for finding meaning amidst melange.

The North American long-term followup studies have not answered any questions about treatment. They have, however, provided some interesting observations. The common occurrence of good outcome without treatment points to the existence of heterogeneity vis-a-vis treatment need and/or response. This, plus the notion of biological treatment resistance, suggests that fruitful advances in understanding schizophrenia may accrue by tracking how drug responsiveness segregates meaningful subgroups. The North American long-term followup studies also suggest that schizophrenia may be quite malleable to prolonged environmental/psychosocial perturbations. These have negative potential if applied too intensively or ambitiously, but positive potential if applied steadily in a supportive, rehabilitative mode in the context of stable and unlimited continuity of care.

This essay, it is hoped, has made the point by example that the more completely samples are described, both at baseline and followup, the more results can be compared meaningfully across studies. Accordingly, while future followup studies should try to incorporate as many of the outlined elements of methodology as possible, accurate and reliable descriptions of sample, diagnosis, and outcome are the sine qua non. At the very least, sample descriptions should include those variables listed in tables 2 and 3. Baseline data should also provide sufficient sign and symptom data for study patients to be diagnosed by any of the systems in table 4 or re-diagnosed by any system for schizophrenia of the future. Since comorbidity is the rule rather than the exception, diagnostic assessment should be dimensional and inclusive, not categorical and exclusive. Outcome, finally, should be recognized as a multivariate construct. Like diagnosis and prognosis, it requires operationalization and calibrated specification for reliable assessment by trained observers.

The North American followup studies have taught us a great deal about schizophrenia thus far. Most important, they have taught us how to construct future studies that maximize a return of meaning.
References


Announcement

The annual meeting of the New Clinical Drug Evaluation Unit (NCDEU), sponsored by the Pharmacologic and Somatic Treatment Research Consortium, Division of Clinical Research of the National Institute of Mental Health (NIMH), will take place on May 26–28, 1987, in Key Biscayne, FL.

This will be the 27th annual meeting of the NCDEU program. The program was established in 1959 by the NIMH as the Early Clinical Drug Evaluation Unit (ECDEU) program with 15 grantees. It has developed into the current format of planned symposia and free communications. Sessions focus on recent research findings in psychopharmacology, methodologic problems and advances in clinical trials in this area, and updates or reviews of significant issues in the field.

For further information, please contact:

Elaine G. Pearl
NCDEU Coordinator
Pharmacologic and Somatic Treatments Research Consortium, NIMH
Room 10C-06
5600 Fishers Lane
Rockville, MD 20857