One Hundred Years of Schizophrenia: A Meta-Analysis of the Outcome Literature

James D. Hegarty, M.D., M.P.H., Ross J. Baldessarini, M.D., Mauricio Tohen, M.D., Dr.P.H., Christine Waternaux, Ph.D., and Godehard Oepen, M.D.

Objective: *This study was undertaken to assess the twentieth-century literature on outcome in schizophrenia for historical trends that might be associated with changes in diagnostic and therapeutic practice and to test the hypothesis that both improved biological treatment and changes in diagnostic criteria have influenced outcome.* *Method:* *Meta-analysis of the international literature on outcome in schizophrenia or dementia praecox from 1895 to 1992 identified 821 studies; 320 of these, with 51,800 subjects in 368 cohorts, met the inclusion criteria for the study.* *Results:* *Only 40.2% of patients were considered improved after follow-ups averaging 5.6 years (range=1–40). Outcome was significantly better when patients were diagnosed according to systems with broad criteria (46.5% were improved) or undefined criteria (41.0% were improved) rather than narrow criteria (27.3% were improved). The proportion of patients who improved increased significantly after mid-century (for 1956–1985 versus 1895–1955, 48.5% versus 35.4%), probably reflecting improved treatment as well as a broadened concept of schizophrenia. However, in the past decade, the average rate of favorable outcome has declined to 36.4%, perhaps reflecting the re-emergence of narrow diagnostic concepts.* *Conclusions:* *Overall, less than half of patients diagnosed with schizophrenia have shown substantial clinical improvement after follow-up averaging nearly 6 years. Despite considerable gains in improvement rates after mid-century, there has been a decline since the 1970s. These historical changes probably reflect improved treatment, shifts in diagnostic criteria, and selection bias related to changes in health care.*

(Am J Psychiatry 1994; 151:1409–1416)

Presented in part at the 145th annual meeting of the American Psychiatric Association, Washington, D.C., May 2–7, 1992. Received Nov. 3, 1992; revisions received Aug. 18, 1993, and Jan. 25, 1994; accepted Feb. 7, 1994. From the Consolidated Department of Psychiatry and the Neuroscience Program, Harvard Medical School, Boston; the Departments of Epidemiology and Biostatistics, Harvard School of Public Health, Boston; and the Laboratories for Psychiatric Research and the Bipolar and Psychotic Disorders Program, Mailman Research Center, McLean Division of the Massachusetts General Hospital. Address reprint requests to Dr. Hegarty, McLean Hospital, 115 Mill St., Belmont, MA 02178.

The authors thank Dr. Franca Centorrino and Dr. Elizabeth M. Sajdel-Sulkowska for assistance with translation of non-English documents and Thelma Charen, M.A., Lyn Dietrich, M.S.L.S., Debbie Pennino, Mila Cason, M.A., and Jeanmarie O'Donnell, B.S., for bibliographic assistance.

Supported in part by NIMH grants MH-31154, MH-47370, and MH-48444 and by a Young Investigator Award from the National Alliance for Research on Schizophrenia and Depression.

Clinical course and outcome in schizophrenia (formerly dementia praecox) have been the subject of extensive investigation through this century. Kraepelin (1) considered dementia praecox to be a chronic or progressive illness leading to severe impairments in cognitive and social functioning, with findings of clinical improvement during follow-up in not more than 17% of cases. While the progress or course of the illness varied, a poor outcome was considered to be almost inevitable eventually or was inherent in the diagnostic conceptualization. For example, in a large early survey of 395 hospitalized patients in 1912, Stearns (2) lamented the "apparent hopelessness of the disease dementia praecox." Early treatments, such as injection of adrenaline (3), inducing fever (4), and vasectomy (5),

proved wholly ineffective, and pessimistic conclusions dominated the prognostic literature for decades. In the 1930s the introduction of insulin coma (6), chemoconvulsive therapy (7), and ECT (8) provided the first effective treatments for psychotic disorders. These more effective interventions, as well as the growth of a broader conceptualization of schizophrenia introduced by Eugen Bleuler in 1911 (9), were widely adopted, and the prognosis for schizophrenia began to appear less bleak by the late 1930s (10). The development of specific antipsychotic medicines, beginning with reserpine and chlorpromazine in the 1950s (11–15), introduced a new era in treatment, and hopes soared for improved outcomes as well as short-term benefits in acute phases of psychotic illness (16–18).

During the 1970s two large, long-term outcome studies (19, 20) found significant clinical improvement in nearly 50% of schizophrenic patients after 10 years. These reports added to a growing literature supporting an improved prognosis for patients with schizophrenia in the contemporary era (21–25). Credit was given to improved treatment as well as deinstitutionalization and new clinical appreciation of psychosocial factors in chronic psychotic disorders, including family relations and expressed emotion (26). It may also be important that most of these studies (19–26) used relatively broad or cross-sectional diagnostic criteria, in contrast to the most recent American diagnostic methods (DSM-III and DSM-III-R), which attend to duration of symptoms and course of illness in what might be termed a "neo-Kraepelinian" trend. Therefore, it is of interest to address historical trends in studies of outcome in schizophrenia and to attempt to ascertain the impact of diagnostic and treatment methods.

Reviews of some of the literature on outcome in schizophrenia have been carried out by others (27–32), but to date there have been no studies incorporating the available international outcome data in a meta-analysis with explicit inclusion and exclusion criteria. Nor has the possible impact of evolving diagnostic concepts been incorporated into such analyses. Accordingly, we undertook the current study, applying recommended contemporary methods of meta-analysis for clinical studies (33) and considering available information on diagnosis and treatment in particular.

METHOD

We undertook a comprehensive review of the twentieth-century literature regarding outcome in schizophrenia. Studies published before 1966 were identified by a title search of *Index Medicus* under the headings Dementia Praecox and Schizophrenia, with the subheadings outcome, course, and prognosis. Historically, literature on schizophrenia was indexed under the subject heading Dementia Praecox for articles listed in the *Current List of Medical Literature* from 1950 to 1953 and under Schizophrenia for articles listed in the *Current List of Medical Literature* from 1954 to 1959 and in *Index Medicus* from 1960 to the present (34–38). Studies published since 1966 were identified through a MEDLINE computerized literature search in which the same title search criteria were used. The entire search encompassed the years 1895 through 1991 and yielded 686 primary reports; additional references (N=135; 16.5%) found in primary reports and meeting the title search criteria or identified as outcome studies also were examined, to provide a total of 821 reports for preliminary consideration.

Inclusion Criteria

Studies identified by the title search were included for analysis if they met the following criteria. 1) They included only patients with diagnoses of schizophrenia or dementia praecox or gave outcome data specific to these diagnostic groups; studies mixing schizophrenic patients with patients who had manic-depressive, schizoaffective, or other psychoses were excluded. 2) The mean length of follow-up was at least 1 year (i.e., cross-sectional or census studies were excluded). 3) Less than 33% of the subjects were lost to follow-up, and there were 15 or more patients at completion of follow-up. 4) Subjects were not selected a priori for good or poor outcome. 5) Explicit numerical data on clinical, social, or vocational outcome were available, so that patients could be assigned to discrete outcome groups (i.e., studies that used correlative relationships or the category "discharged from hospital and not readmitted" were excluded). 6) Both experimental and naturalistic studies had to specify the treatment used throughout the follow-up.

Of the total of 821 studies preliminarily identified by the search criteria, 260 were excluded because they were not outcome studies and eight reports could not be located; the 553 remaining reports were reviewed in detail. The exclusion rate for studies identified by the *Index Medicus* search (1895–1965) was lower (39.5%) than the rate for later studies (1966–1992) found by the computerized MEDLINE search (69.0%), with a moderate correlation between exclusion rate and decade of study ($r=0.56$, $df=8$, n.s.). Most studies excluded after 1965 were eliminated because the follow-up was less than 1 year.

Overall, 320 studies (57.9%) met all of the inclusion criteria for analysis; they represent 51,800 subjects from 368 cohorts and 311,400 person-years of follow-up. Non-English-language reports of studies (N=85) that were translated for this analysis included 42 in German, 12 in French, nine in Italian, seven in Danish, five in Norwegian, three in Polish, three in Russian, three in Spanish, and one in Finnish. In addition, 70 reports on non-English-speaking cohorts were published in English. The translated studies that met the inclusion criteria (N=37) represent 11.6% of the 320-study database. (A complete 327-reference bibliography is available from the first author on request.)

Definition of Outcome

The difficulty of accurately and consistently defining clinical outcome in schizophrenia has been reviewed (39). Given the varied criteria used to express outcome in the hundreds of studies we considered, a generalized standard was required. Our determination of outcome made use of clinical, social, or vocational criteria as reported in the studies analyzed. By consensus, we translated each study's outcome data to ratio measures of improvement (the number of patients considered improved divided by the total number per cohort), using criteria that suggested attainment of substantial levels of functioning and freedom from psychotic symptoms. Thus, patients considered as "improved" in follow-up had to have been described as recovered, in remission, well without residual symptoms, minimally or mildly symptomatic, improved without significant deficit, socially recovered, or working or living independently.

Diagnostic Systems

If specific diagnostic criteria were defined, the studies were assigned to the schizophrenia diagnostic systems of the following: Kraepelin (dementia praecox) (1), Langfeldt (40), Bleuler (9), Schneider (first-rank symptoms) (41), Leonhard (42), Mayer-Gross (43), ICD-8, ICD-9, DSM-II, DSM-III, DSM-III-R, the Present State Examination (PSE) (44), the Feighner criteria (45), and the Research Diagnostic Criteria (RDC) (46). These diagnostic systems were then

assigned to either the so-called narrow category (Kraepelinian, with some implication for long-term course of illness) or broad category (non-Kraepelinian) (47). Diagnostic systems requiring at least 6 months of illness before diagnosis of schizophrenia were considered Kraepelinian and include those of Kraepelin, Langfeldt, and Feighner, as well as DSM-III and DSM-III-R. Non-Kraepelinian systems were defined as more symptom-based, cross-sectional diagnostic approaches with a brief duration requirement (≤1 month) or no requirement for duration of prodromal or active psychotic illness. These include Bleuler's schizophrenia, Leonhard's systematic schizophrenia, Schneider's first-rank criteria, Mayer-Gross's schizophrenia, ICD-8, ICD-9, DSM-II, the PSE/CATEGO system, and the RDC. Studies that did not use clearly defined diagnostic systems were categorized as having unspecified criteria.

Stratification by Decade, Length of Follow-Up, and Treatment

Only cohorts with ≤10 years of follow-up (314 cohorts, or 85.3% of the total of 368) were used for inter-decade comparisons. The mean year of follow-up was calculated as the average of the first and last years of the follow-up period, and decades were defined by their midpoint (e.g., decade 1950 included mean follow-up years falling between 1946 and 1955). Stratification for duration of follow-up included 1 to <2, 2–4, 5–9, 10–19, and ≥20 years. On the basis of treatment used throughout the period of follow-up, four primary treatment groups were defined: 1) nonspecific (placebo trials, psychotherapy, hydrotherapy, fever therapy, and nonneurological surgery), 2) convulsive (electrically or chemically induced seizures or insulin shock), 3) prefrontal lobotomy, and 4) neuroleptic or atypical antipsychotic drugs.

Statistical Analysis

Data are presented as unweighted means and standard deviations unless stated otherwise. For multiple comparisons, means were compared by Duncan's multiple range test (48). A weighted multiple regression model for predicting percentage of patients improved included the following variables: diagnostic system, type of treatment, and length of follow-up. Weights based on number of subjects, with an estimate of variance, were computed for each study by the method of DerSimonian and Laird (49), and overall variance is given as the standard error of the mean (SE). To ensure homogeneity of variance in parametric tests, data were subjected to an arc sine variance-stabilizing transformation (50). Use of standard analysis of variance (ANOVA) and post hoc tests is cited in the Results section. All p values are based on two-tailed tests, with nonsignificance defined as ($p \geq 0.05$).

RESULTS

Table 1 summarizes characteristics of the outcome studies included in this analysis. The number of cohorts of psychotic patients that met the study selection criteria per decade (excluding the incomplete current decade, 1985–1995), averaged 38.7 (SD=27.1), with a maximum of 94 during the 1930s, following the introduction of insulin coma and ECT. The average number of subjects per cohort studied declined significantly over the century, from 250 in 1895–1915 to 56 since 1986 (by decade, $r=-0.98$, N=368, $p<0.0001$). The diagnosis of schizophrenia or dementia praecox in this century was made without explicit diagnostic criteria (unspecified) in 50.0% of the studies meeting the selection criteria. Kraepelinian diagnostic systems were applied in 21.7% of the studies, and specified non-Kraepelinian methods appeared in 28.3%. The duration of follow-up ranged from 1 to 40 years (median=3.0 years) and averaged 5.6 years (SD=6.5). Most of the studies involved relatively short-term follow-up (<5 years in 57.0%). Treatment predominantly involved neuroleptics, followed by convulsive and nonspecific methods and lobotomy.

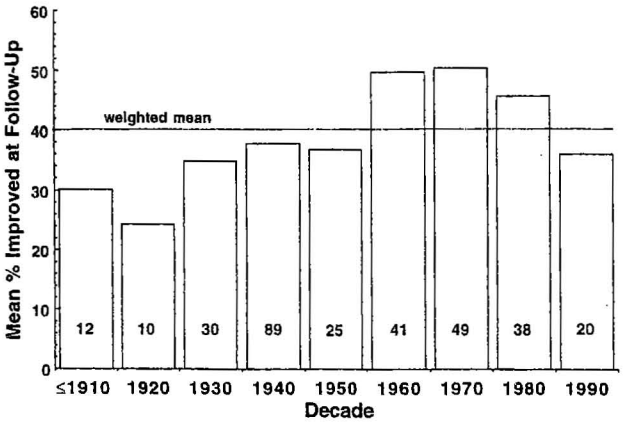
The mean proportion of patients showing a favorable outcome, independent of length of follow-up, across the entire century was 40.2% (SD=17.6%) (N=368 cohorts). A decade-by-decade summary (figure 1) supports the further impression that outcome shifted over

TABLE 1. Characteristics of the Schizophrenia Outcome Studies Analyzed

Item	Cohorts (N=368) N	%
Decade in which studies were completed		
1895–1905	6	1.6
1906–1915	6	1.6
1916–1925	13	3.5
1926–1935	37	10.1
1936–1945	94	25.5
1946–1955	36	9.8
1956–1965	51	13.9
1966–1975	63	17.1
1976–1985	42	11.4
1986–1992	20	5.4
Study location[a]		
United States	137	42.8
United Kingdom	67	20.9
Scandinavia	44	13.7
Germany	24	7.5
Switzerland	11	3.4
Japan	8	2.5
China	5	1.6
Canada	5	1.6
India	4	1.2
Other	15	4.7
Diagnostic system used		
Kraepelinian	80	21.7
DSM-III, DSM-III-R	20	25.0
Feighner	9	11.2
Kraepelin	41	51.2
Langfeldt	10	12.5
Non-Kraepelinian	104	28.3
Bleuler	18	17.3
DSM-II	9	8.7
ICD-8, ICD-9	16	15.4
Leonhard	3	2.9
Mayer-Gross	5	4.8
Present State Examination	33	31.7
Research Diagnostic Criteria	10	9.6
Schneider	10	9.6
Unspecified	184	50.0
Years of follow-up		
1 to <2	98	26.6
2–4	112	30.4
5–9	89	24.2
10–19	48	13.0
≥20	21	5.7
Treatment		
Neuroleptic	173	47.0
Convulsive	101	27.4
Lobotomy	13	3.5
Nonspecific	81	22.0

[a]The numbers for study location are of studies, not cohorts.

FIGURE 1. Mean Percentages of Schizophrenic Patient Cohorts Considered Improved in Follow-Ups of ≤10 Years, by Decades of Studies Reviewed[a]



[a]Horizontal line indicates weighted overall mean for the century (40.1%; N=314 cohorts). Numbers inside bars indicate number of cohorts for each decade. Decades are defined by the midpoint of the decade (e.g., the 1950 decade is 1946–1955). The year of a study was defined as the mean of the first and last years of follow-up (≤1910 includes all studies from 1895 through 1915). There was significant variation in outcome by decade ($F=5.9$, df=8, 305, $p<0.0001$).

time. Prior to 1920, studies consistently involved patients diagnosed as having Kraepelin's dementia praecox, for whom no specific treatment was available. Accordingly, between 1895 and 1925, the weighted average proportion of patients who were considered improved at the end of follow-up was only 27.6% (SE=3.3%). In the 1930 decade, when convulsive therapies were introduced, mean outcomes tended to improve to 34.9% (SE=3.1%). This trend continued after mid-century, with a peak mean in the period 1956–1985 (48.5%, SE=1.7%), which was significantly different from the mean for 1895–1955 (35.4%, SE=1.3%) ($t=5.7$, df=292, $p<0.0001$), perhaps reflecting the introduction of modern antipsychotic drug therapy (51, 52). However, since 1986 the mean likelihood of a favorable outcome has diminished to only 36.4% (SE=4.4%), or a level that is statistically indistinguishable from that found in the first half of the century ($t=0.2$, df=146, n.s.).

Duration of follow-up also was examined as a predictor of clinical outcome. The unweighted likelihood of a favorable outcome tended to decline with longer follow-up, but only slightly and nonsignificantly: 43.7% (SD=21.3%) improved at 1 to <2 years, 40.8% (SD=19.2%) at 2–4 years, 38.3% (SD=16.4%) at 5–9 years, 37.4% (SD=18.6%) at 10–19 years, and 37.3% (SD=13.6%) at ≥20 years ($F=1.5$, df=4, 363, n.s.).

More striking differences were found after stratification for treatment group and diagnostic system (table 2). Unweighted analysis found neuroleptic treatment to be associated with the highest mean percentage of improved subjects (45.6%), followed by convulsive and nonspecific treatments and prefrontal lobotomy. A weighted ANOVA found a highly significant effect of diagnostic system and the impact of all treatments combined, with no significant interaction between treatment and diagnostic system. Post hoc analysis indicated that the combined results of neuroleptic and convulsive treatments were better than those of lobotomy and nonspecific methods.

Studies using broad non-Kraepelinian criteria consistently showed better outcomes than those using narrow Kraepelinian criteria (ratio=1.70; $t=8.3$, df=178, $p<0.0001$) (table 2). This difference held throughout the century—from the early era before effective somatic therapies (up to and including 1925), through the introduction of convulsive therapies (1926–1955), through the introduction of neuroleptics (1956–1975), to the contemporary era (after 1975)—and was highly significant by ANOVA (figure 2).

Finally, a weighted multiple regression model was applied to examine diagnostic system, duration of follow-up, and neuroleptic treatment as predictors of improvement. Diagnostic system was a significant predictor of outcome ($F=24.4$, df=2, 360, $p<0.00001$), as was neuroleptic treatment ($F=31.3$, df=1, 360, $p<0.00001$), but duration of follow-up was not ($F=2.0$, df=4, 360, $p=0.09$).

DISCUSSION

The extensive data we analyzed support several general impressions regarding clinical outcome in schizophrenia during the twentieth century. The main impression is that less than one-half of patients diagnosed with this disorder have shown substantial clinical improvement after follow-up averaging nearly 6 years (table 2). Although there were substantial gains in favorable outcome after the 1920s, rates of improvement declined considerably after the 1970s (figure 1). These biphasic trends may reflect the impact of more effective treatments and historical changes in diagnostic criteria during the past century (table 2 and figure 2), as well as possible effects of sampling bias.

Between the 1920s and the 1970s, favorable outcome in schizophrenia (patients rated clinically improved) doubled, from an average of 24.3% to 50.5% of cases (figure 1). Such gains have been attributed to new treatments (17, 18), particularly neuroleptics since the 1950s (51, 52), and an increased emphasis on family and community interventions since the 1960s (26). In the late 1970s Manfred Bleuler (19) noted the apparent disappearance of the most severe, or catastrophic, forms of schizophrenia and proposed that modern treatment might be responsible. It has been hypothesized that early and consistent treatment with antipsychotic medications may favorably alter the natural course of schizophrenia (51), and a recent analysis by Wyatt (52) of 22 follow-up studies involving first-episode patients (21 of which applied broad diagnostic criteria) found that early intervention with neuroleptics increased the likelihood of a favorable long-term course in such patients. The poor outcome long associated with insidi-

TABLE 2. Percentages of Cohorts of Patients in Schizophrenia Outcome Studies Who Had Improved at Follow-Up, by Treatment and Diagnostic System[a]

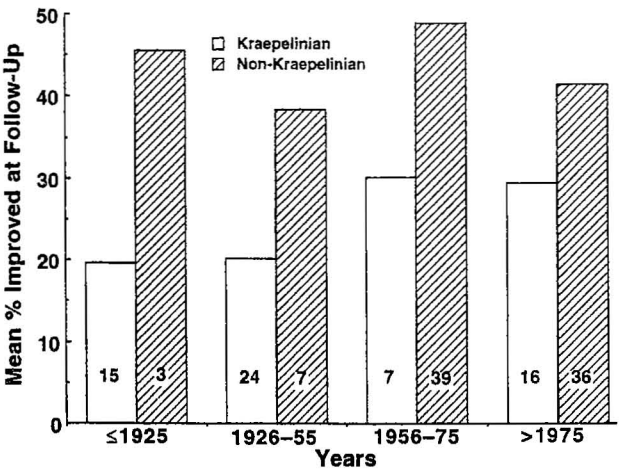
	Cohorts Diagnosed by Kraepelinian Systems			Cohorts Diagnosed by Non-Kraepelinian Systems			Cohorts Diagnosed by Unspecified Criteria			All Cohorts		
		Percentage of Patients Who Improved			Percentage of Patients Who Improved			Percentage of Patients Who Improved			Percentage of Patients Who Improved	
Treatment	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Neuroleptic	37	31.2	17.3	87	48.0	17.6	51	51.9	16.6	175	45.6	18.8
Convulsive	13	26.9	13.8	7	42.2	19.8	81	44.0	18.1	101	41.7	18.5
Nonspecific	30	22.5	11.1	10	34.0	17.5	39	32.4	14.6	79	28.9	14.5
Lobotomy	—	—	—	—	—	—	13	28.1	13.9	13	28.1	13.9
Total	80	27.3	15.1	104	46.5	17.3	184	41.0	16.9	368	40.2	17.6

[a]Weighted analysis of variance found a significant effect for treatment ($F=14.6$, $df=3, 364$, $p<0.0001$) and diagnostic system ($F=28.8$, $df=2, 365$, $p<0.0001$) and no significant interaction between them ($F=1.1$, $df=4, 362$, n.s.). Post hoc analysis indicated that the overall results with neuroleptic and convulsive treatments were better than with lobotomy and nonspecific methods (estimated difference=15.2, SE=2.1, $p<0.0001$) and that outcome in studies using Kraepelinian diagnostic criteria yielded lower improvement rates than studies using non-Kraepelinian criteria (estimated difference=19.2, SE=2.0, $p<0.0001$).

ous-onset schizophrenia may result, in part, from delay in the use of antipsychotic medication. The results of our review suggest, further, that patients diagnosed according to broad or unspecified criteria have shown greater improvement in outcome in the neuroleptic era since the 1950s (table 2 and figure 2). These findings are consistent with the view that changes in the diagnostic criteria for schizophrenia, as well as new treatments, probably account for the improved outcomes reported in mid-century (1930s–1970s).

The dominant diagnostic trend in the 1930s and 1940s was away from the narrow Kraepelinian model, which emphasized chronicity and a debilitating course, and toward a broader Bleulerian concept, focusing on the cross-sectional presentation of particular psychotic symptoms. Notably, Schneider (41, 53) delineated 11 "symptoms of the first rank" that were purported to be pathognomonic of schizophrenia, independent of their duration. There also was growing interest in the psychodynamic components or proposed determinants of psychosis (54). The presumed irreversibility of Kraepelin's dementia praecox was debated, and psychoanalytic or other psychotherapeutic cures were reported (55–57). The result of the reconceptualization from dementia praecox to schizophrenia was that by the 1950s, the diagnosis of schizophrenia had broadened in many centers to include patients who would previously have been diagnosed with manic-depression, involutional psychosis, or other disorders (58, 59). A further impetus to broaden diagnostic criteria may have been the introduction of specific antipsychotic medications. Physicians' inclination to see patients as treatable by a new and effective form of therapy may have encouraged the diagnosis of schizophrenia (60) and contributed to favorable clinical outcomes when relatively neuroleptic-responsive patients or those with disorders with a favorable natural history received antipsychotic medication in the 1950s and 1960s. Similarly, the later introduction of mood-stabilizing agents as well as changes in diagnostic criteria appear to have contributed to a decline in the diagnosis of schizophrenia since 1970 (61).

FIGURE 2. Mean Percentages of Schizophrenic Patient Cohorts Considered Improved in Follow-Ups of ≤10 Years Who Were Diagnosed According to Kraepelinian or non-Kraepelinian Systems, by Eras of Studies Reviewed[a]



[a]Numbers inside bars indicate number of cohorts for each era (total N=147). The difference between outcomes across eras was highly significant with respect to diagnosis ($F=15.2$, $df=1, 142$, $p<0.0001$), with no significant effect of era ($F=1.6$, $df=8, 142$) or interaction of era and diagnosis ($F=0.5$, $df=3, 139$).

By 1970 the shift to a broad, symptom-oriented definition of schizophrenia was coming under increasing criticism (58). A landmark study (59) highlighted the dissimilar diagnostic practices of New York and London psychiatrists of that era, particularly as regards schizophrenia. This work heralded a return to a narrower definition of schizophrenia initiated in the United States by the Feighner-Washington University criteria (45), which later had an influence on DSM-III. The reintroduction of a duration-of-illness requirement—

namely, 6 months of prodromal, active psychotic symptoms or residual state symptoms, with the exclusion of affective illness—operationalized what can be considered neo-Kraepelinian criteria. As the definition of schizophrenia has evolved over the past century, from the relatively narrow or course-oriented Kraepelinian criteria to broader or cross-sectional, symptom-oriented Bleulerian criteria and then back to a neo-Kraepelinian standard (table 1), so also have treatment response and clinical outcome shifted toward more, and then fewer, favorable multi-year clinical outcomes (figure 1). That is, despite routine use of neuroleptics and other contemporary therapies, reported average clinical outcomes have deteriorated in the past decade. We propose that diagnostic changes have had an important impact on outcome in schizophrenia since the introduction of DSM-III, with its return to a more Kraepelinian emphasis on duration and course of illness for diagnosis. Additional factors that may contribute to the recent downward trend in average rates of favorable outcome include biased availability of more chronic or severe cases for study, as an effect of responsiveness to treatment or a result of evolving changes in the organization and financing of psychiatric services.

There are several limitations to the present analysis. These include a high proportion of studies with imprecise and potentially unreliable diagnostic or outcome criteria or clinically heterogeneous samples, even among those not excluded by our a priori selection criteria; the exclusion of studies not presenting crude ratio data with which to assess subject-wise outcome rates; the choices required for pooling diagnostic systems into clusters categorized as narrow (Kraepelinian), broad (non-Kraepelinian), or unspecified; and the effect of large numbers to obscure high variance in outcome among studies (overall variance, as the standard deviation divided by the mean, was 50.0%). In addition, this meta-analysis considered only biological treatments and made no attempt to assess potentially important advances in psychosocial and community-based aspects of comprehensive care (26, 62, 63).

It is possible that uncontrolled bias in studies may arise from the baseline status of the patients selected for follow-up: those first encountered in an acute episode of psychotic illness may have a greater average potential for response to treatment than more chronically ill patients. However, studies identified in this analysis rarely provided information about baseline clinical status. To limit the effect of such possible bias, we excluded short-term, acute treatment studies and included only studies with follow-up duration of at least 1 year.

The validity of a meta-analysis requires that all extant data be included (64). While we made an effort to locate and retrieve all outcome studies of dementia praecox or schizophrenia reported in this century, the search of *Index Medicus* without computer assistance (for publications prior to 1966) is prone to error, and some early studies may have been overlooked in the initial screening process. Another source of bias inherent in most meta-analyses of clinical studies is the absence of unpublished information, which is often suspected of containing a disproportion of negative findings, particularly in studies of treatment effects (65). Moreover, the need to exclude reports that did not permit computation of outcome ratio data led to exclusion of some potentially useful information (in 9.0% of the 553 studies reviewed in detail). In addition, judgments applied to categorizing outcome criteria and diagnostic methods may have limited sensitivity to detect differences among the various treatment groups and diagnostic systems compared. Nevertheless, such risk of type II (false negative) error does not reduce the significance of differences in outcome that were detected. The outcome studies identified were relatively short-term (57.0% of the cohorts were followed <5 years; overall mean=5.6 years), and data on very long-term outcome (>10 years) were not selected, so as to permit analysis of historical trends by decades.

It is important to emphasize that the data available for the present analysis were derived from *clinical* groups. Outcome studies of complete population-based samples would be preferable, but none was identified in our literature search. Subjects identified clinically may not represent fully comparable populations over time. For example, modern treatment may tend to remove persons with favorable responses to treatment from the pool of subjects followed (32, 66). However, the effects of such potential sources of bias in the assessment of outcome in schizophrenia remain for future studies. Finally, the design of this study does not permit direct analysis of the role that deinstitutionalization or other changes in the delivery of psychiatric services may play in the recent decline in favorable outcomes. It is plausible that inadequate funding of community-based services in the decades following deinstitutionalization may have diminished the likelihood of adequate care, or may have tended to add to bias toward retention of severely chronically ill persons in mental health systems (32, 67), in which most follow-up studies have been based. Such hypothesized higher recidivism rates may be driving the apparent trend toward worsening outcomes in recent years (68).

The congruence that we found between outcomes of patients diagnosed according to earlier Kraepelinian criteria and those diagnosed according to contemporary American Psychiatric Association diagnostic systems may appear to support the reliability and predictive value of such narrow criteria or, at least, their similarities. Stephens et al. (69) found that among nine diagnostic systems for schizophrenia, the DSM-III criteria demonstrated the highest correlation with clinical outcome. The DSM-III and DSM-III-R requirement of a minimum of 6 months of illness (including prodromal or residual-type symptoms) may be a particularly critical factor, tending to predict a relatively chronic course or unfavorable response to treatment in an idiopathic psychotic illness; it was re-evaluated in the DSM-IV field trials (70). The contemporary shift toward a narrow, neo-Kraepelinian concept of schizophrenia in the United States may be paralleled by the World Health

Organization's ICD-10, which includes a 1-month duration-of-illness requirement for the diagnosis of schizophrenia (71).

In conclusion, the present meta-analysis of the literature on outcome in schizophrenia through the twentieth century indicates that diagnostic criteria have had a consistent and predictable impact on outcome before and during the era of modern biomedical therapeutics. Notable trends include a significant recent downturn in the likelihood of a favorable outcome with the advent of DSM-III and changes in the provision of psychiatric services to psychotic persons, such that outcome is now approaching levels reported prior to the 1950s, despite the availability of modern treatment. It is not surprising that the use of diagnostic criteria for schizophrenia which require a relatively prolonged psychotic illness, and perhaps imply chronicity, has led to finding a less favorable long-term outcome. In addition to an effect of broad versus narrow diagnosis, the results of this analysis support a favorable impact of modern treatment, particularly the use of neuroleptic agents.

REFERENCES

1. Kraepelin E: Psychiatrie: Ein Lehrbuch für Studierende und Aerzte, 5th ed. Leipzig, University of Leipzig, 1899
2. Stearns AW: The prognosis in dementia praecox. Boston Med Surg J 1912; 167:158–160
3. Stuurman FJ: Experimental adrenaline injections in insane patients. Nederl Maandschr Geneesk (Leiden) 1922; 11:44–60
4. Menzies D: Pyrotherapy in dementia praecox. Lancet 1935; 2: 992–996
5. Stern M, Folsom RP, Ritter IS: Vasectomy and its influence upon 100 cases of dementia praecox at Manhattan State Hospital. State Hosp Q 1925; 10:404–412
6. Sakel M: Schizophreniebehandlung mittels insulin-hypoglykämie sowie hypoglykämischer schock. Wien Med Wchnschr 1935; 84:1211–1215
7. von Meduna L: Über experimentelle campherepilepsie. Arch Psychiatrie 1934; 102:333–339
8. Cerletti U, Bini L: Un nuevo metodo di shockterapie, "l'elettroshock." Boll Acad Med Roma 1938; 64:136–138
9. Bleuler E: Dementia Praecox or the Group of Schizophrenias (1911). Translated by Zinkin J. New York, International Universities Press, 1950
10. von Meduna L, Friedman E: The convulsive-irritative therapy of the psychoses: a survey of more than 3,000 cases. JAMA 1939; 112:501–509
11. Lancaster NP, Jones DH: Chlorpromazine and insulin in psychiatry. Br Med J 1954; 2:565–566
12. Lehmann HE, Hanrahan GE: Chlorpromazine: new inhibiting agent for psychomotor excitement and manic states. AMA Arch Neurol Psychiatry 1954; 72:227–232
13. Goldman D: Treatment of psychotic states with chlorpromazine. JAMA 1955; 57:1274–1276
14. Altschule MD: Use of chlorpromazine and reserpine in mental disorders. N Engl J Med 1956; 254:515–519
15. Winkelman NW Jr: An appraisal of chlorpromazine: general principles for administration based on experience with 1,090 patients. Am J Psychiatry 1957; 113:961–971
16. Cole JO, Goldberg SC, Davis JM: Drugs in the treatment of psychosis: controlled studies, in Psychiatric Drugs. Edited by Solomon P. New York, Grune & Stratton, 1966
17. Ödegård Ö: Changes in the prognosis of functional psychoses since the days of Kraepelin. Br J Psychiatry 1957; 113:813–822
18. Wing JK: Five-year outcome in early schizophrenia. Proc R Soc Med (London) 1966; 59:17–18
19. Bleuler M: The Schizophrenic Disorders: Long-Term Patient and Family Studies. Translated by Clemens SM. New Haven, Conn, Yale University Press, 1978
20. Ciompi L, Mueller C: The Life Course of Aging Schizophrenics: A Long-Term Follow-Up Study Into Old Age. Berlin, Springer, 1978
21. Brown GW, Bone M, Dalison B, Wing JK: Schizophrenia and Social Care. New York, Oxford University Press, 1966
22. Strauss JS, Carpenter WT: The prediction of outcome in schizophrenia: characteristics of outcome. Arch Gen Psychiatry 1972; 27:739–746
23. Hawk AB, Carpenter WT, Strauss JS: Diagnostic criteria and five-year outcome in schizophrenia. Arch Gen Psychiatry 1975; 32:343–347
24. Harding CM, Brooks GW, Ashikagu T, Strauss JS, Breier A: The Vermont longitudinal study of persons with severe mental illness, II: long-term outcome of subjects who retrospectively met DSM-III criteria for schizophrenia. Am J Psychiatry 1987; 144: 727–735
25. Harding CM, Zubin J, Strauss JS: Chronicity in schizophrenia: revisited. Br J Psychiatry 1992; 161(suppl 18):27–37
26. Leff J, Kuipers L, Berkowitz R, Eberlein-Vries R, Sturgeon D: A controlled trial of social intervention in the families of schizophrenic patients. Br J Psychiatry 1982; 141:121–134
27. Zubin J, Sutton S, Salzinger K, Salzinger S, Burdock EI, Peretz D: A biometric approach to prognosis in schizophrenia, in Comparative Epidemiology of Mental Disorders. Edited by Hoch PH, Zubin J. New York, Grune & Stratton, 1961
28. McGlashan TH: A selective review of recent North American long-term follow-up studies of schizophrenia. Schizophr Bull 1988; 14:515–542
29. Stephens JH: Long-term prognosis and follow-up in schizophrenia. Schizophr Bull 1978; 4:25–47
30. Tsuang MT: Long-term outcome in schizophrenia. Trends Neurosci 1982; 5:203–207
31. Davis JM, Andriukaitis S: The natural course of schizophrenia and effective maintenance drug treatment. J Clin Psychopharmacol 1986; 6:2–10
32. Harding CM, Zubin J, Strauss JS: Chronicity in schizophrenia: fact, partial fact, or artifact? Hosp Community Psychiatry 1987; 38:477–486
33. Glass GV, McGaw B, Smith ML: Meta-Analysis in Social Research. Beverly Hills, Calif, Sage, 1981
34. American Medical Association: Subject Headings and Cross References: A Guide to the Classification of Medical Periodical Literature, 2d ed. Chicago, AMA, 1940
35. Armed Forces Medical Library: Subject Heading Authority List Used by the Current List Division, Armed Forces Medical Library. Washington, DC, AFML, 1954
36. National Library of Medicine: Medical Subject Headings, 1st ed. Washington, DC, NLM, 1960
37. National Library of Medicine: Medical Subject Headings, 2nd ed. Washington, DC, NLM, 1963
38. National Library of Medicine: Medical Subject Headings: Supplement to Index Medicus, vol 34, 1993. Bethesda, Md, NLM, 1993
39. Wing JK: Comments on the long-term outcome of schizophrenia. Schizophr Bull 1988; 14:669–673
40. Langfeldt G: The Prognosis in Schizophrenia and the Factors Influencing the Course of the Disease: A Katamnestic Study, Including Individual Re-Examinations in 1936. Acta Psychiatr Neurol Scand Suppl 1937; 13
41. Schneider K: Clinical Psychopathology, 3rd ed. Translated by Hamilton MW, Anderson EW. New York, Grune & Stratton, 1959
42. Leonhard K: Die defektschizophrenen Krankheitsbilder. Leipzig, Leipzig University, 1936
43. Mayer-Gross W: Clinical Psychiatry. Baltimore, Williams & Wilkins, 1956
44. Wing JK, Cooper JE, Sartorius N: The Measurement and Classification of Psychiatric Symptoms: An Instructional Manual for the PSE and CATEGO Programs. New York, Cambridge University Press, 1974
45. Feighner JP, Robins E, Guze SB, Woodruff RA Jr, Winokur G,

Munoz R: Diagnostic criteria for use in psychiatric research. Arch Gen Psychiatry 1972; 26:57–63
46. Spitzer RL, Endicott J, Robins E: Research Diagnostic Criteria (RDC) for a Selected Group of Functional Disorders, 2nd ed. New York, New York State Psychiatric Institute, Biometrics Research, 1975
47. Westermeyer JF, Harrow M: Prognosis and outcome using broad (DSM-II) and narrow (DSM-III) concepts of schizophrenia. Schizophr Bull 1984; 10:624–637
48. Duncan DB: t-Tests and intervals for comparisons suggested by the data. Biometrics 1975; 31:339–359
49. DerSimonian R, Laird N: Meta-analysis in clinical trials. Controlled Clin Trials 1986; 7:177–188
50. Cohen J, Cohen P: Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, 2nd ed. Hillsdale, NJ, Lawrence Erlbaum Associates, 1983, pp 266–267
51. Davis JM: Overview: maintenance therapy in psychiatry, I: schizophrenia. Am J Psychiatry 1975; 132:1237–1245
52. Wyatt RJ: Neuroleptics and the natural course of schizophrenia. Schizophr Bull 1991; 17:325–351
53. Carpenter WT, Strauss JS, Muleh S: Are there pathognomonic symptoms of schizophrenia? an empiric investigation of Schneider's first-rank symptoms. Arch Gen Psychiatry 1973; 28:847–852
54. Sullivan HS: The modified psychoanalytic treatment of schizophrenia. Am J Psychiatry 1931; 88:519–540
55. Lazell EW: The group treatment of dementia praecox. Psychoanal Rev 1921; 8:168–179
56. Hinsie LE: The psychoanalytic treatment of schizophrenia. Psychiatr Q 1927; 1:313–327
57. French TM, Kasanin J: Psychodynamic study of recovery of two schizophrenic cases. Psychoanal Q 1941; 10:1–22
58. Pope HG, Lipinski JF: Differential diagnosis of schizophrenia and manic-depressive illness: a reassessment of the specificity of "schizophrenic" symptoms in the light of current research. Arch Gen Psychiatry 1978; 35:811–828
59. Kendell RE, Cooper JE, Gourlay AJ, Copeland JRM, Sharpe L, Gurland BJ: The diagnostic criteria of American and British psychiatrists. Arch Gen Psychiatry 1971; 25:123–130
60. Baldessarini RJ: Frequency of diagnosis of schizophrenia versus affective disorders from 1944 to 1968. Am J Psychiatry 1970; 127:759–763
61. Stoll AL, Tohen M, Baldessarini RJ, Goodwin DC, Stein S, Katz S, Geenens D, Swinson RP, Goethe JW, McGlashan T: Shifts in the diagnostic frequencies of schizophrenia and major affective disorders from 1972 through 1991: an analysis of six North American psychiatric hospitals, 1972–1988. Am J Psychiatry 1993; 150:1668–1673
62. Hogarty GE, Anderson CM, Reiss DJ, Kornblith SJ, Greenwald DP, Ulrich RF, Carter M: Family psychoeducation, social skills training, and maintenance chemotherapy in the aftercare treatment of schizophrenia. Arch Gen Psychiatry 1991; 48: 340–347
63. Stein LI, Test MA: Training in community living: one-year evaluation, in A symposium: follow-up studies of community care. Edited by Greenblatt M, Budson RD. Am J Psychiatry 1976; 133: 917–918
64. Thompson SG, Pocock SJ: Can meta-analyses be trusted? Lancet 1991; 338:1127–1130
65. Dickersin K, Chan SS, Chalmers TC, Sacks HS, Smith H: Publication bias and clinical trials. Controlled Clin Trials 1987; 8: 343–353
66. Cohen P, Cohen J: The clinician's illusion. Arch Gen Psychiatry 1984; 41:1178–1182
67. Soni SD, Mallik A, Reed P, Gaskell K: Differences between chronic schizophrenic patients in the hospital and in the community. Hosp Community Psychiatry 1992; 43:1233–1238
68. Torrey EF: Economic barriers to widespread implementation of model programs for the seriously mentally ill. Hosp Community Psychiatry 1990; 41:526–531
69. Stephens JH, Astrup C, Carpenter WT, Shaffer JW, Goldberg J: A comparison of nine systems to diagnose schizophrenia. Psychiatry Res 1982; 6:127–143
70. Keith SJ, Matthews SM: The diagnosis of schizophrenia: a review of onset and duration issues. Schizophr Bull 1991; 17:51–67
71. Flaum M, Andreasen NC: Diagnostic criteria for schizophrenia and related disorders: options for DSM-IV. Schizophr Bull 1991; 17:133–156