

The Vermont Longitudinal Study of Persons With Severe Mental Illness, II: Long-Term Outcome of Subjects Who Retrospectively Met *DSM-III* Criteria for Schizophrenia

Courtenay M. Harding, Ph.D., George W. Brooks, M.D., Takamaru Ashikaga, Ph.D.,
John S. Strauss, M.D., and Alan Breier, M.D.

The authors present the findings from a long-term follow-up study of 118 patients from Vermont State Hospital who, when rediagnosed retrospectively, met DSM-III criteria for schizophrenia at their index hospitalization in the mid-1950s. The patients were studied with structured, reliable, multivariate instrument batteries by raters who were blind to information in their records. The rediagnostic process is described, and results of the follow-up are presented. Outcome varied widely, but one-half to two-thirds of the sample had achieved considerable improvement or recovered, in contrast to statements in DSM-III that predict a poor outcome for schizophrenic patients.

(Am J Psychiatry 1987; 144:727-735)

The third edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-III)* of the American Psychiatric Association both reflects and shapes current American thinking about the course and outcome of schizophrenia. Heavily based on the Feighner criteria (1) and the Research Diagnostic Criteria (2), *DSM-III* pictures the schizophrenic patient as a person with increasing residual impairment.

A complete return to premorbid functioning is unusual—so rare, in fact, that some clinicians would question the diagnosis. However, there is always the possibility of full remission or recovery, although its frequency is unknown. The most common course is one of acute exacerbations with increasing residual impairment between episodes. (*DSM-III*, p. 185)

Received Jan. 31, 1986; revised Oct. 7, 1986; accepted Dec. 16, 1986. From the Department of Psychiatry, Yale University School of Medicine, New Haven, Conn.; the College of Medicine and the College of Engineering and Mathematics, University of Vermont, Burlington; and the Clinical Neuroscience Branch, NIMH, Rockville, Md. Address reprint requests to Dr. Harding, 150 CMHC, 34 Park St., New Haven, CT 06519.

Supported by NIMH grants MH-29575, MH-40607, and MH-00340 and by Biomedical grant S0705429 from the College of Medicine, University of Vermont.

Copyright © 1987 American Psychiatric Association.

These impairments are said to include flattened affect, persisting delusions and hallucinations, and increasing inability to carry out everyday functions such as work, social relationships, or basic self-care. Such assumptions influence concepts of etiology (3) and course and outcome (4); in addition, they shape decisions about treatment (5), program implementation (6), economic planning (7), and social policy for mental health service delivery systems (8).

The advent of *DSM-III* has been seen by many clinicians and investigators as a major change in a field heretofore severely hampered in research and treatment relevant to schizophrenia by the lack of reliable definitions of diagnostic categories (9-11). With such a system in place (12), it is now possible to reaffirm or disconfirm the prevalent notions about the long-term course of schizophrenia.

This paper reports findings from the fifth very long-term follow-up study of schizophrenia conducted within the last decade (13-15) and the second such endeavor recently completed in the United States (16). It is the only study to date that has examined the long-term outcome of subjects rediagnosed as meeting the *DSM-III* criteria for schizophrenia.

The Vermont Longitudinal Research Project was a 32-year prospective follow-along study of a clinical research cohort (17-28). The prospectively gathered material has been combined with a systematic retrospective follow-back to document the lives of 97% (N=262) of the 269 original subjects.

In the mid-1950s, when they became subjects in the study, these patients were "middle-aged, poorly educated, lower-class individuals further impoverished by repeated and prolonged hospitalizations" (25, p. 29). Demographic, illness, and hospitalization characteristics of this cohort have been extensively described elsewhere (25-31).

The subjects were originally chosen for a rehabilitation program from the back wards of Vermont State Hospital because of their chronic disabilities and resistance to treatment. The chronicity criterion required subjects to have been disabled for 1 year before entry into the rehabilitation program. The term "disabled" was defined as inability to function in ordinary day-

to-day role capacities. Members of this cohort had been ill for an average of 16 years, totally disabled for an average of 10 years, and continuously hospitalized for 6 years. In addition, most patients had been given phenothiazines for 2½ years without enough improvement to warrant discharge. They were provided with a comprehensive rehabilitation program and released to the community during the mid-to-late 1950s in a planned deinstitutionalization effort (17-25, 31).

In the follow-up data collection period (1980-1982), 97% of the original cohort was extensively studied in a structured and reliable manner (30-31). The catamnestic period for these patients ranged from 22 to 62 years. More detailed descriptions of the methodology, the study sample, and the overall status of the cohort at follow-up may be found in our companion paper in this issue.

Initial results for these subjects, whose original diagnoses had been made according to *DSM-I* criteria, indicated that from one-half to two-thirds of the cohort had significantly improved or recovered (28, 30). These findings were at odds with the prevailing assumptions about the long-term course of schizophrenia. It was possible, however, that this discrepancy had been generated by the use of the loosely formulated *DSM-I* diagnostic guidelines. Therefore, with the publication of *DSM-III* while we were in the midst of our study, we undertook the task of giving a retrospective rediagnosis from case records for each of the 269 patients in order to determine what their *DSM-III* status would have been at the time they were selected for the study.

The present paper examines the process of rediagnosis and assesses the long-term outcome achieved by the group who met the *DSM-III* criteria for schizophrenia at selection. The two hypotheses involved in this aspect of the study were statements of common conceptions about schizophrenia: 1) Members of this cohort diagnosed as having met the *DSM-III* criteria for schizophrenia at index hospitalization would still have signs and symptoms of schizophrenia at follow-up. 2) Members of this cohort diagnosed as having met the *DSM-III* criteria for schizophrenia at index hospitalization would have uniformly poor outcomes in critical areas of functioning such as work, social relations, and self-care at follow-up. Confirmation of these hypotheses would lend support to the validity of the statements about the long-term course and outcome of schizophrenia that are made in *DSM-III*.

REDIAGNOSIS OF PATIENTS AT INDEX HOSPITALIZATION

Originally, 213, or 79%, of the 269 subjects had been given a diagnosis of schizophrenia according to *DSM-I* guidelines. Table 1 presents a breakdown by age, sex, and diagnosis of the entire cohort at entry into the study in the mid-1950s.

We instituted several methods to achieve the retro-

TABLE 1. *DSM-I* Diagnoses of 269 Chronic Psychiatric Patients at Entry Into the Vermont Study in the Mid-1950s

Diagnostic Category	Mean Age (years)	Subjects With Diagnosis (N=269)	
		N	%
Schizophrenia		213	79
Hebephrenic			
Men	36	13	
Women	41	9	
Catatonic			
Men	38	26	
Women	43	39	
Paranoid			
Men	43	48	
Women	45	59	
Undifferentiated			
Men	34	8	
Women	37	11	
Affective disorders		34	13
Men	39	16	
Women	38	18	
Organic disorders		22	8
Men	36	14	
Women	44	8	

spective rediagnosis. First, the two raters selected (J.S.S. and A.B.) were new to the project and blind to the outcome of each subject. The raters participated in two sets of interrater trials on 40 randomly selected cases (15% of the 269 subjects), which were independently assessed in a straight series without any discussion between raters. The case records and standardized record review abstracts from the time of the patient's entry into the study were stripped of all previous diagnostic assignments as well as any information about future episodes, hospitalizations, and other outcome information after index admission. (Index hospitalization was designated as the hospitalization during the 1950s during which transfer to the rehabilitation program occurred.) The *DSM-III* criteria were strictly applied.

The hospital records had been abstracted, as part of the overall goals of the larger project, in a structured and systematic manner by means of a battery of instruments known as the Hospital Record Review Form. This battery contained forms for extracting data about family and early life history, prodromal signs, and all hospital admissions. Interrater trials had revealed it to be a reliable instrument battery (31).

For a signs and symptoms checklist, we used Strauss's Case Record Rating Scale (32) and ratings from the World Health Organization's (WHO) Psychiatric and Personal History Schedule (33). This combination battery recorded behavioral descriptors and symptom dimensions noted by the clinician in recounting his or her impressions of the patient at the time of the original assessment. Case summaries and copies of the original chart information, such as admission and discharge summaries with ward notes but with all references to diagnosis deleted, were included in each diagnostic packet. Structured *DSM-III* diagnostic checklists from WHO and the Chestnut Lodge Fol-

TABLE 2. Follow-Up Status by *DSM-III* Category of 269 Chronic Psychiatric Patients in the Vermont Study Who Were Rediagnosed Retrospectively

Subjects' Follow-Up Status	Number of Subjects in Diagnostic Category						Total	
	Schizophrenia	Schizoaffective Disorder	Affective Disorders	Atypical Psychosis	Other	Organic Disorders	N	%
Alive and interviewed	82	25	29	13	19	10	178	66
Alive; refused participation	4	1	1	0	5	2	13	5
Could not be located	4	1	1	0	1	0	7	3
Deceased	28	4	16	5	7	10	70 ^a	26
Total								
N	118	31	47	18	32	22	268 ^a	
%	44	12	17	7	12	8		100

^aFor one patient there was not enough information to make adequate ratings.

low-up Study (34) were used by those making the rediagnoses to systematically summarize all the evidence for each diagnosis to be assigned.

Concerns about the quality of the records might be raised, because throughout the United States records from most state hospitals are considered to be poor. Vermont State Hospital's records, however, were remarkably complete. Since most of our subjects had also been the subjects of early phenothiazine drug trials before their entry into the rehabilitation program, the records tended to be of good research quality both before and during the institution of the federally funded rehabilitation program in 1957. The records described the evolution of symptoms by using statements from the patients themselves and gave examples of behaviors to illustrate the presence of hallucinations, delusions, catatonic waxy flexibility, and other symptoms. Such clinical notes were entered often by psychiatrists, residents, and other members of the treatment team. There were also mental status reports, past medical histories, results of current physical examinations, medication charts, treatment plans, progress notes, and admission and discharge summaries. In addition, social workers had collected systematic family and personal histories.

We conducted two sets of interrater trials. Complete agreement was achieved on 57% of the first 21 cases. In an analysis of the cases about which there was disagreement, it was found that 56% of the time, the second diagnosis proposed by one rater agreed with the first diagnosis selected by the other rater. Each rater agreed with the eventual consensual diagnosis 71% of the time overall and 75% of the time for schizophrenia. In assessing the level of interrater agreement, after collapsing the data into four diagnostic categories (schizophrenia, schizoaffective disorder, affective disorders, and "other" disorders), we generated an overall kappa coefficient (.35) for the first trial of .40 ($p=.001$) and a kappa of .40 ($p=.02$) for schizophrenia alone. In the second trial set of 19 cases, an overall kappa of .65 ($p<.0001$) was generated; the kappa for schizophrenia was .78 ($p<.0007$). Clearly, there was improvement in levels of agreement after the raters had further experience with the records and the diag-

observed statistic, we concluded that the kappa value fell within the range observed by Spitzer et al. (36).

After application of the *DSM-III* criteria to the entire set of cases, 118 subjects received a diagnosis of schizophrenia (see table 2).

Fifty-four percent (114 of 213) of those who were diagnosed as having schizophrenia according to the *DSM-I* guidelines retained the same diagnosis with the *DSM-III* criteria. (An additional four members of the *DSM-III* schizophrenia group were shifted from the *DSM-I* affective disorders category.) The primary shift from the *DSM-I* category of schizophrenia occurred to the *DSM-III* categories of schizoaffective disorder and atypical psychosis, not to the affective disorders category as expected from the experience of previous investigators.

The process of rediagnosis provided subtype categories for this subsample. The paranoid subtype predominated in both the *DSM-I* (50%, or 107 of 213) and *DSM-III* (61%, or 50 of 82) classification systems. The remaining subtypes included undifferentiated (17%, or 14 of 82), catatonic (13%, or 11 of 82), and disorganized (9%, or seven of 82).

METHOD

Of the 118 subjects who met the *DSM-III* criteria for schizophrenia, at follow-up 70% ($N=82$) were alive and were interviewed, 24% ($N=28$) were deceased, 3% ($N=4$) refused to participate, and 3% were lost to follow-up. It should be noted that these figures are nearly identical to those reported for the larger cohort in our companion paper (see table 1 in that paper).

The present paper focuses on the long-term outcome of the 82 subjects who were alive and were interviewed 20-25 years after their entry into the project, because their data were the most reliable. The catamnestic period for these subjects ranged from 22 to 59 years.

Forty-five percent of the 82 subjects who met the *DSM-III* criteria for schizophrenia at index hospitalization had been hospitalized for more than 6 years before being

hospital from 2 to 6 years, and 31% had been hospitalized less than 2 years.

Demographic analysis of these 82 subjects produced the following information. The group was split evenly between the sexes (41 men and 41 women). Their ages (as of July 1, 1981, which was the midpoint in the data collection period) ranged from 41 to 79 years. It should be noted that 91% (N=75) were above the age of 50; the average age for the group was 61 years. Fifty-five percent (N=45) of the subjects had not completed high school. Sixty-two percent (N=51) had never married, and only 10% (N=8) had remained married. Seventy-six, or 93%, were living in Vermont.

To carry out the follow-up study, our raters conducted two structured and reliable field interviews with each subject to ascertain current status and longitudinal patterns of community tenure. The raters were blind to previously recorded information about the subjects. Additional informants who knew each subject well were also interviewed, and ratings were verified. The six subjects who were not living in Vermont were interviewed with the same protocols. Another structured protocol (the Hospital Record Review Form, described at length elsewhere [31]) was used by a rater blind to all field information to abstract hospital and vocational rehabilitation records.

We used two structured interview batteries from the Vermont Community Questionnaire (30, 31), which included 15 standard scales and schedules, to assess the subjects' levels of functioning in a variety of areas at follow-up and to discern longitudinal shifts and patterns across the 20-25 years since the rehabilitation program began. All batteries were subjected to two sets of interrater trials 6 months apart and were found reliable (30, 31).

As part of the assessment, the two interviewers, who were new to the project and who had 5-8 years of clinical experience each, made ratings that provided a current clinical profile for each subject. The interviewers were blind to diagnostic record information when they made these symptom ratings, after the third hour of contact with each subject. The interviewers used the Research Diagnostic Criteria Screening Interview (36, 37), the Brief Psychiatric Rating Scale (38), and a reduced version of the Mini-Mental State examination (39) to make their assessments. This assessment package was designed to replace the Schedule for Affective Disorders and Schizophrenia (SADS) (40), proposed in our original design (27), because the extra time and costs required for the SADS interview were not funded.

In addition, the Global Assessment Scale (GAS) (41, 42) provided a single score (from 0 to 100) based on level of symptoms and social functioning. The scale's developers divided scores on the instrument into three categories (0-30=poor, 31-60=fair, 61-100=good functioning). The interrater trials generated a Pearson coefficient of .85 ($p < .0001$) for the first set (N=20) and .93 ($p < .0001$) for the second set (N=20) on this scale alone.

The Strauss-Carpenter Levels of Function Scale (43)

was used to identify some of the major components that constitute the overall level of functioning assessed by the GAS. Each of the nine items is scored from 0 (poorest) to 4 (best); they include hospitalizations, symptoms, amount and quality of friendships, amount and quality of work, ability to meet basic needs, fullness of life, and overall level of functioning. (We excluded quality of work because, unlike all the other assessments, it could not be cross-checked by separate informants. A visit to each subject's work site was not deemed to be in the best interests of our subjects, most of whose employers might have been unaware of their early history as state hospital patients.) The results of interrater trials on this instrument alone generated Pearson coefficients of .92 ($p < .0001$) on the first set (N=21) and .92 ($p < .0001$) on the second set (N=18).

RESULTS

For one-half to two-thirds of these subjects who retrospectively met the DSM-III criteria for schizophrenia, long-term outcome was neither downward nor marginal but an evolution into various degrees of productivity, social involvement, wellness, and competent functioning. The more stringent DSM-III diagnostic criteria for schizophrenia failed to produce the expected uniformly poor outcome.

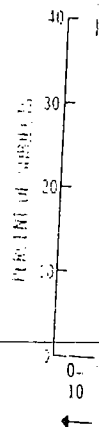
The combined data from the structured instrument battery described earlier, as well as all of the clinical observations obtained in the 3-hour interview sequence, indicated that 68% of the 82 subjects who met the DSM-III criteria for schizophrenia at index hospitalization did not display any further signs or symptoms (either positive or negative) of schizophrenia at follow-up. Forty-five percent of the sample displayed no psychiatric symptoms at all. For another 23%, symptoms had shifted to probable affective or organic disorders. One person was rated as a probable alcohol abuser (see table 3).

Eighty-four percent of the 82 subjects had had psychotropic medications prescribed for them; 75% of these in a low to medium dose range (in chlorpromazine equivalents). Seventy-five percent of the subjects stated they were complying with their regimes, but field interviewers were eventually told, after hours of interview time had elapsed, that the actual compliance pattern was closer to the following: about 25% of the subjects always took their medications, another 25% self-medicated when they had symptoms, and the remaining 34% used none of their medications. Adding the 34% who were noncompliers to the 16% who were currently not receiving any prescriptions for psychotropics means that 50% of the cohort was not using such medication.

A single score for psychological and social functioning was assigned each subject on the basis of the GAS. Figure 1 compares outcome scores of the subjects who met the DSM-III criteria for schizophrenia at index hospitalization with the scores of the subjects who met

TAI
Ven
Rec
—
Diag
No s
Schiz
Po
Ne
I
F
F
Affect
Def
Proi
Post
Organ
Defi
Prot
Poss
Alcohc
Defir
Prob
Possi
Not en

FIGURE
Vermont
Criteria



the DSA
hospitalizati
diagnose
scored 0
scale as
functioni
noted th
criteria fo
ects diag
schizophi

TABLE 3. Psychiatric Status at Follow-Up of 82 Patients in the Vermont Study Originally Diagnosed as Schizophrenic and Rediagnosed According to the RDC

Diagnostic Category	Subjects With Diagnosis (N=82)	
	N	%
No symptoms	37	45
Schizophrenia		
Positive symptoms		
Definite	1	1
Probable	7	9
Possible	0	0
Negative symptoms		
Definite	7	9
Probable	7	9
Possible	0	0
Affective disorders		
Definite	1	1
Probable	8	10
Possible	0	0
Organic disorders		
Definite	1	0
Probable	9	11
Possible	0	0
Alcoholism		
Definite	0	0
Probable	1	1
Possible	0	0
Not enough information to rate	4	4

FIGURE 2. Global Assessment Scale Scores of Subjects in the Vermont Study Who Met Both DSM-I and DSM-III Criteria for Schizophrenia and Subjects Diagnosed as Schizophrenic by DSM-I Who Had Other Diagnoses According to DSM-III

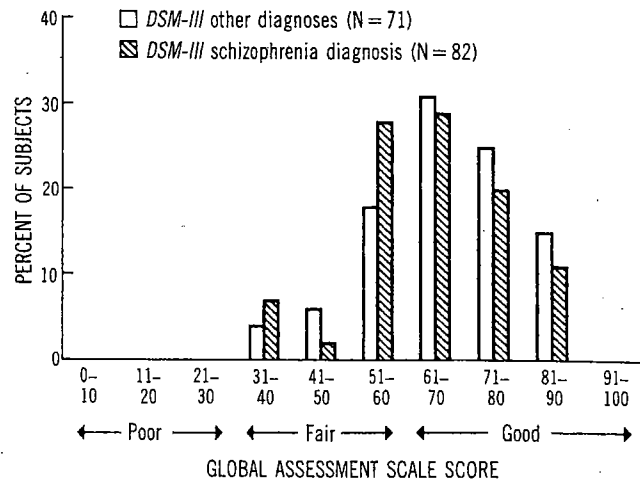


Figure 2 shows the GAS scores of the 82 subjects who met both DSM-I and DSM-III criteria for schizophrenia (including four subjects who were in other categories of DSM-I but who met DSM-III criteria for schizophrenia) and of the subjects who met DSM-I criteria but who were reclassified as fitting some category other than schizophrenia by DSM-III criteria (N=71). A t test for the means of the two groups revealed no significant differences between them ($t = -1.44$, $df = 149$, n.s.).

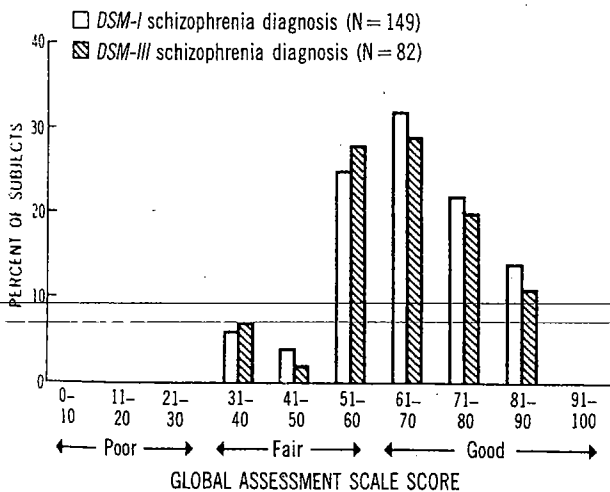
Table 4 shows the findings from the Levels of Function Scale for living subjects originally diagnosed as meeting the DSM-I guidelines for schizophrenia, those for subjects who met the DSM-III criteria for schizophrenia, and those for subjects who met the DSM-III criteria for other categories. For most outcome variables in either diagnostic system, for any of the three groups, two-thirds to four-fifths of the subjects were found to be significantly improved.

The only exception to the high levels of functioning across all diagnostic categories was the rating for employment, which was scored for one-half or fewer of the subjects. However, this rating did not take into account subjects who were retired or elderly.

The major difference between the subjects who met the DSM-III criteria for schizophrenia and those who met the DSM-III criteria for other diagnoses was fewer close friendships for the DSM-III schizophrenia subsample (68% versus 86%) ($\chi^2 = 4.89$, $df = 1$, $p = .03$).

We compared these two groups by using a 2x2 chi-square test with Yates' correction. A small number of cases with missing values were included in the analysis category that reflected the least positive outcome. No significant differences in results were observed when we used this approach and when we used the standard method of excluding cases with missing values.

FIGURE 1. Global Assessment Scale Scores of Subjects in the Vermont Study Who Met DSM-I Criteria and Those Who Met DSM-III Criteria for Schizophrenia at Index Hospitalization



the DSM-I guidelines for schizophrenia at index hospitalization. Sixty percent or more of the subjects diagnosed as schizophrenic by both diagnostic systems scored over 61, designated by the developers of the scale as good functioning. No one scored in the poor functioning category (score of 30 or less). It should be noted that all but four subjects who met the DSM-III criteria for schizophrenia came from the pool of subjects diagnosed as meeting the DSM-I criteria for schizophrenia.

