Depression as a Tracer Condition for the National Study of Medical Care Outcomes

Background Review

Kenneth B. Wells
The research described in this report was supported in part by a grant from the Robert Wood Johnson Foundation and in part by the Henry J. Kaiser Family Foundation under Grant No. 84-2820.

Library of Congress Cataloging in Publication Data

Wells, Kenneth B., 1948-
Depression as a tracer condition for the National Study of Medical Care Outcomes.

"July 1985."
"R-3293-RWJ/HJK."
Includes bibliographies.
8. Henry J. Kaiser Family Foundation. V. Title.
[DNLM: 1. Depressive Disorder. 2. Health Services Research—United States. 3. Outcome and Process Assessment (Health Care). WM 171 W454m]
RC537.W37 1985 362.2'0973 85-17044
ISBN 0-8330-0663-0

The Rand Publication Series: The Report is the principal publication documenting and transmitting Rand’s major research findings and final research results. The Rand Note reports other outputs of sponsored research for general distribution. Publications of The Rand Corporation do not necessarily reflect the opinions or policies of the sponsors of Rand research.

Published by The Rand Corporation
Depression as a Tracer Condition for the National Study of Medical Care Outcomes

Background Review

Kenneth B. Wells

July 1985

Prepared for the
Robert Wood Johnson Foundation
Henry J. Kaiser Family Foundation
FOREWORD

NATIONAL STUDY OF MEDICAL CARE OUTCOMES

The medical care system is being restructured with the goal of containing rising health care costs. The success of competing delivery systems is typically equated with how well they achieve this goal. Little attention has been given to what makes a difference to the patient.

The National Study of Medical Care Outcomes, known for brevity as the Medical Outcomes Study (MOS), has been designed to expand the debate regarding alternatives for the delivery and financing of health care services. In addition to examining how health care costs are contained in three systems of care, the MOS will measure the outcomes of care for adults with chronic disease. This population was chosen for study because they account for the great majority of health care expenditures and because they have the most to gain or lose from the restructuring of the health care system.

A major feature of the MOS is its emphasis on the patient's perspective. Thus, in addition to traditional measures of biomedical parameters, the health outcome assessment includes measures of physical, mental, and social/role functioning as well as general health perceptions. Patient satisfaction with care is also surveyed at regular intervals. The MOS will describe how cost-savings are achieved, whether they occur at the expense of health, and whether differences in outcomes can be explained in terms of patient, provider, practice, and delivery system characteristics.

The MOS will address six primary research questions:

1. **How do patient outcomes vary among different systems of care?**

   The study will evaluate how patient outcomes differ among three distinct ways of organizing and financing health care services:
   - Prepaid group practice form of HMO
   - Large, multispeciality, fee-for-service group practice
   - Small group and solo fee-for-service practice

2. **What is the relationship between system of care and utilization of health care resources?**

   The study will estimate the relationship between system of care and the use of health care resources, with the goal of
describing how any differences in utilization of these resources are achieved.

3. What is the relationship between style of practice and patient outcomes?
Both within and between systems of care, the study will attempt to identify features of practice style that affect patient outcomes. Style of care will be measured in terms of technical process, interpersonal manner, counseling for personal and emotional problems, division of responsibility and control between doctor and patient, information sharing, and continuity and coordination of care.

4. How do patient outcomes and utilization of health care resources vary when care is managed by providers trained in different specialties?
For patients in the fee-for-service system, the study will describe how patient outcomes and utilization of health care resources differ when principal care is provided by physicians and other providers trained in different specialties. Family physicians, general internists, subspecialty internists, psychiatrists, and clinical psychologists will be included.

5. Do the benefits of competing systems of care differ for key subgroups of the population?
The study will test whether the effects of different systems of care are the same for those who are economically disadvantaged or elderly, compared with those who are relatively well off or nonaged.

6. How does chronic disease affect patient functioning?
The study will describe differences in functional status and general health perceptions for patients with selected chronic diseases. The relationship between changes in disease severity and patient functioning over time will also be examined to identify factors that account for differences among patients in disease impact.

The MOS is an observational study of what naturally occurs between patients, physicians and other providers in the settings and under the circumstances they have chosen. About 700 physicians will be sampled from different systems of care in at least four geographic sites. At least one site will be included from each of the four
major census regions. All three systems of care and all provider specialty groups will be studied at each site.

Both cross-sectional and longitudinal studies will be fielded. The cross-sectional study will include all patients seen by participating providers during a 5-day screening period (about 50,000 patients). A subset of patients with specified chronic diseases (tracer conditions) will be enrolled in the longitudinal panel study (about 3,600 patients). The tracer conditions include advanced coronary artery disease, diabetes, high blood pressure, and depression. These conditions were selected because they have substantial functional impact, are relatively prevalent, can be measured in terms of disease severity, and can be affected by healthcare as currently practiced.

Data will be gathered from patients, their physicians, and other healthcare providers. All patient contacts with the healthcare system will be studied over a two-year period. The primary data-collection procedures include: mailed questionnaires that are self-administered at regular intervals, computer-assisted telephone interviews, physician- and patient-completed healthcare visit forms, and patient diaries. MOS instruments and data-collection procedures have been extensively tested prior to their adoption.

The tracer condition methodology will be used to standardize comparisons between systems of care and between provider specialty groups. To achieve this goal, the MOS carefully measures disease severity, comorbidity and other case-mix variables when patients are enrolled. These measures are used to control statistically for case-mix differences when comparing outcomes across systems of care and addressing other major research questions. This methodology will permit the strongest possible conclusions regarding factors most likely to influence patient functional outcomes.

After two years of planning and development, the MOS began a pilot study in Los Angeles in June 1985. This pilot, which simulates formal study conditions, will precede each step of the actual MOS by approximately six months. The first formal study site will enroll providers and patients in January 1986. Data collection will be completed in the summer of 1988. Analyses will begin early in 1986 and will be completed by the end of 1989.

The MOS is sponsored by grants from the Robert Wood Johnson Foundation and the Henry J. Kaiser Family Foundation. The study has been endorsed by the American Academy of Family Physicians, the American College of Physicians, and the American Psychiatric Association. For further information, please contact John E. Ware, Jr., Ph.D., Project Director, Medical Outcomes Study, The Rand Corporation, 1700 Main Street, P.O. Box 2138, Santa Monica, California, 90406-2138.

—John E. Ware, Jr.
PREFACE

The National Study of Medical Care Outcomes (MOS) is in a design development phase. As part of this development, the MOS investigators considered several medical conditions for inclusion as tracers: chronic obstructive pulmonary disease; hypertension; diabetes; arthritis; chronic ischemic heart disease; depression; asthma in children; recurrent otitis media in children; general examination in children; and outpatient medical care of the elderly. For each prospective tracer condition, a consultant prepared a brief report summarizing eleven features of the condition:

1. The definition of the condition
2. The prevalence of the condition in the general population and in the practices of physicians with different training
3. How to identify patients who are to be enrolled, from among patients visiting their physicians
4. How to assess the severity of the disease
5. The functional impact of the disease
6. What treatments are typically applied and what outcomes they are known to lead to in terms of:
   — Disease severity
   — Functional outcomes
7. Whether treatments or outcomes vary as a function of physician specialty or organization of practice
8. What outcomes are relevant for evaluating the care of this condition
9. Any other factors that influence treatments or outcomes that will need to be controlled for (e.g., case-mix control)
10. The course of the disease and treatment as they affect the timing of assessments of outcomes
11. Policy issues related to the condition and its treatment

On the basis of an earlier version of this report, depression was selected as one of four tracer conditions for study in the MOS. This report represents a revised version, in response to reviewer comments and further development of the MOS research questions and design.
EXECUTIVE SUMMARY

1. Definition of major depression: The term "depression" can be used to refer either to the general phenomena of depressed mood or to specific mental disorders. This report primarily concerns specific depressive disorders. Several competing classification schemes provide definitions of the depressive disorders. In the National Study of Medical Care Outcomes (MOS), we use a definition based on the third edition of the American Psychiatric Association's Diagnostic and Statistical Manual (DSM-III). We do so for three reasons: First, DSM-III is the most recently developed classification system and reflects the current state of the art in the understanding of depressive disorders. Second, the MOS should generate results that are useful for both researchers and clinicians, and U.S. clinicians are more likely to be familiar with DSM-III than with schemes developed exclusively for research purposes, such as the Research Diagnostic Criteria (RDC). Although the International Classification of Diseases (Ninth Edition) (ICD-9) is widely used clinically, its classification scheme for the depressive disorders diverges considerably from DSM-III. For example, ICD-9 includes "neurotic depression," but the validity of this category has been challenged by recent research. Third, DSM-III has been accepted by the Health Care Financing Administration (HCFA) as part of the officially adopted classification system for reimbursement by U.S. government programs.

According to DSM-III, major depressive disorder is characterized by one or more major depressive episodes and the absence of manic episodes (e.g., elated mood). A major depressive episode is characterized by persistent depressive mood or loss of interest or pleasure in nearly all usual activities for a period of at least two weeks, accompanied by symptoms in four of the following eight symptom groups: disturbances in appetite or weight; sleep disturbances; psychomotor agitation or retardation; decreased interest or pleasure; decreased energy; feelings of worthlessness or guilt; difficulty concentrating or thinking; and thoughts of death or suicide or suicidal attempts. The diagnosis does not apply when the syndrome is superimposed on certain mental disorders (e.g., schizophrenia, schizophreniform disorder, paranoid disorder), or when due to an organic mental disorder or uncomplicated bereavement.

Our definition of major depression differs from that of DSM-III in that we do not apply the full set of exclusion criteria. Rather than perform a complete diagnostic assessment for schizophrenia, paranoia, and
organic mental disorders, we use a brief screener for mood-incongruent delusions or hallucinations, alcoholism, and cognitive impairment.

The MOS depression tracer will also include some patients with dysthymic disorder, which is characterized by depressed mood or loss of interest in nearly all usual activities for two years or more. However, the depression is not of sufficient severity or persistence to meet the criteria for major depression. We include patients with dysthymic disorder because of its tendency to coexist with episodes of major depression and because of the policy relevance of the disorder (e.g., its associated limitations in functional status and use of health services), especially in general medical settings.

2. Prevalence of major depression: Major depression is one of the most common specific mental disorders (Klerman, 1980a). Epidemiologic studies indicate that 15 to 30 percent of adults experience a clinically significant depressive episode sometime in their lives. The six-month prevalence of major depression (DSM-III criteria) in three sites of the Epidemiologic Catchment Area (ECA) Program is from 2.2 to 3.5 percent (Myers et al., 1984). Major depression is more prevalent in women than in men; in men it increases with age (Boyd and Weissman, 1982).

The majority of persons in a general population with major depression do not receive specific treatment for the disorder (Myers and Weissman, 1980). Nevertheless, major depressive disorder is probably the most common mental disorder in general medical settings (Katon, 1982)—about as common as the most common physical disorders. For example, an anxiety-depression cluster of diagnoses was the fifth most frequent primary diagnosis group for all ambulatory encounters in the National Ambulatory Medical Care Survey in 1977 and 1978 (Rosenblatt et al., 1983).

Dysthymia appears to be about as prevalent as major depression. Its six-month prevalence across three ECA sites varies from 2.1 to 3.8 percent.

Depression is among the most common diagnoses assigned by psychiatrists in private practice. Langsley and Lebaron (1974) found that 14.0 percent of patients seen by practicing psychiatrists had a diagnosis of depressive neurosis. Marmor (1975) found that depressive neurosis was the most common diagnosis assigned by psychiatrists in private practice. (See the full report for a discussion of the relationship of neurotic depression to major depression and dysthymia).

3. Casefinding: For purposes of the MOS, patients with major depression and dysthymia will be identified with a two-stage procedure. At the first stage, we will use a self-assessment measure of depressive symptoms to identify patients with a high likelihood of major
depression. At the second stage, we will use a structured interview to identify patients with a current major depressive episode or dysthymia. The two-stage process is necessary because it would be prohibitively expensive to administer structured interviews to all study patients.

For first-stage screening, we will select a self-assessment measure that has a high sensitivity (≥0.90) for major depression. While a high specificity is also desirable, a more moderate level of specificity could be tolerated for the sake of high sensitivity. Two measures that appear to be appropriate for first-stage screening of major depression are the General Health Questionnaire (GHQ) and the Center for Epidemiologic Studies Depression Scale (CES-D). We are also studying the Mental Health Inventory (MHI) as a brief screener for depression.

The GHQ is a measure of general psychopathology designed specifically as a screening device for psychiatric disorders in primary care settings (Vieweg and Hedlund, 1983). The reliability of the GHQ is high (Cronbach's Alpha 0.80; test-retest at six months 0.90). Hoepner et al. (1979) found that the sensitivity of the GHQ for identifying cases of major depression (RDC) was 100 percent.

The CES-D detects depressive symptoms across diagnostic categories (Radloff, 1977). The language used in the CES-D is very clear, and it appears to be especially useful for patients in general medical settings. Its internal consistency is high (0.85), and its test-retest reliability is acceptable (0.51–0.67). Its sensitivity for identifying all depressive disorders (RDC) in outpatients is 91 percent, but Myers and Weissman (1980) reported a sensitivity of only 64 percent for major depression (RDC) in a community sample. The MHI, based on the General Well-Being Schedule, was specifically designed to measure mental health in the Rand Health Insurance Study (Ware et al., 1979). Its internal consistency exceeds 0.90 and it has excellent construct validity, (Veit and Ware, 1983). One advantage of the measure for use in general medical settings is that it was specifically designed to minimize false positives due to physical health problems. Another advantage for the purposes of the MOS is that a five-item version exists (the MHI-5).

The structured interview for second-stage screening will be the depression section of the Diagnostic Interview Schedule (DIS) of the NIMH (Robins et al., 1981). Using the DIS, diagnoses can be assigned by DSM-III, RDC, or Feighner Criteria. This flexibility of the DIS would allow MOS results to be directly compared with other studies that used one of these three criteria sets. While the DIS can be administered by a trained lay interviewer, alternative interview schedules require a clinician-interviewer. We plan to use a telephone-administered version of the depression section of the DIS. We are
testing the validity of the telephone-administered version against face-
to-face DIS interviews.

4. Severity of the disease: The MOS will include both self- and 
observer assessments of the severity of depression. The self-
assessments provide the patient's perspective. Observer assessments 
are desirable because self-assessments of severity may be invalid during 
the acute phase of illness (Prusoff, Klerman and Paykel, 1972). To be 
useful for the MOS, a measure of severity should include some or all of 
the clinically relevant components of depressive disorders and should 
correlate positively with clinical assessments of severity.

One recommended self-assessment measure of severity of depression 
is the Beck Depression Inventory (BDI) (Beck, 1967). The BDI was 
specifically designed as a measure of severity, and standard cut-off 
scores that reflect different levels of severity have been identified 
(Beck, 1967, Salkind, 1969). The internal consistency is about 0.90. 
The BDI has consistently been shown to have a high correlation with 
observer measures of severity. Other possible measures include the 
depression scales of the Hopkins Symptom Checklist (HSCL) (Parloff, 
Kelman, and Frank, 1954) or the GHQ. In addition, we are evaluating 
the Mental Health Inventory as a measure of the severity of depression 
by testing it against the Hamilton Rating Scales.

The observer assessment of the severity of depression for the MOS 
will be the Hamilton Depression Rating Scale (HRS), the most widely 
used observer scale for assessing the severity of depression (Hamilton, 
1967, 1982). It was originally intended for use by clinicians following a 
clinical interview. The inter-rater reliability of the HRS is 0.90, and 
its correlation with clinical global judgments is between 0.84 and 0.90.

For all tracer conditions, the MOS will include global assessments of 
the severity of disease by the treating physician. Assessments of severity 
of depression by general medical providers would be less valid than 
those of mental health specialists if, as some have suggested, general 
medical providers were less sensitive to depressive symptoms or disor-
ders. If global assessments by provider are included in the depression 
tracer, the dimensions to be rated and the response options should be 
clearly defined. The Global Assessment Scale may serve as a model for 
such an observer scale with clearly defined responses (Endicott et al., 
1976). However, this measure requires the rater to assess both social 
functioning and psychopathology in one scale.

5. Functional impact: All of the depressive disorders are associated 
with excess mortality. Almost all of this excess is due to suicide and 
accidents. Of persons with a unipolar depression (i.e., no history of 
manic episodes), 15 percent commit suicide (Coryell, Noyes, and 
Clancy, 1982). Suicide usually occurs in the first ten years after onset
of major depression. Chronic and intermittent depression may also contribute to the development of cardiovascular disease, cancer, and other physical diseases (Klerman, 1980a).

Depressive disorders are associated with considerable impairment in social functioning, including occupational disability. Severe major depressive episodes are debilitating, and may lead to weight loss, job disability, and severe financial problems. Because depressive episodes may last six months or more, especially if untreated, the strain on social support and personal resources can be considerable. Further, untreated or poorly treated depression may lead to dysthymic disorder (i.e., chronic, intermittent depression) and chronic impairment in social and occupational functioning. Even mildly depressed persons have a reduced ability to maintain intimate relationships and experience reduced sexual interest and performance (Paykel and Weissman, 1973; Klerman, 1980b; Blumenthal and Dielman, 1975). Their reduced social skills may alienate others, and thus cause a loss of social support that deepens the depression (Billings, Cronkite and Moos, 1983).

6. Treatments: The treatments that have been demonstrated to be superior to placebo in treating acute major depressive episodes include tricyclic antidepressants, monoamine oxidase inhibitors, antipsychotic medication, electroconvulsive therapy, and specific psychotherapies (Mindham, 1982; Nies and Robinson, 1982; Kiloh, 1982; Klerman and Schechter, 1982). The tricyclic antidepressants, such as imipramine and amitriptyline, appear to be especially useful for symptoms of melancholia, or so-called endogenous depressions. Monoamine oxidase inhibitors are most effective for patients with atypical depressive features (e.g., weight gain, depression worse in the evening) and mixed anxiety-depression syndromes. Phenothiazines are most useful for patients with psychotic, agitated depression, as an adjunct to antidepressants. Electroconvulsive therapy is the treatment of choice for patients with marked psychomotor retardation or agitation and for acutely suicidal or delusional patients. Lithium and tricyclic antidepressants are effective in preventing recurrent major depressive episodes.

Several specific forms of psychotherapy have been shown to be superior to placebo in the treatment of depression (Weissman, 1979). Additional data on the efficacy of specific psychotherapies, such as cognitive therapy, will soon be available from the Collaborative Studies on Depression sponsored by NIMH. Further, psychotherapy and antidepressants appear to have additive effects as well as differential effects on depressive symptoms. Antidepressant medication appears to have its greatest effect on the somatic symptoms of depression and
psychotherapy on depressed mood and social functioning (DiMascio et al., 1979a).

The effectiveness of these treatments can be measured by the percentage of patients who improve and the amount of improvement in standard measures of depression. In most studies, about 25 percent of depressed patients in placebo groups recovered from their depression during the study period. By comparison, about 60 to 70 percent of those on antidepressants and about 70 to 90 percent of those receiving electroconvulsive therapy recovered. While there are fewer studies of the effect of psychotherapy, those that exist showed that about the same percentage of patients improved on antidepressant medication or psychotherapy. For most of the treatments, the amount of improvement over pretreatment assessments of the severity of depression was more than 50 percent. A recent meta-analysis of studies of the efficacy of antidepressant medications or specific psychotherapies for persons with nonpsychotic unipolar depression found an average effect size of 1.25 standard deviations for psychotherapy and 0.79 of a standard deviation for drug therapy.

7. Variation in process or outcomes with provider type or treatment setting: For tracer conditions other than depression, the MOS patient sample will consist of outpatients treated by general medical providers in a variety of practice settings. Thus, the MOS is primarily a study of care in general medical settings. For a depression tracer, however, we would find it difficult to interpret information about the care of patients with depression in medical settings without data on the care of such patients by mental health specialists as a reference point. Further, some patients receive their care from both general medical physicians and mental health specialists. As a result, we will study patients with depressive disorders who receive their care from either type or both types of provider. Our measures of the process of care, therefore, must be appropriate for both types of provider. Studies of the process of care for mental disorders have focused almost exclusively on the care provided by mental health specialists. These studies have compared well-defined treatment modalities (e.g., cognitive therapy) or have described the process of psychotherapy in considerable detail (e.g., specific interpretations, nonverbal behavior). Such a conceptualization of process would not be feasible for the MOS, and may be unnecessarily detailed for a study that focuses on general medical providers.

As a result, we use a simple conceptual model of the process of care for depression that is based on a synthesis of the literature on the treatment for depression and a problem-oriented approach to the management of medical problems. In the model, the treatment of depression is conceptualized as a sequence of clinical decisions
involving the detection of a depressive disorder, differential diagnosis, identification of target symptoms, selection and monitoring of treatment, and evaluation of clinical outcomes and determination of the need for prophylactic medication. Existing studies suggest that mental health specialists and general medical providers differ at several steps in this sequence, especially in detection of depressive disorders and the selection of a treatment modality. General medical providers may detect depression in a small proportion of their patients who have a major depressive disorder (Nielsen and Williams, 1980). This conclusion has not been confirmed, however, in carefully designed studies of U.S. physicians in private practice. While no studies exist on the detection of major depression by psychiatrists or psychologists, at least one study suggests that psychiatrists are sensitive to their patients' depressive symptoms (Kass et al., 1980).

Concern has been expressed about the appropriateness of the prescribing habits of both psychiatrists and general medical providers. However, the available evidence suggests that psychiatrists are more likely than general medical providers to select an appropriate antidepressant and to prescribe therapeutic doses of antidepressants (Johnson, 1974). Further, while psychiatrists tend to administer individual psychotherapy to most of their patients, general medical providers provide their mental health care in the context of brief office visits that include physical health evaluations and procedures (Wells et al., 1982, Bureau of Health Professions, 1984). These findings strongly suggest that there are major differences in the mental health care delivered by the two provider groups.

The care delivered to depressed patients in HMOs may differ considerably from that in fee-for-service settings. The economic incentives provided in HMOs may affect the process of care in four ways (Luft, 1981). First, while fee-for-service providers gain financially from increased use of their services, HMO providers do not. Typically, their income does not directly depend on the amount of services they provide. Thus, HMO providers probably attempt more than fee-for-service physicians to control use by limiting the frequency and/or the intensity of care (Biltsch and Idzonek, 1978). Second, general medical providers in an HMO may have a greater incentive than fee-for-service physicians to refer to mental health specialists without fear of losing their patients. Third, HMO patients typically have less control over the specific type of treatment they receive. As a result of these factors, patients in an HMO may be more likely to receive time-limited (crisis oriented) or group therapy. Fourth, HMOs may rely heavily on non-physician personnel for provision of mental health services. The use of psychiatrists may be reserved for sicker patients or be limited to the
prescription of medications or to supervision of nonphysicians. As a result, there may be major differences in case-mix of patients of psychiatrists or the role of psychiatrists in fee-for-service and HMO settings.

8. Outcomes: In the MOS, the term “outcome” is used to refer to dimensions of health status and functioning that can be affected by the level or quality of medical care. “Outcomes” are measured at selected intervals during the study and at exit from the study. While these “outcomes” may partly reflect the adequacy of care, they may also reflect other characteristics of the patients (e.g., stage of disease, concurrent illnesses, age) and the environment (e.g., stressful events). The relevant “outcomes” for patients with depressive disorders are the exit diagnosis, severity and type of depressive symptoms, mortality, course of illness, social impairment, and satisfaction with care. Each type of outcome is discussed below.

Diagnostic reassessment at the end of the study is important for several reasons. First, a diagnostic assessment at exit can clarify whether a major depressive episode has resolved or whether the patient is experiencing a new depressive episode. Second, a diagnostic reassessment can document the development of new mental disorders. For example, dysthymic disorder may result from inadequately treated major depression. This is of particular importance because individuals with a residual chronic depression are prone to relapse and may require continued antidepressant medication (Keller et al., 1982b). Third, the reassessment may identify individuals who developed manic symptoms during the course of the study and thus should be classified as having a bipolar disorder. Such individuals may have to be excluded from the analysis because bipolar affective disorders respond differently to treatment than do unipolar depressive disorders and have a different course of illness (Coryell and Winokur, 1982).

Depressive symptoms are important outcomes for several reasons. First, they cause patients considerable distress. Second, different treatments may have differential effects on symptoms. For example, psychotherapy may have its strongest effects on mood and social adjustment, and pharmacotherapy may have a greater effect on associated somatic symptoms (e.g., weight loss, sleep disturbance) (DiMascio et al., 1979a). Third, some symptoms are believed to be especially disabling (e.g., psychotic symptoms) or of prognostic importance (e.g., suicidal ideation).

The duration of major depressive episodes and rate of recurrence are also meaningful treatment outcomes. For example, adequate treatment may shorten episodes, allow more symptom-free periods, or reduce the risk of recurrences. Because recurrences and longer episodes are likely
to result in more functional impairment, these episode characteristics have important consequences for the patient’s quality of life.

Social impairment is usually considered to be the most important nonfatal consequence of depressive disorders. Further, social impairment has been shown to improve with effective treatment of depression (DiMascio et al., 1979a). As a result, persistent impairment (at exit) may indicate inadequate treatment.

About 15 percent of patients with depressive disorders commit suicide. However, it is difficult to obtain reliable information on suicide for two reasons. First, physicians may be reluctant to record suicide as the cause of death on a death certificate. Second, some suicides may be masked as accidents (e.g., single-car accidents).

Enhancing patient satisfaction with care may be viewed as both a goal and outcome of treatment for depression. High patient satisfaction may enhance compliance with treatment, resulting in shorter episodes or fewer recurrences. Dissatisfied patients are more likely to interrupt their treatment by changing doctors (DiMatteo and Friedman, 1982). Satisfaction with psychiatric care has been previously included as a treatment outcome, and is modestly correlated with improvement in symptoms (Edwards et al., 1978; Willer and Miller, 1978). However, studies of the outpatient mental health care delivered by general medical providers suggest that satisfaction with care may be largely independent of the adequacy of the care received (Johnson, 1973, 1974).

Several self-assessment measures are recommended to measure the outcomes of depressive disorders. Techniques to diagnose the disorder and assess the severity of depression are discussed above. In addition, the MOS will include a measure of other symptoms of psychological distress, such as anxiety. The GHQ, HSCL, and MHI would all be useful for this purpose. The HSCL appears to have some advantage because it was specifically developed to monitor changes in symptoms in response to treatment, and considerable evidence exists on its validity for this purpose (Parloff, Kelman and Frank, 1954; Murphy, 1980). Measures of social functioning are being developed for all tracers in the MOS, based on several existing measures.

9. Other covariates: We will include several additional factors as covariates in the MOS analyses: certain characteristics of the major depressive illness, history of treatment, coexisting mental disorders and physical disorders, social supports, psychosocial stressors, attitudes toward mental disorders, demographic characteristics of patients, and health insurance coverage. Each covariate was selected because it may be predictive of the outcomes or because patients of different provider types (e.g., internists versus psychiatrists) differ in the characteristic.
The relevant characteristics of the major depressive illness include (1) the severity of the depression; (2) the presence of psychotic symptoms, melancholia, or symptoms of anxiety; (3) characteristics of the episode of depression; and (4) the level of pre-existing depression. The initial severity of depression may be the patient characteristic that is most predictive of outcome (Rounsaville, Weissman, and Prusoff, 1981). Patients of psychiatrists may have more severe depressions than patients of nonpsychiatrist physicians (Brown et al., 1979). Psychotic features and symptoms of anxiety suggest a poor response to tricyclic antidepressants; melancholia suggests a good response.

Two characteristics of the episode are of prognostic importance: the duration of the episode and the number of prior episodes. There is some evidence that the longer treatment is delayed, the poorer the response to treatment. Further, a history of multiple episodes, especially if there is persistent depressed mood between episodes, indicates a poor prognosis and the need for prophylactic antidepressants or lithium. Preexisting chronic depression suggests a poorer prognosis.

For four reasons, it is important to assess the history of treatment for the current episode (prior to enrollment in the MOS). First, patients who have received treatment may be more severely ill. Second, patients who have not yet received treatment may be more resistant to subsequent treatment because of the delay. Third, patients who have received treatment may have already improved, while patients who have not received treatment may have more room for improvement (i.e., in outcome measures). Fourth, patients of mental health specialists may be more likely to have received some treatment for their depression than are patients of nonpsychiatrist physicians. Without controlling for prior treatment, it may appear that those who have not received treatment (during the study period) improve more.

Coexisting mental and physical disorders may complicate the management of the depressive episode and the interpretation of health status at exit. For example, a major depression that is precipitated by a functional limitation due to chronic illness may fail to improve unless the chronic illness improves. Some major depressive episodes may be directly caused by physical illness. While patients with such disorders would not be categorized as having major depression according to strict DSM-III criteria, some patients with secondary depression might be included in our MOS "depression" sample, because we will not be applying the full set of exclusion criteria. The MOS investigators will include several instruments to assess the presence of chronic physical disease and the level of physical functioning and capacities. In addition, providers will be asked if depression (if present) is caused by underlying physical disease.
Lack of social supports and psychosocial stressors may play a role in the etiology of depression and may affect the course of illness. While social impairment is an outcome of treatment, the level of social support during treatment (i.e., contact with family, friends) may also be predictive of outcome.

Even if patients of mental health specialists and general medical physicians do not differ in case-mix, in terms of subtypes of depression, or severity, these groups may differ in their orientation toward mental health care. Patients of mental health specialists may be more familiar with psychological concepts, more receptive toward psychotherapy, and more likely to accept a diagnosis of a mental disorder and the treatment offered. If such differences occurred, then patients of mental health specialists could have better outcomes of treatment because of their more favorable attitudes, higher motivation, and compliance with treatment.

The patients of mental health specialists may differ on the average from those of general medical providers in age, sex, and educational status (Wells et al., 1982). Some of these demographic factors, such as educational status, may be predictive of outcome (Garfield, 1978). It may be especially important to control for study site in the MOS analysis, because providers in different geographic regions may vary considerably in the mental health care they deliver.

Health insurance coverage is an important determinant of the use of mental health services (Wells et al., 1982). Because of self-selection of coverage, patients of mental health specialists may have more generous coverage for mental health care than patients of general medical providers. However, it is not clear that differences in coverage per se would lead to different mental health “outcomes” at exit from the study.

10. The course of the disease: Major depression is an episodic and often recurrent disorder. While dysthymia is, by definition, a chronic disorder, patients with dysthymia are also at high risk for major depressive episodes. As a result, outcome measurement at an arbitrary time after entry into the study may not be a valid indication of the outcome of care for the index episode. That is, a patient may recover and then have another episode by the end of the MOS.

Three strategies can be used in the MOS to determine the course of major depressive episodes. First, patients can be selected who have only recently developed a depressive episode. Second, at several points in time, patients can provide retrospective accounts of the progression of their symptoms. Third, current depressive symptoms can be assessed at relatively frequent intervals to attempt to capture the fluc-
utuations of the episodes. At a minimum, outcomes should be reassessed at six-month intervals.

11. Policy issues: Depressive disorders are among the most prevalent specific mental disorders in the general population. Further, they are associated with significant mortality (e.g., 15 percent commit suicide) and morbidity (e.g., social impairment). Although effective treatments exist, a majority of patients do not seek specific treatment. A number of policy questions address this lack of treatment. For example, how do reimbursement strategies affect the prevalence of treatment of major depression in the general population? Does more generous reimbursement lead to improved social functioning and fewer disability days?

Many patients with depressive disorders present for treatment to general medical physicians. When they provide treatment, such physicians may be prone to prescribe subclinical doses of antidepressants. It is not known whether nonpsychiatrist physicians provide better care in some types of outpatient settings (e.g., prepaid group practice) than in others. Policy questions relevant to these concerns include: (1) What types of treatment settings are associated with a high quality of care for depressive disorders? (2) What is the level of care for depressive disorders obtained by general medical providers? (3) What subgroups of patients are treated most effectively by these providers? (4) How does the care delivered by general medical providers to patients with depressive disorders differ from that provided to comparable patients by mental health specialists? (5) How do the patients of these two provider groups differ, especially in the severity of their depressive symptoms and their level of functioning?

The costs of care for depressive disorders in the United States have not been estimated, but the cost of all mental health services in 1980 was $14 billion (McGuire, 1981). That figure does not include the costs to society of the social impairment associated with mental disorders.
ACKNOWLEDGMENTS

The author wishes to thank his colleagues John E. Ware, Jr., Anita Stewart, and Barbara Rose of Rand, and Marcia Daniels and Audrey Burnam of UCLA, for their assistance in obtaining relevant literature for this study. Detailed reviews of the manuscript were provided by R. H. Uhlenhuth of the University of Chicago and Carl Taube, Darryl Regier, Jack Burke, Jr., and Barbara Burns of NIMH. Additional consultations on the conceptualization of this report were obtained from Lee Robins, Herbert Pardes, Gerald Klerman, Myrna Weissman, Robert Hirschfeld, Robert Spitzer, Janet Williams, Martin Keller, and Harold Pincus.
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOREWORD</td>
<td>iii</td>
</tr>
<tr>
<td>PREFACE</td>
<td>vii</td>
</tr>
<tr>
<td>EXECUTIVE SUMMARY</td>
<td>ix</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>xxi</td>
</tr>
<tr>
<td>FIGURES AND TABLES</td>
<td>xxv</td>
</tr>
<tr>
<td><strong>Chapter</strong></td>
<td></td>
</tr>
<tr>
<td>1. DEFINITION OF DEPRESSION</td>
<td>1</td>
</tr>
<tr>
<td>Developments in the Conceptualization of Depression</td>
<td>2</td>
</tr>
<tr>
<td>Current Classification Schemes</td>
<td>4</td>
</tr>
<tr>
<td>Selection of Major Depression and Dysthymia as MOS</td>
<td>13</td>
</tr>
<tr>
<td>Tracer Conditions</td>
<td></td>
</tr>
<tr>
<td>2. PREVALENCE OF DEPRESSION</td>
<td>16</td>
</tr>
<tr>
<td>Prevalence in the General Population</td>
<td>16</td>
</tr>
<tr>
<td>Treated Prevalence</td>
<td>19</td>
</tr>
<tr>
<td>3. THE COURSE OF ILLNESS</td>
<td>22</td>
</tr>
<tr>
<td>Major Depression</td>
<td>22</td>
</tr>
<tr>
<td>Course of Dysthymic Disorder</td>
<td>24</td>
</tr>
<tr>
<td>4. TREATMENTS OF DEPRESSION</td>
<td>26</td>
</tr>
<tr>
<td>Pharmacotherapy</td>
<td>27</td>
</tr>
<tr>
<td>Electroconvulsive Therapy</td>
<td>30</td>
</tr>
<tr>
<td>Lithium</td>
<td>31</td>
</tr>
<tr>
<td>Psychosurgery</td>
<td>32</td>
</tr>
<tr>
<td>Psychotherapies</td>
<td>32</td>
</tr>
<tr>
<td>The Size of the Effect of Specific Treatments for Depression</td>
<td>35</td>
</tr>
<tr>
<td>Clinical Predictors of Response to Treatment</td>
<td>38</td>
</tr>
<tr>
<td>5. THE TREATMENTS DELIVERED TO DEPRESSED PATIENTS</td>
<td>41</td>
</tr>
<tr>
<td>Treatment Received for Depressive Disorders, All Providers</td>
<td>43</td>
</tr>
<tr>
<td>Treatment of Depression, General Medical Providers</td>
<td>44</td>
</tr>
<tr>
<td>The Treatment of All Mental Disorders, General Medical Providers</td>
<td>47</td>
</tr>
<tr>
<td>Summary and Implications for the MOS</td>
<td>55</td>
</tr>
</tbody>
</table>
6. A MODEL OF OUTCOMES OF DEPRESSION ........ 57
7. THE PROCESS OF CARE FOR DEPRESSION ...... 60
8. CASEFINDING TECHNIQUES .................... 66
   General Strategy .......................... 66
   Schedule of Affective Disorders and Schizophrenia
   (SADS) ................................... 67
   The Diagnostic Interview Schedule (DIS) ........ 69
9. OTHER MEASURES OF DEPRESSION ............. 73
   Self-Assessment Measures of Depression .......... 74
   Observer Assessments ........................ 82
   Impairment in Social Functioning ............... 89
   Summary and Implications for MOS ............. 90
10. SELECTION OF COVARIATES .................. 91
    Characteristics of the Depressive Disorder ... 91
    Treatment History .......................... 93
    Coexisting Mental Disorders ................ 94
    Coexisting Physical Disease ................ 94
    Social Supports ............................ 95
    Psychosocial Stressors ...................... 95
    Attitudes Toward Mental Health Care and Providers 96
    Demographic Variables ...................... 96
    Insurance Coverage ........................ 97
11. ETHICAL ISSUES FOR THE DEPRESSION ....... 99
    TRACER .................................. 99
REFERENCES .................................. 103
FIGURES

1.1. DSM-III Classification of Affective Disorders .................. 7
5.1. Treatment Plan for Depression ......................... 42

TABLES

1.1. Multiaxial Classification in DSM-III ....................... 6
1.2. Criteria Differences for Major Depressive Episode ........... 11
2.1. Prevalence of Current RDC Diagnoses .................... 17
2.2. Six-Month Prevalence of DIS/DSM-III Major Depressive
     Episode Without Bereavement, by Sex, Age, and Site .... 18
4.1. Definitions of Five Psychotherapies ....................... 33
4.2. Mean Scores for Four Treatment Groups, at Enrollment
     and at 16 Weeks .................................. 37
4.3. Mean Scores for Three Treatment Groups by Outcome
     Measure and Practice Setting ........................ 37
5.1. Services Given and Dispositions of Visits for Patients
     with a Primary Mental Health Diagnosis .................. 53
6.1. Outcome Dimensions for Depressive Disorders ............. 58
7.1. Dimensions of The Process of Care for Depression .......... 61
8.1. Dimensions of Psychopathology: Summary Scales from
     the SADS Interview ................................ 68
8.2. DSM-III Diagnoses Scored by DIS Computer Programs ..... 70
8.3. Comparison of SADS and DIS ............................. 72
9.1. Psychological Distress and Well-Being: Item Content
     of Mental Health Inventory, Adult Version ............... 83
9.2. Item Content of the Hamilton Rating Scale for
     Depression ...................................... 85
     Content and Organization ........................... 88
Chapter 1

DEFINITION OF DEPRESSION

The term "depression" is commonly used to refer either to a specific mental disorder or to a normal mood or emotion. As a normal emotion, depression refers to feelings of hopelessness, discouragement, and frustration that accompany experiences of life. In this context, depression may serve an adaptive function. It is a distress signal, a call for help, and a retreat from danger. As a clinical disorder, or set of disorders, depression refers to a more severe and persistent disturbance of mood accompanied by other, nonmood-related symptoms marked by changes in thinking, behavior, and physiology. Clinical depressions are associated with significant functional limitations, human suffering, and mortality. Many of the problems in defining clinical depression arise from the difficulties in finding a clear distinction between normal depressive mood and clinically meaningful depression.

Several authors, notably Boyd and Weissman (1982), emphasize the important distinction between depressive symptoms and depressive disorders. Depressive symptoms may be viewed as a final common pathway for the expression of distress—whether due to a disturbance in environment, metabolism, genetics, etc. From this conceptual framework, clinically significant depression is defined primarily by the number or intensity of depressive symptoms. Most general population surveys of "depression" are based on this framework; that is, the number, frequency, and severity of depressive symptoms are simply tabulated to indicate the "amount" of depression. As Boyd and Weissman (1982) note, such a conceptualization was prominent prior to 1970, partly because it avoided confusion about the meaning of specific diagnostic terms and because reliable measures of specific depressive disorders were not available.

By contrast, depressive disorders are distinct clinical syndromes, defined by a unique pattern of depressive symptoms and distinguished in part by family history, etiology, and/or treatment response. These disorders have been the focus of considerable clinical and research attention in the past three decades (Klerman, 1980a).

The relation of specific depressive disorders to the more general phenomenon of depressive symptoms is uncertain. While virtually all persons with depressive disorders have depressive symptoms, some
persons with depressive symptoms do not meet the criteria for a specific depressive disorder. Both groups are of some policy significance. For example, specific depressive disorders are associated with significant morbidity and mortality. Depressive symptoms, even in the absence of a disorder, may lead to use of medical services (Klerman, 1980b). This report concerns the diagnosis and management of major depression, a specific depressive syndrome, rather than the more general phenomenon of depressive symptoms.

This chapter reviews the definition of depressive disorders. First, it describes important historical changes in the conceptualization and classification of depression. Second, it summarizes the definition of the unipolar depressive disorders according to the Diagnostic and Statistical Manual, Third Edition (DSM-III) of the American Psychiatric Association (1980). Third, it compares the definition of major depression in that Manual with the earlier definition in DSM-II (1968) and the Research Diagnostic Criteria (RDC) (Spitzer, Endicott, and Robins, 1978).

DEVELOPMENTS IN THE CONCEPTUALIZATION OF DEPRESSION

The conceptualization of depressive disorders has changed considerably over the last 30 years. While diagnostic classification schemes developed prior to 1970 emphasized the distinctions between psychotic and neurotic depressions and between endogenous and reactive depressions, current diagnostic schemes emphasize the distinction between unipolar and bipolar depressions. For example, the category of "neurotic depression" does not appear in DSM-III. As a historical context for understanding the definition of depressive disorders, this chapter briefly reviews this shift in the conceptualization of depression.

Since the influential work of Bleuler (1951), most mental health clinicians have used the term "affective disorders" to refer to a continuum of serious disturbances of mood with pure depression at one end and pure mania at the other. Diagnostic schemes developed in the 1960s, such as DSM-II, incorporated this formulation. Among the depressive disorders, the principal distinction made by clinicians was that between psychotic and neurotic depression. Psychotic depression generally referred to (1) any significant disturbance in higher-level cognitive functions (i.e., memory, language, orientation, perception, and thinking) associated with depression; (2) any impairment in reality-testing with depression; or (3) severe limitations in social or functional status due to depression (Klerman, 1980a). Neurotic depression, by
contrast, referred to a milder depression with less impairment in social functioning.

The concept of neurotic depression as a distinct disorder has recently fallen out of favor, largely because the term has been imprecisely used. For example, Klerman et al. (1979) identified seven uses of the term “neurotic depression”: (1) all nonpsychotic depressions; (2) depression with little or no impairment in social functioning; (3) depression without vegetative or “endogenous” signs of depression (e.g., early morning awakening, weight loss, psychomotor retardation, guilt); (4) the presence of a specific constellation of symptoms, including self-pity, irritability, reactivity to the environment, and fluctuating symptoms; (5) a depressive response to a situational stress; (6) worsened symptoms in the context of a longstanding depressive personality; and (7) depressive response to unconscious psychological conflicts (e.g., over loss or dependency). To avoid confusion about the meaning of “neurotic,” several recent classification schemes avoid the term.

Another historically important distinction among types of depressive disorders is that between endogenous and reactive depressions. In the past, endogenous depressions were defined as more severe depressions that were associated with a specific symptom cluster (e.g., early morning awakening, weight loss, guilt, psychomotor retardation) and were not precipitated by psychosocial stress. Reactive depressions were viewed as milder depressions that were precipitated by psychosocial stress. Endogenous depressions were thought to be a phenomenon of mid- and late life, and reactive depressions were thought to be more prevalent in early adulthood. Further, persons with endogenous depressions were thought to have more stable premorbid personalities than did those with reactive depressions. Psychotic depressions were thought to be endogenous, and neurotic depressions, reactive.

Recently, the endogenous-reactive distinction has been seriously questioned. In particular, there is only a slight correlation between the severity of depression and the presence of precipitating factors. Further, the “endogenous” symptom cluster can occur with or without psychotic features (Klerman, 1980a).

Thus the conceptual model of subtypes of depression, as reflected by the psychotic-endogenous versus neurotic-reactive dichotomy, is no longer widely used, especially in the diagnostic schemes developed in the United States. As will be seen below, some of the descriptive aspects of this model (such as the “endogenous” symptom cluster) remain in current diagnostic schemes.

One of the most important among newer concepts of depression is the distinction between unipolar and bipolar depressive disorders. Leonhard, Korf, and Schulz (1962) proposed the separation of patients
with only depression episodes (unipolar) from those who also have manic episodes (bipolar). Subsequent research findings supported this distinction on the basis of family history (Gershon et al., 1976; Winokur, 1978), biological studies (Goodwin et al., 1978; Carroll, Curtis, and Mendels, 1976), treatment response (Goodwin et al., 1972; Fieve, Kumbaraci, and Dunner, 1976; Mendels, 1976; Quitkin et al., 1976), and clinical course and prognosis (Angst et al., 1973; Brodie and Leff, 1971; Winokur, 1978). As a whole, these studies suggest that bipolar depression probably consists of a few discrete disorders, while unipolar depression probably consists of many heterogeneous disorders that share a common clinical presentation (Boyd and Weissman, 1982). Both DSM III and the RDC recognize the distinction between bipolar and unipolar disorders, although these systems do not use this exact terminology.

Another important distinction among depressive disorders in some current diagnostic schemes is that between primary and secondary depression (Robins and Guze, 1972). Primary affective disorders are those in which the only previous mental disorders are mania or depression. Secondary depressive disorders are preceded by other psychiatric or physical illness. While the RDC specifically acknowledges this distinction, there is no direct counterpart in DSM-III.

CURRENT CLASSIFICATION SCHEMES

Prior to the 1970s, the classification schemes for mental disorders, such as the International Classification of Diseases and DSM-II, provided descriptions of clinical disorders but did not include specific criteria for assigning or excluding a diagnosis. Since the 1970s, clinicians and researchers in the United States and Great Britain have developed classification systems that include such explicit criteria. In the United States, the most commonly used systems are the Feighner Criteria (Feighner et al., 1972), the RDC, and DSM-III. These three systems represent a continuous development, with each subsequent system incorporating concepts and research findings not available during the development of the previous system (Boyd and Weissman, 1982). The Feighner Criteria and RDC were developed primarily for research purposes. For example, the RDC was developed as part of the Collaborative Project on the Psychobiology of Depression of the National Institute of Mental Health (NIMH). By contrast, DSM-III was developed by the Committee on Nomenclature and Statistics of the American Psychiatric Association to provide a standard nomenclature for clinicians.
Concurrent with the development of these systems in the United States, researchers in Great Britain developed the Present State Examination (PSE), which identifies clinical syndromes and disorders that are roughly equivalent to the mental disorders section of the International Classification of Diseases, Ninth Edition (ICD-9) (Wing, Cooper, and Sartorius, 1974). The ICD-9 is widely used throughout the world, and it continues to use the distinction between psychotic and neurotic depression. Wing has recently cautioned investigators against using the PSE as a definitive diagnostic instrument (Wing, 1983).

For purposes of the National Study of Medical Care Outcomes (MOS), the author recommends the DSM-III conceptualization of the depressive disorders for four reasons. First, the DSM-III is the most recently developed version and reflects the current state of the art in understanding depressive disorders. Second, the MOS should generate results that are useful for both researchers and clinicians. Clinicians in the United States are much more likely to become familiar with the DSM-III than with the RDC. The DSM-III has been accepted by the Health Care Financing Administration (HCFA) as a subset of the classification system for reimbursement of health services by U.S. government programs. Third, the DSM-III model of depressive disorders is somewhat simpler than that of the RDC. Fourth, the reliability of both the RDC and the DSM-III has been tested and found to be acceptable (Spitzer, Endicott, and Robins, 1978; American Psychiatric Association, 1980; Hyler, Williams, and Spitzer, 1982). As a result, the researcher is free to choose the scheme most suited to particular research needs.

**DSM-III and Unipolar Depression**

The DSM-III utilizes a multiaxial classification system. In effect, it is a comprehensive health evaluation from a mental health perspective (see Table 1.1). The complete spectrum of mental disorders is represented on Axes I and II. Axis I concerns specific clinical syndromes and reasons for consultation. Axis II concerns personality disorders and specific developmental disorders. The physical disorders are represented on Axis III. On Axis IV, the severity of psychosocial disorders is coded, and Axis V reflects the highest level of adaptive functioning in the past year.

According to DSM-III (p. 205), the essential feature of the affective disorders is:
Table 1.1
MULTIAXIAL CLASSIFICATION IN DSM-III

<table>
<thead>
<tr>
<th>Axis</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Clinical syndromes (e.g., all affective disorders)</td>
</tr>
<tr>
<td></td>
<td>Conditions not attributable to a mental disorder that are a focus of attention or treatment</td>
</tr>
<tr>
<td>II</td>
<td>Personality disorders</td>
</tr>
<tr>
<td></td>
<td>Specific developmental disorders</td>
</tr>
<tr>
<td>III</td>
<td>Physical disorders and conditions</td>
</tr>
<tr>
<td>IV</td>
<td>Severity of psychosocial stressor (as perceived by clinician or observer)</td>
</tr>
<tr>
<td>V</td>
<td>Highest level of adaptive functioning in the past year</td>
</tr>
</tbody>
</table>

a disturbance of mood, accompanied by a full or partial manic or depressive syndrome that is not due to any other physical or mental disorders. The manic and depressive syndromes each consist of characteristic symptoms that tend to occur together.

DSM-III divides the affective disorders into major affective disorders, other specific affective disorders, and atypical affective disorders (see Fig. 1.1). Each category is subdivided into bipolar disorders and unipolar (depressive) disorders. The three unipolar depressive disorders are major depressive disorder, dysthymic disorder, and atypical depressive disorder. An additional relevant diagnosis is adjustment disorder with depressive features. Below, we summarize the characteristics of the unipolar depressive disorders and adjustment disorder (based on DSM-III).

Major Depressive Disorder. Major depressive disorder is characterized by one or more major depressive episodes and the absence of manic episodes (e.g., elated mood). A major depressive episode is characterized by persistent depressive mood or loss of interest or pleasure in nearly all usual activities, accompanied by other depressive symptoms. The other symptoms include disturbances in appetite, weight, and sleep; psychomotor agitation or retardation; decreased energy; feelings of worthlessness or guilt; difficulty concentrating or thinking; and thoughts of death or suicide or suicidal attempts. DSM-III specifies that four of eight specific depressive symptoms must be present nearly every day for a period of at least two weeks to make a diagnosis of
Fig. 1.1—DSM-III classification of affective disorders
major depressive disorder in adults. The diagnosis does not apply when the syndrome is superimposed on certain psychiatric conditions (i.e., schizophrenia, schizophreniform disorder, paranoid disorder) or when caused by an organic mental disorder or uncomplicated bereavement.

In the DSM-III, a psychological reaction to a functional impairment associated with a physical illness may result in a major depressive episode. The major depression is coded on Axis I, the physical disorder on Axis 3, and the severity of the stressor on Axis 4. However, when the depression is directly caused by a physical illness (e.g., a metastatic carcinoma, vascular disease involving the brain) the diagnosis of major depression is excluded.

In the DSM-III, the presence of psychotic symptoms denotes a specific subtype of major depressive disorders. Psychotic features of a major depression are defined as gross impairment in reality-testing as manifested by delusions, hallucinations, or depressive stupor (i.e., the individual is mute and unresponsive). Psychotic features may be mood congruent or incongruent. Mood-congruent features include depressive stupor and delusions or hallucinations consistent with the depressive content. For example, delusions that one's body is rotting, or auditory hallucinations telling the depressed person to commit suicide, are mood-congruent psychotic features. Mood-incongruent psychotic features are delusions or hallucinations that do not involve personal inadequacy, guilt, disease, death, nihilism, or deserved punishment. For example, thought-broadcasting and delusions of being controlled by others are mood-incongruent psychotic features.

Major depressive disorders may occur with or without the presence of “melancholia,” a symptom complex that indicates a particularly severe type of major depression likely to respond to drug therapy. Melancholia is characterized by anhedonia (lack of pleasure in nearly all activities) and three or more of the following: (1) a distinct quality of depressed mood clearly different from normal “depression”; (2) depression worse in the morning; (3) early morning awakening; (4) marked psychomotor retardation or agitation; (5) significant anorexia or weight loss; (6) excessive or inappropriate guilt. Melancholia is similar to the old concept of endogenous depression, except that melancholia may be present whether or not the major depression is precipitated by life circumstances.

Dysthymic Disorder. As in major depressive disorder, a dysthymic disorder is characterized by depressed mood or loss of interest in nearly all usual activities. However, the depression is not of sufficient severity or persistence to meet the criteria for major depression. For adults, the disorder should be present at least two years, but there may be periods of normal mood lasting a few days to weeks, but not exceeding
a few months, at a time. The diagnosis requires three associated symptoms of depression during the depressive periods.

Dysthymic disorder and major depression may occur in the same individual. A person with dysthymic disorder may eventually develop a major depressive episode. Alternatively, a major depression may not completely resolve, leaving a chronic intermittent mild depression (i.e., dysthymic disorder).

**Adjustment Disorder with Depression.** Adjustment disorders are reactions to specific psychosocial stressors that occur within three months of the onset of the stressor. The reaction resolves when the stress subsides or when a new level of adaptation is achieved. If the principal symptoms of the disorder are those of depression, then a diagnosis of adjustment disorder with depressed mood is indicated. However, a full syndrome of major depression in response to a psychosocial stressor is categorized as major depression, and a pattern of dysthymia following a psychosocial stressor is categorized as a dysthymic disorder.

**Atypical Depressive Disorder.** This disorder is characterized by clinically significant symptoms of depression that do not meet the criteria of major depression or dysthymic or adjustment disorder with depression. In effect, it is a residual category. Nevertheless, this category may be of some policy significance. For example, Klerman (1980b) has suggested that “masked” depression is a type of atypical depression. Patients with masked depression usually do not complain of depressed mood; rather they present to general medical providers with somatic complaints associated with depression. Such patients may be especially high users of medical care.

**Comparison of DSM-III and DSM-II for Unipolar Depression**

This section summarizes some important differences in the conceptualization of the affective disorders between DSM-II and DSM-III. For a fuller discussion, see DSM-III, pp. 375–377.

1. In DSM-II, the category of major affective disorders refers specifically to affective psychoses and excludes disorders precipitated by life circumstances (i.e., the category refers to endogenous affective psychoses). By contrast, the affective disorders are grouped together in DSM-III, regardless of the presence or absence of psychotic features or precipitating life experiences.

2. DSM-II includes the category “involutional melancholia” to refer to psychotic depressions in mid- or late life. DSM-III
does not have a separate category for this because of the lack of evidence that such depressions differ from depressions that occur at other life stages.

3. DSM-II groups psychotic depressions and mania into one category (manic-depressive illness) with subtypes for manic, depressive, circular, or mixed. DSM-III acknowledges bipolar and unipolar (depressive) affective disorders as separate subcategories. In DSM-III, the diagnosis of a bipolar disorder is made whenever a manic episode is present, even in the absence of a history of depression. That is, there is no pure manic subtype, as is found in the DSM-II. Research findings show that almost all individuals with mania eventually develop depressive episodes.

4. DSM-II includes a “depressive neurosis” category. In DSM-III, any of three categories may be equivalent to neurotic depression: major depression without melancholia or psychotic features, dysthmic disorder, or adjustment disorder with depressed mood.

Comparison of the DSM-III and RDC Definitions of Depression

The RDC was the predecessor of the DSM-III. Numerous studies of depression, notably the NIMH Collaborative Studies on Depression, have used the RDC to identify depressive disorders. Thus it is important to understand the relationship between the DSM-III and the RDC definitions. Williams and Spitzer (1982) and Stoltzman et al. (1981) have provided detailed comparisons of the two systems. Table 1.2, derived from Stoltzman et al., summarizes the differences for the diagnosis of major depression:

1. The RDC has two sets of criteria: one for a “definite” diagnosis and one for a “probable” diagnosis. The DSM-III criteria are similar to the criteria for a “probable” diagnosis in the RDC.

2. The DSM-III requires that associated symptoms of depression be present for two weeks. The RDC requires depressive mood for a duration of one week (probable diagnosis) or two weeks (definite diagnosis). However, there is no duration requirement for associated symptoms. As a result, the RDC may generate more false negatives for major depression than does the DSM-III.
Table 1.2
CRITERIA DIFFERENCES FOR MAJOR DEPRESSIVE EPISODE

<table>
<thead>
<tr>
<th>Criteria</th>
<th>RDC</th>
<th>DSM-III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss of interest or pleasure treated</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>as equivalent to low mood</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of appetite</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>includes increased appetite</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of weight</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>includes weight gain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of symptom groups required</td>
<td>4 for a current (probable)</td>
<td>4 diagnosis</td>
</tr>
<tr>
<td>Duration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symptoms</td>
<td>—</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Episode</td>
<td>1 week</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Depressed mood</td>
<td>1 week</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Preemptions</td>
<td>Schizophrenia;</td>
<td>Bereavement, organic</td>
</tr>
<tr>
<td></td>
<td>Schizophrenia-like symp-</td>
<td>mental disorder; prior</td>
</tr>
<tr>
<td></td>
<td>toms; other nonaffective</td>
<td>onset of schizophrenia,</td>
</tr>
<tr>
<td></td>
<td>psychotic features for 1</td>
<td>schizophreniaiform or</td>
</tr>
<tr>
<td></td>
<td>month without low mood</td>
<td>paranoid disorder,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>psychotic between affec-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tive episodes</td>
</tr>
<tr>
<td>Distinction between definite and prob-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>able diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weaker criteria for past episodes</td>
<td>+</td>
<td>—</td>
</tr>
</tbody>
</table>


3. In the RDC, a history of schizophrenia, symptoms of schizophrenia, or nonaffective disorders exclude the diagnosis of major depression. In the DSM-III, the exclusion criteria are somewhat more lenient; mood-incongruent psychotic features or bizarre behavior exclude the diagnosis if they predominate the clinical picture before or after the episode. A diagnosis of major depression is excluded if the episode is superimposed on schizophrenia or paranoid disorders.

4. The RDC requires impairment in functioning or evidence of help-seeking behavior for a diagnosis of major depression.
These requirements were deleted in the development of the DSM-III because its developers thought that a full major depressive disorder was probably of social significance even in the absence of functional impairment or help-seeking behavior.

5. The RDC delineates ten subtypes of major depression: primary, secondary, simple, recurrent, psychotic, incapacitating, endogenous, situational, agitated, and retarded major depressive disorder. These subtypes are not mutually exclusive, and several of them have counterparts in the DSM-III (e.g., recurrent, psychotic, endogenous), but others do not (e.g., incapacitating, primary, secondary, situational). In general, those that do not have counterparts in the DSM-III were eliminated because research evidence challenged the clinical usefulness of the classification. For example, major depressions that are precipitated by a psychosocial stressor appear to be no different from those that are not. As a result, a situational subtype is not listed in the DSM-III. The primary/secondary distinction of the RDC distinguishes depressions preceded by another mental disorder from those not so preceded. When the DSM-III was developed, the prognostic implication of this distinction was unclear, and it was not included in the DSM-III. There are also slight differences between the RDC and the DSM-III in the criteria that define psychotic and endogenous (melancholia) features of major depression.

6. The two systems differ in the terminology they use to classify minor or chronic depressions. The RDC defines minor depressive disorder as a relatively sustained depressed mood that does not meet the criteria of major depression. The RDC requires evidence of functional impairment or of help-seeking behavior for this diagnosis. The RDC also defines intermittent depressive disorder, which essentially is minor depression interrupted by periods of normal mood. These two categories would be classified in the DSM-III as adjustment disorder with depressed mood, atypical depression, or dysthymic disorder.

In spite of the differences between the RDC and the DSM-III criteria for major depression, the two systems may identify similar groups of patients. Singerman et al. (1981) administered the NIMH Diagnostic Interview Schedule (DIS) to 216 persons and assigned diagnoses by the DSM-III, RDC, and Feighner Criteria. There was perfect
agreement for the diagnosis of a major depressive episode by the DSM-III and the RDC (87 cases).

SELECTION OF MAJOR DEPRESSION AND DYSTHYMIA AS MOS TRACER CONDITIONS

We have selected major depression and dysthymia as tracer conditions for the MOS based on their relatively high prevalence in the general population and in the practices of general medical physicians and mental health specialists, the degree of associated impairment in functioning, the rich research history for depressive disorders, and the high reliability of available diagnostic instruments for diagnosing these disorders. Each reason is discussed in detail in subsequent sections. We have selected two depressive disorders rather than major depression alone because the two disorders commonly coexist and because mild chronic depression is of great interest to general medical physicians. Because dysthymia is a recently defined disorder, relatively little research on it is available. Thus, much of this report describes previous literature on major depression or on unipolar depressive disorders combined.

The definition of depressive disorders used in the MOS is based on the DSM-III classification system. Our definition differs, however, from the DSM-III criteria in several respects. First, we do not apply the exact exclusion criteria of the DSM-III because of the difficulty and expense of identifying diagnoses of schizophrenia and organic causes of depression. However, we will exclude persons with preexisting mood-incongruent hallucinations or delusions and persons with evidence of severe cognitive dysfunction. As in the DSM-III, we will exclude persons with major depression due to uncomplicated bereavement.

Impact and Policy Relevance of Depression

Unipolar depression is associated with significant excess mortality, almost all of it due to suicide and accidents. Of persons with unipolar depression, 15 percent commit suicide (Guze and Robins, 1970; Coryell, Noyes, and Clancy, 1982). About half of all suicides occur in persons with an affective disorder. Suicide typically occurs in the first ten years after onset of major depression (Tsuang and Woolson, 1977). Chronic and intermittent depression may also contribute to the development of cardiovascular disease, cancer, and other physical diseases (Klerman, 1980a; Avery and Winokur, 1976).
Unipolar depression is associated with considerable morbidity, especially impairment in social and occupational functioning. Severe major depressive episodes are debilitating and may lead to weight loss, job disability, and severe financial problems. Because depressive episodes may last six months or more, especially if untreated, the strain on social support and personal resources can be considerable. Further, untreated or poorly treated depression may lead to dysthymic disorder and chronic social and occupational dysfunctioning. Even mildly depressed persons have a reduced ability to maintain intimate relationships and experience a reduced sexual interest and performance (Paykel and Weissman, 1973; Klerman, 1980b; Blumenthal and Diefman, 1975). The reduced social skills of depressed persons may alienate others. The subsequent decreased social support can further intensify depression (Billings, Cronkite and Moos, 1983), leading to a destructive positive feedback loop between poor social supports and depression.

Depressed persons use considerable health care resources. Depressive symptoms are associated with an increased use of psychotropic medications (Craig and Van Natta, 1978), especially minor tranquilizers. Depressed persons may present to general medical providers with somatic symptoms or nonspecific complaints (Klerman, 1980b). High use of medical care resources may result, if the depression is unrecognized (Katon, 1982). For example, among patients in primary care settings, those with mental disorders use about twice the medical services used by those without mental disorders (Houpt et al., 1980). Those with depressive disorders may be especially prone to use medical services (Watts, 1982; Weissman, Myers, and Thompson, 1981).

The policy significance of depressive disorders is enhanced by the relatively high prevalence of these disorders (see chap. 2). Epidemiologic studies indicate that 15 to 30 percent of adults experience a clinically significant depressive episode during their lifetime (Klerman, 1980a). Further, studies consistently show that depression-related diagnoses are among the most prevalent primary diagnoses assigned by physicians in office practice (Katon, 1982). However, less than a quarter of the patients with a depressive disorder are evaluated by a mental health specialist (Weissman, Myers, and Thompson, 1981). The remainder see a clergyman or general medical provider (nonpsychiatrist physicians, nurses) or receive no health care.

The care delivered by general medical providers to persons with depressive disorders is of particular importance to policymakers because of the high proportion of such persons who receive their only health care from these providers. Medical providers may detect clinically significant depression in a low proportion of their patients with
these disorders (Nielsen and Williams, 1980). Further, when depression is detected, psychotherapy is provided to only a small proportion of patients; antidepressant medication may be prescribed in subclinical dosages (Johnson, 1973, 1974). These findings are of particular interest because antidepressant medication and several specific types of psychotherapy are efficacious in the treatment of major depression (Weissman, 1979; Mindham, 1982).

Although the costs of care for major depression in the United States have not been estimated, $14 billion was spent on all mental health services in 1980 (McGuire, 1981). Of this, one-third was spent on ambulatory psychotherapy. These figures do not include societal costs of the social impairment, especially disability, associated with mental disorders.
Chapter 2

PREVALENCE OF DEPRESSION

PREVALENCE IN THE GENERAL POPULATION

Major depression is generally recognized as one of the most prevalent specific mental disorders. Klerman (1980a) has estimated that 18 to 23 percent of females and 8 to 11 percent of males have a major depressive episode during their adult years. Boyd and Weissman (1982) reviewed 24 community studies of the prevalence of unipolar depression. Across studies, the point prevalence for unipolar depression was between 1.8 and 3.2 percent for men and between 2.0 and 9.3 percent for women.

Studies of the prevalence of depression in general populations have differed widely in study design and assessment techniques. Prior to the 1970s, most studies used treatment records and/or self-report surveys of psychological symptoms to identify persons in treatment for depression with symptoms of depression. In the last decade, several major psychiatric epidemiology studies have been conducted that identify specific psychiatric diagnoses, such as major depression, according to specific classification systems, such as the RDC or DSM-III, using highly structured interview protocols, such as the SADS-L or the DIS. Weissman, Myers, and Harding (1978), in a study of RDC (SADS-L) psychiatric diagnoses in the New Haven general population, found a point prevalence of 4.3 percent for probable major depression and a prevalence of 6.9 percent for probable major and minor depression combined (Table 2.1). Roberts and Vernon (1982) used the RDC (SADS-L) to study a general population in Alameda County, California. The annual prevalence rate in their sample was 4.7 percent for probable major depression and 2.5 percent for minor depression; the point prevalences were 2.1 and 1.3, respectively.

Robins et al. (1984) and Myers et al. (1984) recently reported data on the prevalence of specific psychiatric disorders, including major depression, in three sites of the Epidemiologic Catchment Area Program (ECA) of NIMH. The ECA is the first study of the prevalence of the DSM-III diagnoses in the general population. The assessment instrument was the DIS (Regier and Meyers, 1984). The lifetime prevalence of major depressive episode varied from 3.7 to 6.7 in the three sites, while that of dysthymia varied from 2.1 to 3.8. By contrast, the
Table 2.1
Prevalence of Current RDC Diagnoses

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Percentage with Definite or Probable Disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major depression</td>
<td>4.3</td>
</tr>
<tr>
<td>Minor depression</td>
<td>2.6</td>
</tr>
<tr>
<td>Manic disorder and hypomania</td>
<td>0.0</td>
</tr>
<tr>
<td>Generalized anxiety</td>
<td>2.6</td>
</tr>
<tr>
<td>Phobic disorder</td>
<td>1.4</td>
</tr>
<tr>
<td>Panic disorder</td>
<td>0.4</td>
</tr>
<tr>
<td>Obsessive/compulsive disorder</td>
<td>0.0</td>
</tr>
<tr>
<td>Alcoholism</td>
<td>2.6</td>
</tr>
<tr>
<td>Drug abuse</td>
<td>1.0</td>
</tr>
<tr>
<td>Borderline features</td>
<td>0.2</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>0.4</td>
</tr>
<tr>
<td>Schizoaffective disorder</td>
<td>0.0</td>
</tr>
<tr>
<td>Unspecified psychosis</td>
<td>0.2</td>
</tr>
<tr>
<td>Other psychiatric disorder</td>
<td>1.2</td>
</tr>
</tbody>
</table>

SOURCE: Adapted from Weissman, Myers, and Harding (1978).

The lifetime prevalence of alcohol abuse/dependence was 11.5 to 15.7, and that of phobia was 7.8 to 23.3. The lifetime prevalence of major depressive disorder was roughly twice as high for women as for men. The lifetime prevalence was also significantly higher for adults under 65 than for adults 65 or older. Robins and her colleagues offered two explanations for this finding: Either the disorder is becoming more common, or those over 65 had a poorer memory for their depressive episodes. The lifetime prevalence rate for major depressive episode was not significantly different for blacks and nonblacks or as a function of education.

Myers et al. (1984) reported the six-month prevalence rates from the same data. Across the three sites, the prevalence of major depressive episode within the preceding six months varied from 2.2 to 3.5. As with the lifetime diagnosis, rates were higher for those under 65 than for those 65 or older and for women than for men (Table 2.2).

Subgroups at Risk

As noted above, depressive disorders are more prevalent in particular subpopulations (Boyd and Weissman, 1982). Most studies indicate that the prevalence in the general population is roughly twice as high
Table 2.2
SIX-MONTH PREVALENCE OF DIS/DSM-III MAJOR DEPRESSIVE EPISODE WITHOUT BEREAVEMENT, BY SEX, AGE, AND SITE
(In percent)

<table>
<thead>
<tr>
<th>Site</th>
<th>18-24</th>
<th>25-44</th>
<th>45-64</th>
<th>65+</th>
<th>Total Men</th>
<th>18-24</th>
<th>25-44</th>
<th>45-64</th>
<th>65+</th>
<th>Total Women</th>
<th>Both Sexes</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Haven</td>
<td>3.9</td>
<td>2.7</td>
<td>1.4</td>
<td>0.5</td>
<td>2.2</td>
<td>6.1</td>
<td>7.4</td>
<td>2.2</td>
<td>1.6</td>
<td>4.6</td>
<td>3.5</td>
</tr>
<tr>
<td>Baltimore</td>
<td>1.1</td>
<td>1.6</td>
<td>1.5</td>
<td>0.3</td>
<td>1.3</td>
<td>3.0</td>
<td>4.5</td>
<td>2.4</td>
<td>1.3</td>
<td>3.0</td>
<td>2.2</td>
</tr>
<tr>
<td>St. Louis</td>
<td>1.1</td>
<td>2.8</td>
<td>1.3</td>
<td>0.1</td>
<td>1.7</td>
<td>5.2</td>
<td>5.2</td>
<td>4.9</td>
<td>1.0</td>
<td>4.5</td>
<td>3.2</td>
</tr>
</tbody>
</table>

SOURCE: Adapted from Myers et al. (1984).

for women as for men. Weissman and Klerman (1977) reported a similar sex difference in the prevalence of RDC depressive disorders in treatment settings.

Boyd and Weissman (1982) described an interaction of age and sex in the prevalence of unipolar depression. The peak prevalence in women is between ages 35 and 45. For men, the prevalence seems to increase gradually with age. However, Robins et al. (1984) and Myers et al. (1984) did not find such an interaction using data from the ECA. Different socioeconomic classes appear to have the same prevalence of depressive disorders, although one study found that, among women with children, those from the lower classes have a higher rate of unipolar depression (Brown and Harris, 1978).

In addition to these sociodemographic factors, it is now well established that aspects of the family or personal history contribute to a high risk for unipolar depression. First-degree relatives of persons with unipolar depression are roughly twice as likely to have unipolar depression as are the first-degree relatives of matched controls (Winokur, 1979). The first-degree family members of alcoholics are also more likely to have a unipolar depression. Further, persons without close significant others may be especially prone to depression. For example, Brown and Harris (1978) found that women who did not have intimate others were four times as likely to develop depression after a stressful life circumstance as were those with intimate relationships.

In general, life stress has been associated with an increased prevalence of unipolar depression (Paykel, 1982). However, depressed persons may be more likely to perceive or report stressful events, making it difficult to evaluate the causal nature of the relationship. Finally, women appear to be at greater risk for unipolar depressive disorders.
and other mental disorders for six months after giving birth (Weissman and Klerman, 1977; Boyd and Weissman, 1982).

TREATED PREVALENCE

The majority of persons with depressive disorders do not receive specific treatment for depression. Weissman, Myers, and Thompson (1981) found that only 6 percent of those with a major or minor depression by the RDC (SADS-L) ever saw a psychiatrist, and only 17 percent visited any mental health specialist. Within the six months prior to the study, only 14 percent sought help from clergy, medical social workers, or nurses, and 9 percent received outpatient mental health care from a nonpsychiatrist physician. Shapiro et al. (1984) reported the use of health services by participants in the ECA. Of those with an affective disorder in the preceding six months by the DSM-III (DIS), 31.2 to 31.7 percent had received mental health care from a mental health specialist or general medical provider.

Although relatively few persons with depressive disorders receive mental health care, almost all are in contact with the health care system. For example, in the ECA, 76.5 to 78.6 percent of those with a recent affective disorder had received some health care in the six months prior to their ECA interview.

Depression in General Medical Practices

Depression is thought to be among the most common presenting problems in general medical practice (Watts, 1982; Katon, 1982). Unfortunately, studies have not reported data on the prevalence of specific mental disorders in the offices of private fee-for-service mental health specialists or nonpsychiatrist physicians. However, studies have used the RDC (SADS-L) to examine the prevalence of specific mental disorders in medical outpatient clinics and inpatient settings.

Hoeper et al. (1979) studied the mental disorders in a primary-care population of a multispecialty clinic. In a two-stage identification process, they first screened patients with the General Health Questionnaire (GHQ), a survey measure of symptoms of psychological distress. Second, they administered the SADS-L to 83 percent of the GHQ positives (i.e., those with a score over 4) and 12 percent of the GHQ negatives. According to the RDC (SADS-L), the weighted prevalence of any mental disorder was 26.7 percent. The estimates for selected specific disorders were: major depression (5.8 percent), phobic disorder (5.8 percent), intermittent depression (5.0 percent), labile personality
(3.7 percent), minor depression (3.4 percent), cyclothymic personality (2.0 percent) and generalized anxiety (1.6 percent). Hooper and his colleagues noted that the prevalence of mental disorders in a hospital-based general medical clinic might be higher than that in private office practice.

Raft et al. (1975) used self-report measures and a clinical interview to identify depression in outpatients who were prescribed psychotic medications at a general hospital. Of these patients, 29 percent fit the Feighner criteria for depression and an additional 12 percent had "masked depression" (i.e., presenting with physical symptoms rather than depressed mood).

Other studies of depression in primary-care settings have examined the prevalence of depressive symptoms (from self-administered questionnaires), physician diagnoses of depression, or physicians' reports of depression (Katon, 1982). Several studies reported the prevalence of combined anxiety and depression. For example, Nielson and Williams (1980) evaluated depressive symptoms as measured by the Beck Depression Inventory (BDI) in medical outpatients of a prepaid health program; 5.5 percent had at least moderate depression, and 12.2 percent had at least a mild depression. Salkind (1969) found that 25 percent of 80 consecutive patients in his own practice were depressed (as measured by a cutoff score of 17 on the BDI). Seller, Blascovich, and Lenkei (1981) found that 14 percent of 222 outpatients in a university family practice clinic were depressed (using a cutoff score of 21 on the BDI). Studies of medical inpatients using the BDI identified 20 to 24 percent as depressed (Schwab et al., 1967; Moffic and Paykel, 1975). Although these studies did not include an assessment of the prevalence of specific disorders, they emphasize the importance of depressive phenomena in medical practice.

Depression is probably about as common in general medical settings as the most common specific physical disorders. For example, Rosenblatt et al. (1983) analyzed data from the 1977 and 1978 National Ambulatory Medical Care Survey (NAMCS). They determined the most frequent primary diagnoses (assigned by the treating physicians) in ambulatory care encounters. An anxiety-depression cluster of diagnoses was the fifth most frequent diagnosis group and accounted for 3.1 percent of all encounters (including outpatient psychiatric visits). By comparison, hypertension accounted for 4.4 percent of encounters, ischemic heart disease for 2.6 percent, and diabetes mellitus for 1.7 percent. Their estimate of the percentage of encounters related to depression or anxiety is probably conservative, because they analyzed only the primary diagnosis for each encounter and because physicians may
under-report mental health diagnoses to protect patient confidentiality (Towery, Sharfstein, and Goldberg, 1980).

Prevalence in Outpatient Mental Health Specialty Settings

Published studies have not included the prevalence of specific depressive disorders as defined by RDC or DSM-III in private mental health specialty practice. Mezzich, Coffman, and Goodpastor (1982) obtained DSM-III diagnoses on 1,111 consecutive patients who applied for care at a 24-hour psychiatric walk-in clinic during a six-month period. Of these patients, 20.7 percent had a diagnosis of primary affective disorder. While information is not available on the distribution of subclassifications, presumably most of these would be either major or minor depressive disorders or dysthymia. By comparison, 12.2 percent had a primary schizophrenic disorder, 2.5 percent an anxiety disorder, and 11.2 percent an adjustment disorder.

Langsley and Lebaron (1974) studied 3,151 patients seen by 97 members of a state psychiatric society during a six-month period. The most common diagnosis (assigned by the psychiatrists) was schizophrenia (18.2 percent of patients); 14.0 percent had a diagnosis of depressive neurosis.

Marmor (1975) summarized the results of a survey conducted for the American Psychiatric Association (APA) and the National Association for Mental Health. The study included a 10 percent random sample of psychiatrists who had indicated on an earlier APA survey that they spent 15 or more hours per week seeing private patients. The response rate was 75 percent; 29 percent of respondents were psychoanalysts. The psychiatrists were asked to provide information on themselves and the last 10 patients they had seen. Of the patients described, 43 percent were male, and the mean age was 35. Depressive neurosis was the most common primary diagnosis, assigned to 23 percent of patients. Of these 23 percent, over two-thirds were estimated to have moderate to severe functional impairment and thus may have had a major depressive episode. However, only a small percentage of the patients had a diagnosis of affective psychosis.

In summary, although few studies have used specific diagnostic criteria to determine the prevalence of depressive disorders in outpatient medical settings or practice of mental health specialists, the existing studies indicate that these disorders are among the most common mental disorders in a general population and that a diagnosis of "depression" is among the most common diagnoses assigned by providers in medical and mental health outpatient settings.
Chapter 3

THE COURSE OF ILLNESS

A substantial literature exists on the course of depressive illness. However, the generalizability of past data is uncertain for two reasons. First, some studies span several periods of change in the definition and treatment of depression. Second, in some studies, the influence of treatment on the observed course of illness has not been clear. Fortunately, several studies have recently provided data on the course of major depressive illness as defined by the RDC.

MAJOR DEPRESSION

Major depression is often considered a recurrent, episodic illness. The course of illness can be described in terms of the age at onset, the duration of episodes, and the rate of recurrence or relapse.

The age of onset for major depression appears to be evenly distributed throughout adult life, and can even occur in infancy (Klerman, 1980a). This wide range of age at onset may result partly from the heterogeneity of the etiologies of unipolar disorders. Specific subtypes may eventually be shown to have more specific ages of onset (Coryell and Winokur, 1982).

The duration of an episode of major depression depends on both the natural history of the illness and the treatment administered during the episode. On the average, depressive episodes persist longer in persons with unipolar than with bipolar disease. Prior to specific pharmacologic treatment, bipolar episodes had a mean duration of 7 to 13 months (Coryell and Winokur, 1982). Recent studies of bipolar episodes have reported a mean duration of 2 to 3 months (Angst et al., 1973). A later age of onset is associated with longer episodes and more frequent recurrences.

While major depression is often considered a recurrent illness, some persons have only a single episode. Coryell and Winokur (1982) summarized data from eleven studies that reported the percentage of unipolar depressives who had only one depressive episode. The length of follow-up in these studies varied from 1 to 30 years. Across studies, the percentage of persons with one episode of unipolar depression

22
varied from 5 to 60 percent, but most of the studies reported percentages above 23 percent.

Keller and his associates (1982a,b; 1983a,b) reported data on the rate of relapse for patients enrolled in the NIMH Collaborative Study on the Psychobiology of Depression. A unipolar depressive disorder (classified by the RDC) was the diagnosis for 81 percent of the sample, and 74 percent had a primary major depressive disorder. About 40 percent had a psychotic depression. At entry to the study, 31 percent had their depressive episode superimposed on a chronic (milder) depression. To be considered recovered, patients had to have no symptoms of major depression, or only one or two mild depressive symptoms for eight consecutive weeks.

Using a life-table analysis, Keller and his associates found that the probability of a relapse was highest in the first three months following recovery (a monthly rate of 9 percent for each of the first three months). The probability of relapse steadily declined thereafter to about 1 percent per month. Roughly one-quarter of the patients who recovered in the first year had a relapse within six months of recovery.

Keller et al. (1982b) determined the predictors of the rate of relapse. Patients whose depressive episode was superimposed on a chronic underlying depression were much more likely to relapse within four weeks of recovery than were patients without a preexisting chronic depression. Patients with a history of three or more previous depressive episodes were also more likely to relapse during the first six months after recovery. This was especially true for those without an underlying chronic depression.

In a subsequent analysis, Keller et al. (1983a) determined that those with secondary depressions (RDC type) and an older age of onset (of a first episode) were more likely to relapse. Most of the patients with secondary depressions had prior histories of drug or alcohol abuse. The authors also determined that the duration of a well period following relapse is generally not as long as that of the previous well period.

Keller et al. (1983b) presented data on the course of a major depressive episode for those with and without a history of chronic minor depression. Recovery time for the acute episode was significantly shorter for those with underlying chronic depression. For example, 83 percent of those with a chronic underlying depression had recovered from acute episodes by 52 weeks after entry into the study, in contrast to 69 percent of those without the underlying depression. However, only 43 percent of those with a chronic underlying depression had recovered from their chronic disorder at 104 weeks after entry. As the authors noted, it is important to distinguish between recovery from the acute episode and from the underlying chronic depression when
evaluating the outcomes of patients with "double depression" (i.e., both a major depressive disorder and chronic depression). Keller et al. (1983b) found that patients with double depression are at much higher risk of having multiple, recurrent episodes.

COURSE OF DYSTHYMIC DISORDER

Dysthymic disorder is a newly defined disorder (i.e., it has no direct counterpart in previous classification schemes). As a result, conclusions about the course of illness for this disorder are drawn from studies of neurotic depression, nonpsychotic depression, depressive personalities, etc. (Klerman, 1980b).

Many patients with chronic depression date the beginning of their depressive symptoms to young adulthood, but the most common period of onset is maturity and old age. Typically, there is an insidious onset of depressive symptoms and impairment in social function prior to the development of the dysthymia. Klerman (1980b) noted that hospitalization is currently uncommon for dysthymic disorder. The most important impairments are in the area of social functioning, manifested by marital problems, reduced ability in sexual functioning, reduced interest in hobbies, and social withdrawal.

Implications for the MOS Design

Because major depression is an episodic disorder, health status assessment at an arbitrary time (e.g., exit) after entry into the MOS may not indicate a valid outcome of the course of a given episode. That is, a patient may recover and then have another episode by the end of the MOS. While it can be argued that a recurrence is evidence of poor management (e.g., the lack of prophylactic antidepressants), some patients with adequate management may also have recurrences. The outcomes measured at the end of the study should be evaluated with consideration of the history of episodes.

Keller and his associates were able to describe the clinical course of a depressive episode because the patients were being closely followed by a clinical/research team in a psychiatric setting. As a result, they were able to determine recovery dates and the occurrence of a relapse. Such a strategy would not be feasible for the MOS.

Three strategies can be used in the MOS to determine the course of major depressive episodes. First, patients can be selected who have only recently developed a depressive episode (e.g., within six weeks). Second, at several points in time, patients can provide retrospective
accounts of the progression of their symptoms. Third, current depressive symptoms can be assessed at relatively frequent intervals to attempt to capture the fluctuations of the episodes. At a minimum, outcomes will be reassessed at six-month intervals.
Chapter 4

TREATMENTS OF DEPRESSION

The treatment of unipolar depressive disorders has been the focus of considerable research attention. Progress in the treatment of depression has been among the most significant accomplishments of mental health clinicians and researchers in the last three decades.

Treatments of depression can be classified into pharmacotherapy (i.e., tricyclic antidepressants, monoamine oxidase inhibitors, antipsychotics, lithium, and minor tranquilizers), electroconvulsive therapy (ECT), psychotherapy, and other treatments (e.g., sleep deprivation, psychosurgery). Several literature reviews of the efficacy of these approaches have recently appeared (Weissman, 1979; Davis, 1980; Paykel, 1982a; Rehm and Kornblith, 1979). Several of these authors noted the paucity of adequately controlled clinical trials, even for the efficacy of antidepressants and ECT. Important methodological limitations of many studies have included: diagnostic heterogeneity of samples, lack of a true placebo group, insufficient specification of treatment or outcome, and small sample sizes. Despite these limitations, there is widespread agreement concerning the efficacy of antidepressant medication, the role of drugs in prophylaxis of recurrent depression, and the efficacy of ECT for certain kinds of unipolar depression. Further, sufficient evidence has accumulated to indicate that specific psychotherapies are efficacious in treating unipolar depression. Moreover, the effects of psychotherapy may be additive to those of medication.

This chapter summarizes studies on the efficacy of treatment for major depression, partly based on the reviews in Paykel (1982a). Detailed information on the administration of psychotherapy is not provided because of the large number of available models. The chapter concludes by summarizing information on the size of treatment effects and clinical predictors of the response to treatment.
PHARMACOTHERAPY

Tricyclic Antidepressants

The tricyclic antidepressants have been used for over twenty years and are widely considered efficacious in treating depression. Imipramine and amitriptyline are the "gold standards" with which other drugs are compared, partly because they were the first to be tested against a placebo and because no other drug has been consistently shown to be superior to these in relieving depression.

There have been several comparisons of imipramine or amitriptyline with placebo (Ball and Kiloh, 1959; Kenning, Richardson, and Tucker, 1960; Friedman, Demowbray, and Hamilton, 1961; Roulet et al., 1962; Rees, Brown, and Benaim, 1961). The patients in these studies were inpatients or outpatients with severe depression. On the basis of their reviews of controlled trials of imipramine and placebo, Klerman and Cole (1965) and Rogers and Clay (1975) concluded that the effectiveness of imipramine in acute endogenous depression was proven, but its role with chronic neurotic depression was uncertain.

In a large proportion (roughly 40 percent) of properly conducted controlled trials, imipramine was not more effective than placebo (Mindham, 1982). However, studies affirming positive effects of imipramine outnumber those that do not. Further, the positive effects are generally large. For example, in one of the earliest studies of depressed outpatients (Ball and Kiloh, 1959), 74 percent of those with endogenous depression improved compared with 22 percent receiving placebo. Amitriptyline has also been demonstrated to be superior to placebo (Garry and Leonard, 1963; Skarbek and Smedberg, 1962). Its efficacy is comparable to that of imipramine.

In addition to studies of these two drugs, there have been numerous studies of other tricyclic antidepressants, comparing one drug with another or with a placebo. Examples of these drugs include desipramine, trimipramine, nortriptyline, protriptyline, opipramol, dothiepin, doxepin, butriptyline, and clomipramine. These drugs vary in their clinical potency and side effects, but are of similar efficacy (Mindham, 1982).

Mindham reviewed clinical trials of other substances resembling the tricyclics in chemical structure (iprindole, dibenzepin, maprotiline and mianserin) as well as trials of drugs chemically distinct from the tricyclics (vloxazine, flupenthixol, tofenacin, and nomifensine). Most appear to be similar to imipramine in efficacy. However, flupenthixol appears to be especially effective for the treatment of symptoms of
anxiety and depression, but it has not yet been tested in severely depressed patients. Tofenacin has also not yet been adequately studied.

The tricyclics have also been studied as prophylaxis against recurrences of depression (Mindham, Howland, and Shepherd, 1973; Paykel et al., 1975; Prien, Klett, and Caffey, 1973; Klerman et al., 1974; Coppen et al., 1978; Quitkin et al., 1978). These studies demonstrate that continued use of tricyclics after a recovery from depressive episode significantly decreased the risk of relapse. For example, Mindham, Howland, and Shepherd (1973) found that 50 percent of patients on a placebo relapsed within six months, but only 22 percent of those on antidepressants did so.

In the last decade, numerous studies have evaluated the usefulness of plasma levels of tricyclic drugs for monitoring therapy. To date the clinical usefulness of blood levels has not been adequately demonstrated (Mindham, 1982).

Course of Treatment With Tricyclics. Tricyclics are generally started at low doses (25 mg BID or TID) and gradually increased to a therapeutic dose. In adults, the equivalent of 150 mg of imipramine is considered a minimum dose for an adequate trial (Davis, 1980). The dose may be increased to 200–300 mg/day if necessary. A clinical response may be seen in the first week, but perhaps not for four to six weeks. The tricyclics appear to be most effective for the associated symptoms of depression, especially for "endogenous" symptoms (e.g., melancholia). Common side effects, especially at high doses, are sedation and atropinic effects such as dry mouth, sweating, constipation, and urinary retention. The most serious side effects are cardiovascular, including orthostatic hypotension, palpitations, arrhythmias, and tachycardia.

If a clinical response occurs, medication should be continued for several months (recommendations include 8 to 10 weeks, six months, etc.). Continued treatment depends in part on the clinical picture. As noted by Mindham (1982), patients with incomplete resolutions of a depressive episode may benefit from continued administration of antidepressants.

A number of studies have addressed the choice of specific antidepressants (Mindham, 1982). The greatest evidence for efficacy exists for amitriptyline and imipramine. Some antidepressants have a sedative effect (amitriptyline, trimipramine), some a stimulant effect (nortriptyline, desipramine), and some neither effect (imipramine). Drugs with sedative properties are recommended for patients with agitated depression; drugs with stimulant properties, for persons with retarded depression. While some of the newer tricyclics and other antidepressants are reported to have fewer side effects, there have not
been enough studies to demonstrate a clear preference among the tricyclic antidepressants.

**Monoamine Oxidase Inhibitors**

The monoamine oxidase inhibitors (MAOIs) are also used in the treatment of depression, but their use is more controversial (Nies and Robinson, 1982). They are classified into hydroxines, such as phenylazine, and non-hydroxines, such as tranylcypromine.

MAOIs are not widely used (Nies and Robinson, 1982). First, there are some serious side effects such as hypertensive crises, hepatotoxicity, and a lupus-like syndrome. The hypertensive crises occur most commonly when foods containing tyramine (e.g., cheddar cheese, red wines) or drugs containing pressor agents (e.g., decongestants) are ingested. Second, there has been concern about the efficacy of these drugs for depression. The early controlled trials of MAOIs suffered from serious methodologic flaws (Robinson et al., 1973). Nevertheless, MAOIs have been found to be more effective than placebos for outpatients with mild or moderate depression (Robinson et al., 1973; Ravaris et al., 1976; Johnstone and Marsh, 1973). The MAOIs appear to be particularly effective for patients with atypical features (e.g., weight gain, depression worse in the evening) and mixed anxiety-depression syndromes (Nies and Robinson, 1982; Paykel et al., 1979; Liebowitz et al., 1984).

Clinical guidelines (doses, length of treatment) are not as well established for the MAOIs as for the tricyclic antidepressants (Davis, 1980). Starting doses (which vary with the particular drug) are increased gradually over two to three weeks or longer. MAOIs should be reserved for intelligent, cooperative patients in whom there is little risk of an overdose (Nies and Robinson, 1982).

**Other Drug Treatment**

As noted by Davis (1980), phenothiazines have been demonstrated to be effective antidepressants in several controlled clinical trials. These drugs are believed to be especially useful for psychotic, agitated depression (Fink, Klein, and Kramer, 1965) or for nonpsychotic depression with symptoms of both anxiety and depression. In such instances, they are often used in combination with a tricyclic antidepressant.

Minor tranquilizers are used for the anxiety and insomnia associated with depression. However, those drugs should only be used for a few weeks because of the development of tolerance and the risk of increased insomnia, depression, and suicide. Minor tranquilizers are
useful for persons having trouble falling asleep, but not for persons with early morning awakening (Davis, 1980; Klerman, 1982).

Stimulants have also been used to treat depression. Currently, methylphenidate is occasionally used in conjunction with tricyclics for its mood-stimulating effect. However, problems with dependence, side effects, tolerance, and low clinical efficacy have virtually precluded the usefulness of stimulants in treating depression.

ELECTROCONVULSIVE THERAPY

Electroconvulsive therapy (ECT) was a widely used treatment for many psychiatric inpatients (especially for mania, schizophrenia, and depression) prior to the availability of antidepressants and neuroleptics (antipsychotics). Considerable public controversy has surrounded the use of ECT because of beliefs that it is dangerous and used for mind-control. Modified ECT, as practiced today, is an extremely safe and effective procedure. In fact, ECT is the most appropriate treatment for patients with marked psychomotor retardation or agitation, and for acutely suicidal or delusional patients (Kiloh, 1982).

There are several acceptable controlled trials of bitemporal ECT against placebo (Ulett, Smith, and Gleser, 1956; Kiloh, Child, and Latner, 1966a,b; Greenblatt, Grosser and Wechsler, 1964; Clinical Psychiatry Committee, 1965; Wittenborn et al., 1962; Wilson et al., 1962–1963). The patient populations generally were thought to have endogenous or neurotic depressions by the study investigators. The response rate for ECT was generally 70 to 90 percent, as compared with 60 to 70 percent for the antidepressants, and 25 percent or so for placebo groups. In most comparisons with antidepressant drugs, ECT was efficacious for a greater percentage of patients, and the onset of improvement was much faster (Kiloh, 1982; Brandon et al., 1984).

Studies addressing the indications for ECT have shown that the best response occurs for patients with endogenous features (e.g., melancholia) and without a history of precipitating stressful events. Coryell and Zimmerman (1984) recently found, however, that melancholia (DSM-III) did not predict the outcome of ECT but that familial subtype did. Studies have reached conflicting conclusions as to whether ECT is more effective for depressed patients with psychotic symptoms than for other depressed patients (Rich et al., 1984; Brandon et al., 1984).

Much work has been done on perfecting the technique, selecting the optimal stimulus, and evaluating unilateral ECT against bilateral ECT. The principal side effect is memory loss, which is usually temporary and not severe; this can be minimized with unilateral ECT. The
mortality risk of a course of ECT is about 0.03, despite the fact that ECT is usually reserved for very disturbed or acutely medically ill patients (Kihol, 1982).

ECT is customarily given two to three times a week, but unilateral ECT can probably be given more often. The total number of treatments required varies, but typically is five to eight. A course can be extended, if necessary, to 20 treatments or beyond. Extra treatments after relapse are not helpful (Kiholh, 1982).

LITHIUM

Lithium has not yet been widely studied for its antidepressant value (Coppen, Metcalfe and Wood, 1982). Several studies suggest that it has value in acute depression (Mendels, Secunda, and Dyson, 1972; Watanabe, Ishino and Otsuki, 1975), but it may be more effective for the depressive episodes of bipolar patients (Coppen, Metcalfe and Wood, 1982; NIMH Consensus Development Conference, 1984).

Currently, the chief indication for lithium in unipolar depression is in prophylaxis of recurrent depression (Coppen, Metcalfe and Wood, 1982). Schou (1974) has reviewed studies of lithium prophylaxis. Lithium prophylaxis appears to be of value only for patients with primary “endogenous” depression, not mild or secondary depression. The prophylactic effect appears to be strong. For example, Bastrup et al. (1970) found that none of those patients maintained on lithium after a depressive episode experienced a relapse, as opposed to 53 percent of unipolar patients without lithium. Coppen et al. (1971) followed persons with unipolar depression for over two years. Those on lithium spent 5 percent of their time with depressive symptoms, as opposed to 30 percent for those without lithium.

Until recently, lithium prophylaxis was reserved for patients with at least three unipolar episodes. It is not clear when lithium prophylaxis can be stopped without risk of relapse (Angst, 1980).

Recently, several psychiatrists have advocated the use of lithium carbonate and thyroid (T₄ or T₃) in combination with a tricyclic antidepressant when the latter has not produced a full remission. The efficacy of this combination has not yet been fully evaluated.

Lithium can result in dry mouth, thirst, polydipsia, and tremor. Only about 5 percent of patients experience severe symptoms. While many patients complain of poor memory, a similar proportion complain of poor memory when using antidepressants. Hyperthyroidism and chronic nephropathy may occur. Because of the increased incidence of congenital anomalies, lithium is contraindicated in pregnancy and in
those patients with heart failure, kidney disease, and thyroid function (Coppen, Metcalfe, and Wood, 1982).

PSYCHOSURGERY

Psychosurgery is a rare treatment for depression. However, it apparently is indicated in cases of intractable depression (e.g., for ten years) with severe social dysfunction. The appropriate procedure is stereotactic subcaudate tractotomy, although stereotactic limbic leucotomy has also been used. Relapse after surgery is rare (Kelly, 1982).

PSYCHOTHERAPIES

An extensive literature describes the nature and efficacy of psychotherapy in general. Despite the considerable historical controversy over its effectiveness, several recent reviews concluded that psychotherapy is effective for relief of mental disorders (Advisory Panel on Psychotherapies, 1980; American Psychiatric Association's Commission on Psychotherapies, 1982; Garfield and Bergin, 1978). Until recently, few studies addressed the efficacy of specific psychotherapies for particular mental disorders such as major depression.

There are many reasons for this lack of data. First, operational definitions of mental disorders were developed only in the last decade. Without such definitions, most studies used heterogeneous patient groups. Second, the available types of psychotherapy are very numerous, operational definitions of goals and methods have only recently been developed for a few specific forms of psychotherapy. Third, the number of studies required to establish the efficacy of treatment for even a single therapy may be quite large. For example, a large percentage of the controlled studies of imipramine failed to show its superiority over placebo.

Currently, there are a number of ongoing, large-scale studies of the efficacy of specific psychotherapy for treatment for depression, e.g., as part of the NIMH collaborative studies of depression. Weissman (1979) and Klerman and Schechter (1982) provide excellent reviews of the efficacy of psychotherapy in comparison with drug therapy for the treatment of depression. Weissman reviewed 17 randomized controlled trials that tested the efficacy of behavioral, cognitive, group, interpersonal, and marital therapy in samples of depressed patients. The treatments are defined in Table 4.1. While the studies did not use the RDC or DSM-III criteria, the patient samples consisted exclusively of persons with depressive illnesses.
Table 4.1
DEFINITIONS OF FIVE PSYCHOTHERAPIES

1. Cognitive therapy: a treatment that focuses on correcting maladaptive cognitive patterns, such as low self-concept.

2. Behavioral therapy: identifies the reinforcers for specific depressive behaviors; provides social skills therapy, verbal skills, problem solving, enhances pleasant experiences.

3. Interpersonal psychotherapy (ITP): goals are to improve functioning by enhancing coping with stress, restoring morale, and attending to social and personal problems.

4. Group therapy: a therapist and a group of patients focus on the problem of each patient and on the group process.

5. Marital therapy: the goals include facilitating communication, understanding and classifying the nature of the relationship, sharing and developing expectations, etc.

SOURCE: Adapted from Weissman (1979).

Weissman examined the evidence for the efficacy of the five types of psychotherapy and for all psychotherapies combined. Five studies of cognitive therapy found it to be superior to behavior therapy, insight-oriented group psychotherapy, or being on a waiting list; one study showed cognitive therapy to be superior to imipramine. The four studies of behavior therapy all reported significant improvement over being on a waiting list. Interpersonal psychotherapy (ITP) was more efficacious than either placebo or minimal treatment; group and marital therapy were superior to placebo.

Considering all 17 studies, Weissman found that in eight of nine comparisons with placebo, waiting list, or minimal treatment, a specific psychotherapy treatment was superior. Five of the studies compared drugs and psychotherapy. In one such study, psychotherapy was equal to drugs; another found no difference; in three studies, drugs were superior to psychotherapy for symptom reduction and prevention of relapse, but psychotherapy was superior to drugs for problems of mood, guilt, and social function.

In most studies, the effects of drugs and psychotherapy were additive. That is, the best response was found in patients having both psychotherapy and drugs. This conclusion is consistent with that of Luborsky, Singer, and Luborsky (1975), who reviewed controlled studies of the efficacy of psychotherapy in general.

A recent set of studies from the Boston-New Haven Collaboration Depression Project has shed further light on the relative value of drugs and psychotherapy in the treatment of depression. Weissman et al.
(1979) studied outpatients with unipolar, nonpsychotic, acute, primary, major depression. The subjects were randomly assigned to amitriptyline, short-term ITP, a combination of the two treatments, or non-scheduled supportive psychotherapy. The 81 patients were followed for 16 weeks. The combination group had the best outcome (by a variety of measures); the nonscheduled supportive therapy had the worst outcome; and drug therapy and ITP alone had an intermediate efficacy.

In this study, drugs and therapy had differential effects on outcomes (DiMascio et al., 1979a). Amitriptyline had its effects on the vegetative-motor symptoms of depression, but psychotherapy affected mood, suicidal ideation, occupational functioning, and interests. Further, the effects of psychotherapy occurred later—at four to eight weeks instead of in the first few weeks.

Prusoff et al. (1980) found that the efficacy of treatment depended on the RDC subtype of depression (endogenous versus situational). The combination treatment was efficacious for both groups of patients; situational depression responded to both treatments together or drugs alone but not to psychotherapy alone.

Recently, other comparisons of treatment for depression have emerged. Bellack, Hersen, and Himmelhoch (1981) compared amitriptyline, social skills training plus amitriptyline, social skills training plus placebo, and psychotherapy plus placebo. Significant clinical improvement was found for each treatment group. However, the rate of dropout from the study was highest for the group assigned to drug treatment alone and lowest for the group assigned to social skills training. As a result, when considering all patients assigned to any treatment group, the social skills group had a much higher proportion of patients who improved. In this study, the combination of drugs and psychotherapy was not more effective than either treatment alone.

Zeiss, Levinsohn, and Munoz (1979) compared interpersonal psychotherapy, behavioral treatment, cognitive therapy, and pleasant events enhancement. All the methods led to improvement, but the outcomes were not specific to the objectives of a particular method. Rather the improvement was general over a variety of outcomes. Roth et al. (1982) found that a combination of antidepressants and self-control therapy (a treatment related to both cognitive therapy and behavior modification) resulted in more improvement in depressive symptoms than did self-control therapy alone.

Data on the efficacy of psychodynamically oriented individual psychotherapy for treatment of specific depressive disorders have not yet been reported. The general efficacy of these treatments has been evaluated, however (see Garfield and Bergin, 1973). Because depression is the most prevalent mental disorder in most psychiatric
Treatment settings, past studies of the general efficacy of psychotherapy probably included many persons with dysthyMIC disorder and major depression (Klerman and Schechter, 1982). There have been no studies of the efficacy of the counseling provided by nonpsychiatric physicians for the treatment of depressive disorders.

THE SIZE OF THE EFFECT OF SPECIFIC TREATMENTS FOR DEPRESSION

Studies of the efficacy of treatments for depression varied in outcome measures, patient populations (self-selected or clinician-referred outpatients and inpatients), and study settings (research wards, hospital outpatient clinics). Thus it is not surprising that estimates of the magnitude of effects of specific treatments vary considerably among studies. Several studies reported mean outcome scores for different treatment groups, but did not report standard errors. As a result, the precision of some estimates cannot be evaluated.

Many studies evaluated the efficacy of treatment in terms of the percentage of patients who recover (as indicated by clinical assessments or cut-off scores on measures of depression). However, efficacy has also been expressed in terms of rates of treatment failure, determined from life-table analysis (Weissman et al., 1979), the amount of improvement (at outcome assessment) over a pretreatment measure of depression, and changes in depression as a percentage of the maximum possible improvement on a particular measure of depression.

Several studies have used the technique of meta-analysis to estimate effects of specific treatments of mental disorders across several outcome studies. A comprehensive review of the size of the effect of psychotherapy in general was conducted by Smith and Glass (1977). They evaluated 375 controlled studies of psychotherapy. For each study, they calculated the mean difference between treatment and control groups for each outcome measure, divided by the standard deviation of the control group in that measure. Across all studies, the average effect size for psychotherapy was 0.68 of a standard deviation. They also examined the effect size for specific types of outcome measures. For anxiety and self-esteem, the average effect size was 0.97 and 0.90 of a standard deviation, respectively. For social adjustment, the average effect size was 0.31 of a standard deviation.

Steinbrueck, Maxwell, and Howard (1983) conducted a meta-analysis of 56 outcome studies to compare the effectiveness of psychotherapy and drug therapy for treatment of adults with nonpsychotic unipolar depression. Their analysis excluded studies where the
treatment was a combination of drug therapy and psychotherapy. The specific types of psychotherapy studied were behavioral therapy, social-learning interpersonal therapy, cognitive therapy, a combination of these three methods, and marital therapy. The studies analyzed included 35 drugs, most commonly imipramine and amitriptyline. The average effect size for psychotherapy was 1.22 standard deviations; for drugs, the average effect size was 0.61 standard deviations. This difference is statistically significant. The only other characteristic of study design that was significantly related to effect size was number of weeks in treatment.

To illustrate the magnitude of treatment effects for particular outcome measures (e.g., Hamilton Rating Scale), the author reviewed several recent studies of the efficacy tricylic antidepressants (Feighner et al., 1983; Guelfi et al., 1983; De Wilde, Mertens, and Wakelin, 1983; Nelson et al., 1982; Brown et al., 1982). The approximate range of effects of a four-to-six-week trial of antidepressant in these studies is a 40 to 70 percent reduction in the Hamilton Rating Scale (HRS), a 40 to 60 percent drop in the Raskin Clinical Rating Scales and in clinical global impressions, and a 25 percent drop in the total score of the Hopkins Symptom Check List (HSCL).

Weisman et al. (1979) and DiMascio et al. (1979a) used symptomatic failure rates (determined from life table analysis) as an outcome measure. In their study, patients with nonbipolar, nonpsychotic, acute primary major depression were randomly assigned to ITP alone, ITP plus pharmacotherapy, pharmacotherapy alone, or unscheduled psychotherapy alone. Symptomatic failure rates at 16 weeks were 4 percent for the ITP plus drug group, 12 percent for ITP alone, 21 percent for pharmacotherapy alone, and 48 percent for unscheduled psychotherapy. The change in mean total scores for the Raskin Depression Scale (range 3 to 15) and the HRS are given in Table 4.2. Although standard errors are not provided, we observe about a 40 percent reduction in HRS scores for ITP alone and pharmacotherapy alone and roughly a 70 percent reduction for these two groups combined.

Blackburn et al. (1981) randomly assigned 64 patients with major depression to cognitive therapy alone or cognitive therapy plus antidepressants. The patients included psychiatric inpatients and outpatients in general practice. They evaluated efficacy in terms of percentage change in BDI and HRS. Table 4.3 gives the percentage change in depression scores, adjusted for duration of illness after 20 weeks of treatment.

For each group, the percentage change for both measures was large and in a range (50 to 80 percent) similar to that reported by DiMascio
Table 4.2
MEAN SCORES FOR FOUR TREATMENT GROUPS,
AT ENROLLMENT AND AT 16 WEEKS

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Raasen Depression Scale</th>
<th>Hamilton Rating Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Initial</td>
<td>16 Weeks</td>
</tr>
<tr>
<td>ITP + drug</td>
<td>8.5</td>
<td>-4.8</td>
</tr>
<tr>
<td>ITP alone</td>
<td>6.6</td>
<td>6.2</td>
</tr>
<tr>
<td>Drug alone</td>
<td>8.5</td>
<td>6.7</td>
</tr>
<tr>
<td>Nonscheduled treatment</td>
<td>8.7</td>
<td>7.5</td>
</tr>
</tbody>
</table>

SOURCE: Adapted from DiMascio et al. (1979a), Table 4.

Table 4.3
MEAN SCORES FOR THREE TREATMENT GROUPS BY OUTCOME MEASURE AND PRACTICE SETTING

<table>
<thead>
<tr>
<th>Type of Treatment</th>
<th>Beck Depression Inventory</th>
<th>Hamilton Rating Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>General Practice</td>
<td>General Practice</td>
</tr>
<tr>
<td>Cognitive therapy</td>
<td>47.7</td>
<td>84.4</td>
</tr>
<tr>
<td>Pharmacotherapy</td>
<td>59.5</td>
<td>13.8</td>
</tr>
<tr>
<td>Both</td>
<td>78.8</td>
<td>71.7</td>
</tr>
</tbody>
</table>

SOURCE: Adapted from Blackburn et al. (1981).

et al. (1979a). The low percentage change for pharmacotherapy in general practice settings is noteworthy. One explanation for this result is that the general practice patients had subtypes of major depression that were less responsive to pharmacotherapy. Another explanation may be poorer compliance with medication in this group. These results emphasize the importance of controlling for differences in case-mix in comparisons of psychiatric and general practice settings, as well as the importance of assessing compliance with medication.

While it is difficult to draw conclusions from only a few studies, it appears that, among those who accept treatment, one can
conservatively expect about a 40 to 50 percent improvement in HRS or BDI scores as a result of antidepressant medication or certain specific forms of psychotherapy. A thorough review of the subject would be necessary to confirm this effect size in a larger number of studies or to reach conclusions on which aspects of research design affect the estimate of the magnitude of treatment effects when using particular outcome measures.

CLINICAL PREDICTORS OF RESPONSE TO TREATMENT

As noted above, not all patients with specific depressive disorders respond equally well to the various treatments for depression. We conclude this chapter on the treatment of depression by summarizing data on four prognostic indicators: (1) psychotic features, (2) anhedonia, (3) endogenous features (melancholia), and (4) anxiety. These four features may be of importance in controlling for case-mix in the MOS.

Several authors have evaluated the prognostic significance of delusions. Coryell and Tsuang (1982) reviewed studies that found a poor response to antidepressant medication when delusions were present. They cautioned that naturalistic studies of the prognostic significance of delusions are difficult to evaluate because patients with delusions are more likely to receive ECT, the most successful treatment for severe depression. Thus, a poorer prognosis could be masked by more effective treatment. They reviewed data on a 40-year follow-up of persons previously hospitalized with a diagnosis of “depression” and compared data on those with and without delusions. After the index hospitalization, those with delusions had significantly poorer response to treatment (on crude measures of occupational and social status and depressive symptoms). However, their 40-year follow-up study showed similar mortality and morbidity rates for both groups. They concluded that delusions indicate a poor short-term prognosis but not a poor long-term prognosis. Methodologic flaws in the design, such as incomplete knowledge of treatment history and different standards to evaluate the short- and long-term outcomes, preclude firm conclusions.

Coryell et al. (1984) recently compared outcomes during a 6-to-24-month follow-up for patients in the NIMH Collaborative Study of the Psychobiology of Depression. The subjects included 24 patients with schizoaffective disorder (depressed type), 56 with psychotic major depression, and 274 with nonpsychotic major depression (RDC criteria). The psychotic depressed patients had significantly poorer outcomes on measures of psychosocial functioning than did the
nonpsychotic depressed patients. However, survival curves were similar for the two groups.

Other studies suggest that both the short- and long-term prognoses are poor for patients with delusional depression. For example, Roose et al. (1983), using a retrospective case-control design, found that the prevalence of suicide in a psychiatric facility was five times higher for depressed patients with delusions than for those without. Brown et al. (1982) found a poorer response to tricyclic antidepressant medication in depressives with psychotic features.

Another feature of psychotic symptoms studied in relation to treatment response is mood congruence versus incongruence. Coryell, Tsuang, and McDaniel (1982) reported that mood-incongruence indicates a poorer response to treatment.

The presence of severe anhedonia (lack of interest or pleasure in usual activities) has been evaluated in relation to outcomes. Using an anhedonia scale, Fawcett et al. (1983) identified 12 percent of inpatients with major depression as having severe anhedonia. While this group had significantly greater depression by the BDI and more social impairment than those without severe anhedonia, they recovered more quickly. As a result, their outcomes were the same as those without anhedonia, with one exception: Those who initially had anhedonia remained more anhedonic at follow-up.

A third clinical feature of some prognostic significance is the presence of “endogenous” features, or melancholia (in DSM-III terminology). The “endogenous” symptom cluster or “melancholia” (DSM-III) seems to denote a distinct subtype of major depression and predicts a good response to tricyclic antidepressants (Paykel, 1972; Bielski and Friedel, 1976; Raskin and Crook, 1976; Rao and Coppen, 1979), good response to ECT (Avery and Lubrano, 1979), and a good short- and long-term prognosis (Kay et al., 1969; Paykel, Klerman and Prusoff, 1974).

A fourth clinical feature of prognostic importance is the presence of anxiety (in depressive disorders). Roth and Mountjoy (1982) have provided an excellent review of the nature of overlap of anxiety and depression. As noted by these authors, a subgroup of patients with depressive disorders also appears to have significant anxiety. Fawcett and Kravitz (1983) evaluated the prevalence of anxiety disorders and symptoms of anxiety in a sample of persons with major depression (RDC/SADS). Symptoms of anxiety were present in about 40 percent of the sample and 29 percent had a history of panic attacks. The sample size, however, was not reported.

The presence of symptoms of anxiety during an episode of major depression may have treatment implications. Among persons with
nonpsychotic unipolar disorders, those with significant anxiety respond poorly to tricyclic antidepressants, but respond favorably to MAO-inhibitors (Paykel, 1972; Paykel and Prusoff, 1977).
Chapter 5

THE TREATMENTS DELIVERED TO DEPRESSED PATIENTS

To what extent are procedures of known efficacy in the treatment of depressive disorders used in the practices of mental health specialists and general medical providers? Very few studies have addressed this issue using depressed patient groups as identified by diagnostic criteria (RDC, DSM-III). However, several past studies provided data on the ability of general medical providers to detect symptoms of depression and on the medications they prescribe for patients diagnosed as depressed. Relatively few studies have described the treatments that mental health specialists in private practice administer to depressed patients.

In reviewing studies of the treatment of mental disorders by specific provider groups, it is helpful to refer to a simple provider-oriented model of a plan for treatment of depression (Fig. 5.1). In this model, the provider first detects cases of depression and then performs a fuller assessment, including a determination of the presence of contributive organic factors and a fuller psychiatric assessment. This assessment leads both to a determination of the subtype of depression and to the identification of target symptoms, i.e., salient symptoms that can be used to monitor improvement. The provider selects a treatment regimen on the basis of the subtype of depression and other aspects of the clinical presentation. For example, the presence of acute suicidal intent suggests hospitalization and ECT. The provider then monitors the response to treatment. The parameters to be monitored include compliance with medication, side effects of treatment, response of target symptoms, and follow-up on referrals to other providers. When the depression has resolved, the patient's status is documented and a decision is made about the appropriateness of prophylactic medication (antidepressants or lithium).

This model of a treatment plan does not include algorithms for the appropriateness of decisions at each step. Rather it is intended to provide a conceptual framework to organize the empirical findings discussed below on the management of depression.

This chapter reviews studies of the treatments received by persons with depressive illness as defined by the RDC or DSM-III criteria regardless of the provider type visited, and the mental health care
Fig. 5.1—Treatment plan for depression
delivered by general medical providers and psychiatrists to patients with depression and to those with any mental disorders, as defined by a variety of criteria (research criteria, survey measures, physician diagnosis). We focus on psychiatrists and nonpsychiatrist physicians because both these provider groups can prescribe medication for depressive disorders.

**TREATMENT RECEIVED FOR DEPRESSIVE DISORDERS, ALL PROVIDERS**

Weissman and Myers (1980) and Weissman, Myers and Thompson (1981) described treatments received by persons in a community sample who had an RDC diagnosis of major or minor depression. During the year prior to the study, 17.1 percent had seen a mental health professional, 6 percent a psychiatrist, and 8.6 percent a nonpsychiatric physician about their depression. Of persons with depression, 55.2 percent had used psychotropic drugs; 34.5 percent used minor tranquilizers; 17.2 percent, sleeping pills; and 17.2 percent, antidepressants. It is not clear from the data how much of the drug treatment was prescribed by psychiatrists and how much by other physicians. Nevertheless, it is noteworthy that, among the depressed patients who received any psychotropic drugs, only 31 percent received the most efficacious drugs (antidepressants).

Keller et al. (1982b) studied the treatments received by 217 patients who were enrolled in the NIMH Collaborative Study of Depression with an RDC diagnosis of a current episode of definite major depressive disorder of at least one month's duration. They obtained self-reports (confirmed, when possible) on the treatments received prior to enrollment. Unlike the sample used by Weissman and her colleagues (1981), this sample was not drawn from a general population but was obtained through the normal referral mechanisms of the participating health care settings. Keller and his colleagues found that, prior to enrollment in the study, 67 percent of the sample had received psychotherapy; 55 percent had received antianxiety medication; 25 percent had received antipsychotic medication; and 34 percent had received antidepressant medication for at least four consecutive weeks. Only 12 percent had received antidepressants at daily doses equivalent to 150 mg or more of imipramine. While patients with psychotic depression had received more ECT and neuroleptics than had nonpsychotic patients, 25 percent of psychotic patients received neither neuroleptics nor antidepressants. Among those with a nonpsychotic major depression, even those with episodes of longer duration, greater severity at
enrollment (by the HRS), a history of recent suicide attempt, or the endogenous subtype were unlikely to have received high-intensity treatment. The data, however, do not allow a comparison of treatments delivered by psychiatrists, other mental health specialists, and nonpsychiatrist physicians.

TREATMENT OF DEPRESSION, GENERAL MEDICAL PROVIDERS

A few studies provide data on the detection of depression (Step 1, Fig. 5.1), the assessment of contributing organic factors (Step 2), the selection of a treatment regimen (Step 4), and the monitoring of treatment (Step 5) in general medical settings.

Studies of the detection of depression in general medical settings are of particular policy interest because of concerns that general medical providers may detect a low percentage of patients with mental disorders. These studies compared the diagnoses assigned to patients by their general medical providers (family practice residents, medical faculty members and senior medical students) with an independent assessment of depression, such as patients’ scores on a self-assessment measure of depression. Most of the studies were conducted in general medical outpatient clinics. In each study, the providers detected depression in less than half of the patients thought to have clinically significant depression.

Sellor, Blascovich, and Lenkei (1981) evaluated residents’ assessments of depression as indicated by the presence on the chart of a diagnosis of depression or prescription of antidepressant medication. Patients’ symptoms of depression were determined by the BDI. The 222 patients in this study represented a random sample of adult clinic patients stratified by age and sex. The sensitivity of the chart diagnosis was only 0.17, but the specificity was high (0.91). Female patients were more likely to be diagnosed, even controlling for their depression scores.

Nielsen and Williams (1980) compared evidence of a diagnosis of depression in the charts of 526 medical outpatients with the results of the BDI; 12.2 percent of the sample had at least mild depression, and 5.5 percent had moderate or severe depression. Among those with moderate depression, a diagnosis of depression was present in only one chart. However, other indicators that the physician was aware of some psychopathology was present in 50 percent of the charts of moderately depressed patients.
Physicians are significantly more likely to detect depressive symptoms if the patient is of high socioeconomic status or is female (Schwab et al., 1967; Schottstaedt, Deckert, and Schneider, 1971). The severity of illness is related to likelihood of detection. For example, Cavanaugh (1983) found that first-year medical residents detected 22 percent of patients with mild depression (BDI scores of 13 to 20), 29 percent of those with moderate depression (scores of 21 to 30), and 38 percent of those with severe depression (scores of 31 and over).

Katerndahl (1981) provided data on the assessment by family practitioners of organic factors that may contribute to depression (Fig. 5.1, Step 2). He reviewed the medical records of patients enrolled in a university family practice center. Of patients diagnosed as depressed during a one-year period, 47.8 percent also had a nonpsychiatric disorder that may cause depression, and 43.3 percent had been prescribed drugs that may cause depression. However, only 7.6 percent of the charts indicated that the physicians recognized the presence of these associated factors. It is difficult to draw conclusions from this study because it has not been replicated and because of the difficulty in drawing inferences on the physicians’ knowledge or behavior from medical records. However, the data suggest that general medical providers may not use or communicate important information relevant to evaluations of their depressed patients.

A few studies provided data on the specific mental health care delivered by general medical providers to the patients they diagnose as depressed (Fig. 5.1, Step 4). Johnson (1973) evaluated 73 patients with a new depressive episode from the practices of 14 general practitioners in England. The patients were identified by their physicians as having depression, and the diagnosis was subsequently confirmed by Johnson. To obtain information on treatment, the patients were interviewed at 4 to 6 weeks and 16 to 18 weeks after their diagnostic interview.

Of the 73 patients, 72 were prescribed an antidepressant drug, but only 25 percent were prescribed a daily dose equivalent to 75 mg or more of amitriptyline. Of 29 patients who received a prescription for an antidepressant after their second research interview, 17 (59 percent) had stopped taking it without their doctor’s knowledge. Reasons for stopping medication included feeling better (26 percent), prescription running out (21 percent), and skepticism of the value of pills (21 percent).

Only 3 percent received psychotherapy. No social agencies took part in the treatment of any of the patients, and modification of social problems (e.g., occupational, marital) was not part of treatment. Yet half of the patients reported being unable to follow their normal lives, and 81 percent indicated a recent stressful social event.
Moreover, the depressed patients seldom visited their physicians during the four-month period. For example, in the 4-to-6-week period between the first and second interview, 53 percent of the depressed patients had consulted their physician only once. While only 8 percent of patients expressed dissatisfaction with their physician, only 15 percent acknowledged receiving any personal help or support from their doctor in the depressive illness.

Johnson (1974) also provided data on the characteristics of the pre-existing relationship between general practitioners and the depressed patients. He used data reported in his 1973 paper and additional data from two samples of medical outpatients who had received a clinical diagnosis of depression. Half of all patients with depression had been registered with their doctor for five years or less, but one-fourth had been registered for ten years or more. Only one-third regarded themselves as well-known to their family doctor; two-fifths said they were hardly known or totally unknown to the family doctor. Those who continued in treatment for their current depressive illness were significantly more likely to have been well-known to their doctor prior to the onset of illness.

This study also provided data on treatment patterns in private practice. Although 92 percent of patients received antidepressants at their initial consultation, this percentage dropped to 50 percent at subsequent consultations. In all groups, physicians prescribed low doses of antidepressants. Further, physicians tended to reduce the dose at subsequent consultations. Two-fifths of patients who continued in treatment for more than three months were no longer seeing their general practitioners and received their repeat prescriptions without an interview. Drug defaulting was a major problem in all patient groups. For example, after one month, 65 percent of new patients had ceased to take their medication regularly. Lack of communication between patient and doctor was one of the reasons given for defaulting. Finally, depending on the patient sample, only 10 to 24 percent were ever referred to a psychiatrist.

A few studies conducted in the United States provided evidence on the tendency of general medical providers to prescribe psychotropic drugs inappropriately for depressed patients (Gullick and King, 1979; Tollefson et al., 1983). However, the patient samples in these studies were quite small. Gullick and King (1979) used a structured psychiatric interview to determine the presence of psychiatric disorders in 139 consecutive patients who presented at a marital and sexual counseling center. Reports of psychoactive drug use were confirmed by urine assays for 100 patients. For the 20 patients with a current unipolar depression, the most commonly used drug was valium. Only four
patients were on antidepressants. Two-thirds of the patients had received their drugs from a primary care physician and one-fifth from a psychiatrist. Tollefson et al. (1983) applied DSM-III criteria to 30 consecutive patients diagnosed by family practice residents as having depression. Thirteen had a major depressive illness. Of these, five had been prescribed an antidepressant.

THE TREATMENT OF ALL MENTAL DISORDERS, GENERAL MEDICAL PROVIDERS

A substantial body of literature has included data on the detection and management of all mental or emotional disorders by general medical providers. Because major depression is the most prevalent mental disorder in these settings, the results of these studies contribute to an understanding of how depressive disorders are managed. Generally, these studies support two conclusions of the previous section: (1) general medical providers may detect a relatively low percentage of cases of mental disorder and (2) management consists largely of psychotropic medication. These studies provide additional information on factors associated with detection and on referral practices.

Goldberg (1980) summarized data from two studies (in England and the United States) on the detection of mental disorders. In both studies, physician reports of mental illness were compared with the patients' scores on self-reports of general mental health (Cornell Medical Index, GHQ). For both studies, the correlation between physician assessments and true cases of mental disorder (as defined by the patient survey measure) was very low. While the "true prevalence" was 20 to 50 percent across individual practices, physicians' estimates varied from 0 to 80 percent. Physicians with high estimates were as inaccurate as physicians with low estimates.

Kessler, Cleary, and Burke (1985) reported data on the detection of mental disorders, as defined by SADS-L criteria, by primary care physicians. Detection was identified by diagnoses, treatments, and referrals listed in the medical record. About 30 percent of those with a mental disorder that persisted over six months had any evidence of "detection" in the medical record during that period.

Shepherd et al. (1966) and Marks, Goldberg, and Hillier (1979) in England, and Goldberg et al. (1982) in the United States, studied the determinants of the variation in physician's estimates of the prevalence of mental disorders in their practices. Shepherd and his colleagues found that a measure of patient turnover ("mobility index") and a measure of the importance to the physician of psychogenic factors in
disease ("psychosomatic score") accounted for 51 percent of the variance in level of detection. Marks, Goldberg, and Hillier (1979) accounted for 53 percent of the variance by a measure of "general interest and concern" and the physicians' orientation toward psychiatry. Marks and his colleagues also studied determinants of the accuracy of detection; 59 percent of the accuracy of physicians (compared with the GHQ) in detecting mental illness was accounted for by "interest and concern" and "conservatism." Conservatism was a measure of personal inflexibility and authoritarianism.

Goldberg et al. (1982) studied the accuracy of family practice residents in detecting mental disorders as defined by a cut-off score on the GHQ. Interviews with patients were videotaped and rated by independent observers. Residents who were more accurate in detecting patients with mental illness tended to be more self-confident and extroverted, used more eye contact with patients, tended to begin clinical interviews with open-ended questions, and became more directive as interviews progressed. Interview behaviors explained 30 percent of the variance in accuracy. The correlation between accuracy and level of detection (i.e., percentage of patients thought to have mental illness) was insignificant. Those who diagnosed more mental illness (regardless of accuracy) were more "psychologically oriented," i.e., they asked more psychosocial questions, were more empathic, and were more sensitive to verbal cues in interviews.

Hankin (1980) reviewed the literature on the management in general medical settings of patients diagnosed as having any mental disorder or emotional problem. The majority of patients with emotional disorders received pharmacotherapy, and the percentage of patients treated with psychotropic medication varied from 29 to 79 percent, depending on the study. For example, Fink et al. (1967) found that 57 percent of patients with a diagnosis of mental disorder were treated by drugs alone. Rosen et al. (1972) found that physicians in five general hospital clinics used psychotropic drugs alone for 15 percent; drugs plus therapy for 35 percent; and therapy, drugs, and suggested environmental change for 10 percent of patients with mental disorders.

Data from NCHS indicate that only 2 percent of visits to primary care practitioners include a psychotherapy procedure (Brown, Regier, and Balter, 1979). However, this 2 percent represents 27 percent of all psychotherapy sessions, due to the large volume of medical care visits relative to visits to specialty mental health providers. The percentage of patients with mental diagnoses in primary care settings who receive any psychotherapy has been estimated at 60 to 84 percent, but psychotherapy is received in less than a quarter of visits from these patients (Brown, Regier, and Balter, 1979). Little is known about the nature or
quality of psychotherapy given or the outcomes of psychotherapies. A few studies suggest that such treatment is short-term (one to four sessions) (Fink et al., 1967; Wells et al., 1982). Aldrich (1965) studied general practitioners in the United States and Great Britain and found they were less likely to use face-to-face counseling as a management strategy than were psychiatrists. Nevertheless, several studies have suggested that the majority of patients receiving psychotherapy from medical physicians are satisfied with their care (Locke, 1966; Johnson, 1974).

The referral of patients to psychiatrists by general physicians has been studied in the United States and Great Britain. About 1 to 2.5 percent of all patients in medical outpatient settings are referred to psychiatrists, and the percentage of patients with a mental diagnosis referred varies from 5 percent (Shepherd et al., 1966) to 44.5 percent (Hilkevitch, 1965). Referral rates tend to be lower in Great Britain, probably due to clearer separation of medical and psychiatric services and different views of the general physician’s role in providing psychological services (Hankin and Oktay, 1979). Some characteristics of patients associated with increased likelihood of being referred to a psychiatrist (given that a psychiatric diagnosis is present) have been identified. These include being male and being younger (Hopkins and Cooper, 1969; Shepherd et al., 1966). Patient request for referral also makes a referral more likely as does the presence of psychotic symptoms (Fink et al., 1967).

Studies of the effect of physicians’ attitudes toward psychiatry and mental disorders on use of psychiatric referral services present conflicting results. Shortell and Daniel (1974) found that physicians with positive attitudes toward psychiatry are more likely to refer patients to psychiatrists. Hull (1979) found that, when other factors are controlled for, physicians with greater knowledge about psychotropic drugs and those who see the management of mental disorders as within the role of the general physician are less likely to refer to psychiatrists (by self-reports of referral practices). In addition, Hull found that physicians from high social class backgrounds, Anglos, non-Catholics, and those with more nonpaying patients were less likely to refer to psychiatrists.

Management of Depression, Psychiatrists

Studies reporting on the treatments provided by psychiatrists include quality assurance studies, surveys of practicing psychiatrists, studies of utilization in specific insurance plans or treatment settings, and analyses of data from the National Ambulatory Medical Care
Survey (NAMCS) or the Rand Health Insurance Study (HIS). These studies provide data on the medications and psychotherapies selected. Few data are available on other components of a treatment plan (Fig. 5.1). However, one study provided data on therapists' awareness of specific dimensions of psychopathology (including depression) in their patients (Kass et al., 1980).

Kass and his colleagues (1980) interviewed 32 patients who visited the Bronx Municipal Hospital Center Emergency room (for any reason) and who were in ongoing psychotherapy (at least once a week) at the hospital outpatient clinic. They obtained scores on the HSCL for each patient and immediately telephoned the therapist to obtain ratings on the nine dimensions of psychopathology reflected in the scale. They then mailed the therapists a survey asking them to rate the patients on the items of the HSCL. The therapists were aware of depression in 94 percent of patients who scored above the cut-off for the depression subscale. Therapists were less accurate in recognizing psychotic (35 percent) and obsessive-compulsive (16 percent) pathology.

Awareness of psychopathology did not vary with demographic characteristics of patients. Several clinical and treatment factors were related to awareness: Patients with a diagnosis of borderline personality, who were fearful of offending the therapist or who felt their therapists were not empathic, were less likely to have a therapist who was aware of their psychopathology. The small sample size and unusual method of sampling precludes firm conclusions from this study. However, the results suggest that mental health specialists may be more aware of symptoms of depression in their patients than are general medical providers.

Studies that provide data on the treatments selected by psychiatrists for their depressed patients have focused on the appropriateness of medications prescribed and the type of psychotherapy administered. For example, Langley and Lebaron (1974) surveyed members of a local psychiatric society about the treatment delivered to all patients seen over a six-month period; 97 members provided reports on 3,151 patients, of whom 378 had a diagnosis of depression. For depressed patients treated in the hospital, the length of stay ranged from 1 to 99 days, with a median of 6 days. For depressed patients treated on an outpatient basis, the number of office visits per patient per year varied from 1 to 390, with a median of 14; 69 percent were treated with psychotropic medication, and 2.4 percent were treated with ECT.

Most studies of psychiatrists' prescribing practices were designed to assess the quality of hospital inpatient care. These studies concluded that psychiatrists may make excessive use of antipsychotics, e.g., may use too many drugs for the same patient, and may use anti-
parkinsonian drugs unnecessarily (Altman et al., 1972; Schroeder, Caffey and Lorei, 1977; Hartmann, Allison, and Hartig, 1979; Winstead et al., 1976; Michel and Kolakowska, 1981; Eastaugh, 1980). However, Prien, Balter, and Caffey (1978) discussed the need for caution in inferring inappropriate use of psychotropic medication from hospital survey data. For example, in many of these studies, data were not available to ascertain fully the appropriateness of medications prescribed for a particular patient. In fact, some of the studies that emphasized the inappropriateness of psychiatrists' prescribing habits also provided data suggesting that psychiatrists may use antidepressants more appropriately than do general medical providers.

Michel and Kolakowska (1981) examined the care delivered to 511 inpatients in two British psychiatric hospitals. They concluded that psychiatrists tended to prescribe too many psychoactive drugs for each patient and tended to prescribe antidepressants for some patients who did not have a diagnosis of affective disorder. Nevertheless, in this study, every patient with a diagnosis of depression was prescribed lithium, an antidepressant, or a neuroleptic. Further, when antidepressants were prescribed, drugs of known efficacy were most often selected (i.e., amitriptyline, imipramine). Only one-third of patients who were prescribed antidepressants received daily doses lower than the equivalent of 100 mg of imipramine. Those with lower dosages were either elderly or had a diagnosis other than an affective disorder. Lower dosages may have been appropriate for these patients.

Studies of psychiatrists' choice of psychotherapy indicate that individual (face-to-face) psychotherapy is the predominant treatment modality in outpatient settings. Marmor (1975) reported the results of surveying a 10 percent random sample of psychiatrists who spent at least 15 hours a week in private practice. The psychiatrists provided data on the last ten patients seen. These 440 respondents provided individual psychotherapy for 81 percent of their patients, group therapy for 10 percent, conjoint therapy for 5 percent, and family therapy for 4 percent. Psychoanalysts were slightly more likely to employ individual therapy than were child psychiatrists and general psychiatrists.

In this study, patients of nonanalysts who were seen in individual therapy had from one to twelve sessions or more per month; three-fourths were seen fewer than nine times a month. Unfortunately, Marmor did not provide data on the psychotropic medication used nor the use of psychotherapy for specific diagnostic groups. The finding that psychiatrists largely provide individual psychotherapy is consistent with the results of utilization studies. For example, Reed, Myers, and Scheidemandel (1972) reported that in 1970, Michigan Blue Cross and
Blue Shield paid $3,459,366 to private practitioners for mental health care. Of this total, $3,327,363 was to cover the cost of individual psychotherapy.

Wells et al. (1982) reported a similar finding using data from the Rand HIS. In this study, a random sample of the nonaged, noninstitutionalized civilian population in six U.S. sites was randomly assigned to health insurance plans that varied in the amount of cost-sharing for all health services. The enrollee's use of services and their health status were followed for three- or five-year enrollment periods. In the second year, 7.6 to 9.2 percent of enrollees used outpatient mental health services. Roughly half of these users ever saw a mental health specialist for their care, and about a quarter ever saw a psychiatrist. Individual psychotherapy was received by 88 percent of users of mental health specialists. By contrast, only 6 percent of those who used general medical providers for their mental health care received any specific psychological services or therapeutic listening (as indicated by claims forms).

Two recent studies provided more specific information on the treatments and dispositions used by psychiatrists and nonpsychiatrists for patients with any mental disorder, using National Ambulatory Medical Care Surveys (NAMCS). In these studies, physicians kept special patient records for one week.

Verbrugge (1984) found that 5.4 percent of all adult visits in 1975 had a primary diagnosis of a mental disorder. Table 5.1 summarizes the services given and the dispositions of these visits. Roughly 25 percent of visits to psychiatrists (as compared with 54 to 60 percent of clients to nonpsychiatrists) included a drug prescription. Data on prescription of psychotropic drugs are not given in the paper. Over 90 percent of visits to psychiatrists included psychotherapy or therapeutic listening, compared with roughly 15 percent of visits to nonpsychiatrists. By contrast, nonpsychiatrists performed physical examinations and procedures during a much higher proportion of these visits.

Verbrugge found differences in the kind of treatments performed by nonpsychiatrists, depending on the sex of the patient. The sex differences were most marked for patients with a diagnosis of anxiety or depression. For example, women with a diagnosis of anxiety were more likely to be examined, have their blood pressure taken, and receive a drug prescription than were men. These women were more likely to be given a return visit, while men were more likely to be given medical advice, therapeutic listening, or a referral elsewhere. Sex differences were not found for visits to psychiatrists.

The Bureau of Health Professions (1984) provided data on the characteristics of mental health visits delivered by psychiatrists and nonpsychiatrist physicians, using data from the 1980 and 1981 NAMC
Table 5.1
SERVICES GIVEN AND DISPOSITIONS OF VISITS FOR PATIENTS WITH A PRIMARY MENTAL HEALTH DIAGNOSIS

<table>
<thead>
<tr>
<th>Services and Dispositions</th>
<th>Type of Provider</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Psychiatrists</td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
<td>Men</td>
</tr>
<tr>
<td>No services</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>Limited history/examination</td>
<td>2.9</td>
<td>2.8</td>
<td>52.6</td>
<td>50.2</td>
<td></td>
</tr>
<tr>
<td>General history/examination</td>
<td>0.9</td>
<td>2.1</td>
<td>15.5</td>
<td>19.0</td>
<td></td>
</tr>
<tr>
<td>Lab test</td>
<td>0.3</td>
<td>0.6</td>
<td>16.3</td>
<td>20.0</td>
<td></td>
</tr>
<tr>
<td>Blood pressure check</td>
<td>2.1</td>
<td>1.2</td>
<td>52.2</td>
<td>42.9</td>
<td></td>
</tr>
<tr>
<td>EKG</td>
<td>0.1</td>
<td>0.0</td>
<td>3.9</td>
<td>7.1</td>
<td></td>
</tr>
<tr>
<td>Hearing test</td>
<td>0.0</td>
<td>0.0</td>
<td>2.4</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>Vision test</td>
<td>0.0</td>
<td>0.0</td>
<td>2.9</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Endoscopy</td>
<td>0.0</td>
<td>0.0</td>
<td>0.8</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>X-ray</td>
<td>0.3</td>
<td>0.2</td>
<td>5.7</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>Office surgery</td>
<td>0.1</td>
<td>0.0</td>
<td>0.1</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>Drug prescription</td>
<td>25.3</td>
<td>24.6</td>
<td>58.9</td>
<td>53.8</td>
<td></td>
</tr>
<tr>
<td>Injection</td>
<td>1.1</td>
<td>1.1</td>
<td>20.5</td>
<td>20.3</td>
<td></td>
</tr>
<tr>
<td>Immunization/desensitization</td>
<td>0.1</td>
<td>0.0</td>
<td>1.1</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>Physiotherapy</td>
<td>0.3</td>
<td>0.4</td>
<td>2.9</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>Medical counseling</td>
<td>5.8</td>
<td>6.9</td>
<td>21.8</td>
<td>20.4</td>
<td></td>
</tr>
<tr>
<td>Psychotherapy/therapeutic listening</td>
<td>93.2</td>
<td>91.0</td>
<td>24.7</td>
<td>15.6</td>
<td></td>
</tr>
<tr>
<td>Other service</td>
<td>8.7</td>
<td>13.1</td>
<td>4.9</td>
<td>7.0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dispositions (% of visits)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No follow-up plans</td>
<td>2.4</td>
<td>3.3</td>
<td>8.8</td>
<td>10.0</td>
<td></td>
</tr>
<tr>
<td>Appointment</td>
<td>92.1</td>
<td>91.1</td>
<td>52.2</td>
<td>47.9</td>
<td></td>
</tr>
<tr>
<td>PRN</td>
<td>8.1</td>
<td>5.9</td>
<td>34.5</td>
<td>34.5</td>
<td></td>
</tr>
<tr>
<td>Telephone</td>
<td>4.3</td>
<td>2.9</td>
<td>4.6</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td>Referral</td>
<td>1.1</td>
<td>1.0</td>
<td>4.6</td>
<td>6.8</td>
<td></td>
</tr>
<tr>
<td>Return to referring physician</td>
<td>0.1</td>
<td>0.4</td>
<td>0.5</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>Admit to hospital</td>
<td>0.8</td>
<td>0.8</td>
<td>0.7</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>Other disposition</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
<td>1.5</td>
<td></td>
</tr>
</tbody>
</table>


surveys. Among the findings of the bureau: (1) Nonpsychiatrists perform many more diagnostic services, such as physical exams, during their mental health visits; (2) nonpsychiatrists prescribe psychotropic drugs in a greater proportion of their office visits that have a psychiatric diagnosis than do psychiatrists; (3) nonpsychiatrists are more likely to prescribe anxiolytic drugs during a mental health visit; and (4) mental health visits delivered by nonpsychiatrist physicians were, on the average, about 20 minutes in duration, compared with 44 minutes for visits to psychiatrists. Because the report does not analyze differences
in care provided to persons (including all their mental health visits), it is not possible to fully evaluate the significance of differences in style of care. For example, it is not clear that a person visiting one provider type rather than the other would be more or less likely to ever receive a psychotropic medication. Nevertheless, these studies strongly point to marked differences in the style of mental health care delivered by psychiatrists and nonpsychiatrist physicians.

Mental Health Treatment in Health Maintenance Organizations (HMOs) and Fee-for-Service (FFS) Settings

The care delivered to depressed patients in HMOs may differ considerably from that in FFS settings. The economic incentives provided in HMOs may affect the process of care in four ways (Luft, 1981). First, while FFS providers gain financially from increased use of their services, HMO providers do not. Typically, income of HMO providers does not depend directly on the number of services they provide. Thus, they may attempt to control use by limiting the frequency and/or the intensity of care (Bilker and Idzonek, 1978). Second, general medical providers in an HMO may have a greater incentive than FFS physicians to refer to mental health specialists without fear of losing their patients. Third, HMO patients typically have less control over the specific type of treatment they receive. As a result of these factors, patients in an HMO may be more likely to receive time-limited (crisis-oriented) or group therapy. Fourth, HMOs may rely heavily on nonphysician personnel to provide mental health services. The use of psychiatrists may be reserved for sicker patients or be limited to the prescription of medications or supervision of nonphysicians. As a result, there may be major differences in psychiatrists’ case-mix or in the role of psychiatrists in FFS and HMO settings.

Studies comparing patterns of use in HMOs and FFS settings support the hypothesis that differences in incentive structure lead to differences in the care delivered. Those studies have found reduced inpatient and increased outpatient use in HMOs. (See Luft, 1981, for a review of the medical literature.) For example, Craig and Patterson (1981) performed a secondary analysis of data from four studies of prepaid group plans and compared these data with data on use in the high-option and low-option Blue Cross/Blue Shield Federal Employee Plans (FEPs). Expenditures for inpatient services under HMO plans were less than those under the low-option FEPs. Expenditures for out-
patient services in prepaid plans equaled or exceeded expenditures in the high-option FEPs.

Billings and Moos recently reported data on the treatments delivered to 412 depressed patients in psychiatric treatment facilities, outpatient programs of an HMO, a CMHC, and a VA, and inpatient programs of a university teaching hospital and a VA. All patients were adults with major or minor unipolar depression according to the RDC. Even after controlling for differences among programs in patient characteristics (sociodemographic factors, severity of depression), there were significant differences among programs in the treatments delivered. The HMO outpatients made fewer mental health visits and were less likely to receive psychoactive or antidepressant medication than outpatients in the VA or CMHC (Billings and Moos, 1984).

Studies comparing utilization in HMO and fee-for-service plans have not focused on the use of mental health services furnished by general medical providers. Such services would be especially interesting to study because HMOs vary considerably in the degree to which general medical providers are encouraged to provide mental health services. For example, some HMOs specifically encourage their family physicians to provide mental health care.

SUMMARY AND IMPLICATIONS FOR THE MOS

The treatment of depression can be conceptualized as a sequence of clinical decisions involving the detection of depression, differential diagnosis and identification of target symptoms, selection and monitoring of treatment, and evaluation of outcomes and the need for prophylaxis. While the practices of mental health specialists and general medical providers may differ at several steps in this sequence, the available data largely concern the detection of symptoms or "cases" and the selection of a treatment modality. In particular, general medical providers may detect depression in a small proportion of their patients who have a depressive disorder or depressive symptoms. However, those studies have largely focused on medical outpatient clinic settings and physicians in training. More experienced physicians or those in private offices may have a different level of detection. While no studies exist on the detection of major depression by psychiatrists or psychologists, at least one study suggests that psychiatrists are sensitive to their patients' depressive symptoms.

Concern has been expressed about the appropriateness of the practices of both psychiatrists and general medical providers in prescribing psychotropic medication. However, the available evidence suggests
that psychiatrists are more likely than are general medical providers to select an appropriate antidepressant and to prescribe therapeutic doses of antidepressants. Further, while psychiatrists tend to administer individual psychotherapy to most of their patients, general medical providers provide their mental health care in the context of brief office visits that include physical health evaluations and procedures.

These findings strongly suggest that there are major differences in the mental health care delivered by the two provider groups. Moreover, the mental health care delivered in HMOs may differ in important respects from that delivered in fee-for-service settings. Thus, the pattern of care for depression varies considerably among the provider groups and health care settings of interest for the MOS.

Moreover, the variation in care observed among provider groups is very likely to be associated with major differences in the outcomes of care. For example, antidepressant medication is a very effective treatment for major depression. However, general medical providers may be more likely than psychiatrists to prescribe inadequate doses. Further, if nonpsychiatric physicians have a low rate of detection of depressive disorders, their depressed patients may be less likely to receive antidepressant medications. Empirical data on this issue are not available, however.

Mental health specialists and general medical providers appear to differ considerably in their manner of counseling patients with mental disorders, but the implications of this difference are uncertain. As discussed in Chap. 7, a number of investigators have suggested that all effective psychotherapies share a common set of beneficial factors (Rounsaville, Weissman, and Prusoff, 1981; Garfield and Bergin, 1978; Frank, 1974). Such factors include the nature of the therapeutic alliance and the physician’s ability to provide general emotional support. Because general medical providers may accomplish some of these objectives even in the brief therapy they provide, their minimal therapy may be quite effective. On the other hand, Weissman et al. (1979) found that a specific type of psychotherapy was more efficacious than unscheduled supportive therapy. If talking with a general medical provider is equivalent to unscheduled supportive therapy, then one would expect the more rigorous psychotherapy provided by mental health specialists to be associated with better clinical outcomes.
Chapter 6

A MODEL OF OUTCOMES OF DEPRESSION

In the MOS, the term “outcome” refers to dimensions of patients’ health status or health-related attitudes and behaviors that may be affected by the level or quality of medical care received. These “outcomes” are measured at exit from the study and at specified intervals during the study (e.g., every six months). “Outcomes” may be affected by characteristics of the target disorder (e.g., a chronic versus an acute disorder), the individual (e.g., concurrent disorders), the environment (e.g., stressful life events), and the care received (e.g., medication, psychotherapy).

In the MOS, observed differences between groups of patients in “outcomes” are not necessarily attributable to differences in medical care received. By contrast, differences in outcome between groups in a randomized controlled trial can be attributed to the experimental intervention. In the MOS, the ability to attribute differences in outcome to one factor instead of others depends on our ability to identify outcomes that are likely to be affected by medical care and other factors that can affect those outcomes. This chapter integrates conclusions of previous chapters to propose a model of the “outcomes” for the major depression tracer. Chapter 10 concerns the identification of factors other than medical care that affect these outcomes.

The major outcome dimensions for depressive disorders are exit diagnosis, level (severity) of symptoms, mortality, course of illness, social impairment, and satisfaction with care (Table 6.1). This chapter provides references to support conclusions only for outcomes not discussed in previous chapters. The importance of social impairment as an outcome is discussed in detail at the end of Chap. 8. The importance of patient satisfaction with care is not discussed in depth because other MOS investigators are examining this concept for use in the MOS as a whole.

Few past studies of depression have assessed diagnostic classification at the end of the study, but it is important to do so for several reasons. First, diagnostic reassessment can clarify whether the index episode has resolved or whether the patient is experiencing a new depressive episode. Second, it can document the development of new psychiatric disorders. For example, dysthymic disorder may result from
Table 6.1
OUTCOME DIMENSIONS FOR DEPRESSIVE DISORDERS

<table>
<thead>
<tr>
<th>Diagnostic Classification</th>
<th>Course of Episode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major depression present?</td>
<td>Length of episode</td>
</tr>
<tr>
<td>Dysthymia present?</td>
<td>Recurrences</td>
</tr>
<tr>
<td>Evidence of bipolar disease</td>
<td>Symptom-free periods</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severity of Symptoms</td>
</tr>
<tr>
<td>Depressed mood</td>
</tr>
<tr>
<td>Somatic symptoms</td>
</tr>
<tr>
<td>Melancholia</td>
</tr>
<tr>
<td>Hypochondriasis</td>
</tr>
<tr>
<td>Anxiety</td>
</tr>
<tr>
<td>Psychotic Symptoms</td>
</tr>
<tr>
<td>Suicidal ideation</td>
</tr>
</tbody>
</table>

| Satisfaction with Care |

Inadequately treated major depression. This is particularly important because a patient with a residual chronic depression is prone to relapse and may require continued antidepressant medication. Third, the reassessment may identify people who developed manic symptoms during the course of the study and thus should be classified as having a bipolar disorder. Such people may have to be excluded from the analysis because bipolar disease responds differently to treatment than does unipolar disease, and has a different course of illness.

Depressive symptoms are important outcomes for several reasons. First, they cause patients considerable distress. Second, different treatments may have differential effects on symptoms. For example, psychotherapy may have its strongest effects on mood and social adjustment; pharmacotherapy may have a greater effect on associated somatic symptoms (e.g., weight loss, sleep disturbance). Third, some symptoms are believed to be especially disabling (e.g., psychotic symptoms) or of prognostic importance (e.g., suicidal ideation).

Some characteristics of major depressive episodes may also be considered meaningful treatment outcomes. For example, adequate treatment may shorten episodes, allow more symptom-free periods, or reduce the risk of recurrences. Because recurrences and larger episodes are likely to result in more functional impairment, they have important consequences for the patient’s quality of life.

About 15 percent of patients with depressive disorders commit suicide, but in spite of that frequency, it is difficult to obtain reliable
information on suicide. Physicians may be reluctant to record suicide as the cause of death on a death certificate, and some suicides may be masked as accidents (e.g., single-car accidents).

Numerous studies have found a positive association between depression (i.e., unipolar depressive disorder or depressive symptoms) and poor social resources and functioning (Flaherty et al., 1983; Billings, Cronkite, and Moos, 1983; Brown, Bhrocham, and Harris, 1975; Clayton, Halikes, and Maurice, 1974; Warheit, 1979; Costello, 1982; Brown and Harris, 1978). Some researchers have hypothesized that depressed persons alienate others and that they lack the social skills necessary for developing good relationships (Lewinsohn, 1974; Coyne, 1976). Others propose that the lack of social supports worsens the negative impact of stressful life events, which in turn lead to a greater risk for depression (Brown and Harris, 1978). Finally, the lack of social supports may itself predispose to depression (Warheit, 1979; Chan, 1977). Few studies have used prospective designs that would help clarify the extent to which poor social supports should be considered a cause or consequence of depression.

From a clinical perspective, social impairment is usually considered the most important nonfatal consequence of a major depressive episode (Klerman, 1980a). Major depression commonly results in sexual difficulties, which can intensify personal distress and marital and family discord. In addition to the personal suffering that attends social impairment, there are major policy implications of the disability or reduced occupational functioning associated with depression. Further, social impairment has been shown to improve with effective treatment of depression. As a result, persistent impairment (at study end-point) may indicate inadequate treatment.

The outcome model also includes patient satisfaction with care. Satisfaction with psychiatric care is modestly correlated with improvement in symptoms (Edwards et al., 1978; Willer and Miller, 1978). However, studies of the mental health care delivered by general medical providers suggest that satisfaction with care may be largely independent of the adequacy of care received (Johnson 1973, 1974). Satisfaction with care is itself an important treatment outcome because major depression is a disorder where continuity of care is important. However, patients who are dissatisfied with their care may be more likely to change doctors (DiMatteo and Friedman, 1982).
Chapter 7

THE PROCESS OF CARE FOR DEPRESSION

Chapter 4 discussed the treatment of major depression in terms of a treatment plan or a logical progression of clinical decisions over time. The concept of the process of care as used in the MOS refers to all components of care present in a particular encounter between a patient and a provider, whether or not that care reflects a logical treatment plan. An important objective of the MOS is to describe the process of care delivered by several physician groups in prepaid and fee-for-service settings.

The process of care for depression can be evaluated from several approaches. For example, previous studies of the process of care have examined the selection of a specific treatment modality (e.g., psychotropic drugs or counseling); the structural aspects of care (e.g., treatment setting, frequency or duration of office visits); the content of office visits (e.g., medical history, physical examination, explanations of treatment); and the nature of the physician/patient relationship. Within these general categories, the process of care can be evaluated at different levels of detail. For example, evaluations of drug treatment can focus on the selection of a specific drug, the appropriateness of the prescribed regimen, or the quality of instructions given by the physician.

In discussing the literature on the process of care for depression, we refer to a model of the process of care (Table 7.1). The model represents a synthesis of the literature on counseling and psychotherapy. Its dimensions consist of the diagnostic evaluation, the treatment plan, the general intensity of care, specific psychotherapeutic activities, and the general affective quality of the physician-patient relationship. The diagnostic evaluation and treatment plan dimensions are similar to the treatment plan discussed in Chap. 4. “General intensity of care” describes the amount of care in terms of duration, number, and frequency of sessions. In addition, the initiation of treatment is an indication of the intensity of the provider’s approach. These constructs were empirically shown to represent a single dimension of counseling (for health habits) in Wells, Ware and Lewis (1984).

The dimensions for specific psychotherapeutic activities and the general affective quality of the physician/patient relationship describe
### Table 7.1
**DIMENSIONS OF THE PROCESS OF CARE FOR DEPRESSION**

<table>
<thead>
<tr>
<th>Diagnostic evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychiatric history</td>
</tr>
<tr>
<td>Evaluation of depressive symptoms, functional status</td>
</tr>
<tr>
<td>Physical examination, procedures</td>
</tr>
<tr>
<td>Evaluation of organic contributing factors</td>
</tr>
<tr>
<td>Treatment plan</td>
</tr>
<tr>
<td>Psychotropic drug; appropriate dose</td>
</tr>
<tr>
<td>Dosage</td>
</tr>
<tr>
<td>Type counseling or psychotherapy</td>
</tr>
<tr>
<td>Follow-up</td>
</tr>
<tr>
<td>General intensity of care</td>
</tr>
<tr>
<td>Duration of care</td>
</tr>
<tr>
<td>Frequency of visits</td>
</tr>
<tr>
<td>Duration of sessions</td>
</tr>
<tr>
<td>Provider- versus patient-initiated care</td>
</tr>
<tr>
<td>Specific psychotherapeutic activities</td>
</tr>
<tr>
<td>General activity versus passivity</td>
</tr>
<tr>
<td>Sharing treatment plan</td>
</tr>
<tr>
<td>Giving advice or specific suggestions</td>
</tr>
<tr>
<td>Interpretations (accuracy)</td>
</tr>
<tr>
<td>General affective quality of physician/patient relationship</td>
</tr>
<tr>
<td>Physician's warmth</td>
</tr>
<tr>
<td>Positive regard for patient</td>
</tr>
<tr>
<td>Accuracy of understanding (empathy)</td>
</tr>
</tbody>
</table>

The particular therapeutic style of the psychotherapeutic relationship. These are the two aspects of process most commonly evaluated in studies of the process of psychotherapy.

Of the five dimensions, those of the treatment plan and types of psychotherapeutic activities have been extensively studied in relation to outcomes for major depressive disorders. Recent studies of the efficacy of specific psychotherapies (e.g., cognitive therapy) in the treatment of depression may be considered studies of the effects of both treatment plan and specific psychotherapeutic activities on outcome.

One study of the treatment of major depression has examined the relation of specific psychotherapeutic activities to outcomes (Rounsaville, Weissman, and Prusoff, 1981). This study was part of the controlled trial of ITP described above (Weissman et al., 1979; DiMascio et al., 1979). In that study, psychotherapists recorded aspects of the
process of each session immediately after the session. The aspects of process studied included the problems addressed in the visit and the extent of use of eight therapeutic techniques: non-judgmental exploration; active questioning; clarification; direct advice; assistance in decision-making; insight development; facilitation of emotional responses; and other techniques. The investigators also noted the total number of sessions provided.

The outcome measures included the Social Adjustment Scale Self-Report, and depressive symptoms, as indicated by scores on the HRS (derived from a clinical interview). Three components of process were significantly related to improved Hamilton depression scores in bivariate analyses: less use of explorative (listening) techniques, less discussion of major problem areas, and more use of assistance in decision-making. However, when pre-treatment characteristics of the patient were statistically controlled for, there were no significant associations between these components of process and outcome. The pre-treatment characteristics, measured by clinicians’ judgments after the initial session, included the patient’s aptitude for psychotherapy, emotional freedom, acute depression, general emotional health, and intellectual achievement.

As noted by Rounsaville and his colleagues, the patients receiving specific treatment for depression improved in relation to a control group, indicating a link between outcome and process at a gross level. The authors indicated that other aspects of process previously linked to outcomes were omitted from their study. These authors noted that several large-scale studies of the process of care for depressive disorders are in progress as part of the NIMH Collaborative Studies on Depression.

There is a large literature on the relation of process and outcome for psychotherapy in general. While these studies did not focus specifically on depressed patients, they are nevertheless an important source of information in support of a model for process. This body of literature includes several large-scale studies, such as the Vanderbilt Psychotherapy project (Strupp and Hadley, 1979), the Penn Psychotherapy Project (Luborsky et al., 1980), the Menninger Foundation Psychotherapy Research Project (Kernberg et al., 1972), the Wisconsin Project (Rogers et al., 1967), and others. These studies evaluated process largely in terms of psychotherapeutic activities and physician/patient relationships. They tended to use detailed evaluations of process, including observer’s scores derived from session transcripts and clinician and patient assessments. Several reviews of this literature have recently appeared (see Garfield and Bergin, 1978). These reviews identify components of the process of care that might be useful for the
MOS. This chapter concentrates on studies of patients' and therapists' perceptions of the professional relationship or of concrete structural aspects of sessions.

Orlinsky and Howard (1978) reviewed studies testing the association between therapeutic outcome and the number and frequency of sessions and the total duration of treatment. Of 33 studies of the total number of sessions, 20 reported a positive significant association with outcome, 7 reported an insignificant relationship, and 6 reported a positive curvilinear relationship. Of 22 studies that concerned the duration of treatment, 12 found a significant positive association, 9 found an insignificant association, and 1 found a significant negative association. Of 16 studies of frequency of sessions and outcome, about half found a significant positive association, about half an insignificant association, and 1 a significant negative association. Orlinsky and Howard noted that the studies varied widely in samples, methods, and control of pretreatment characteristics, especially health status. This is particularly important because patients who receive more intensive treatment may also have poorer mental health (Ware et al., 1984). Nevertheless, Orlinsky and Howard concluded that intensity of care is probably positively associated with outcome.

A large number of studies have addressed the affective nature of the therapist/patient relationship. As noted by Parloff, Waskow, and Wolfe (1978), virtually all schools of psychotherapy postulate that a positive therapeutic alliance is a necessary condition for successful psychotherapy. One of the most widely studied models of this alliance was proposed by Rogers (1957). In his original scheme, he proposed six conditions as necessary and sufficient for effective treatment: (1) contact between the patient and therapist; (2) a state of vulnerability in the patient; (3) genuineness in the therapist; (4) unconditional positive regard for the patient (e.g., warmth and acceptance); (5) accurate understanding of the patient (empathy); and (6) the patient's perception of these qualities in the therapist.

As discussed by Parloff, Waskow, and Wolfe, numerous studies have tested the components of the original model. Most of these studies employed observer assessments of the therapists' qualities. The bulk of the evidence suggests that accurate empathy, nonpossessive warmth, and genuineness play an important role in outcome. However, there are some contradictory findings, and the generalizability of this conclusion is uncertain (Mitchell, Bozarth, and Krauff, 1977). Nevertheless, studies of patients' perceptions of these qualities in the therapist have fairly consistently demonstrated a positive relationship to outcome (Gurman, 1977). Parloff, Waskow, and Wolfe (1978) cautioned that the causal nature of this association is unclear because patients who
are improving may be more likely to perceive positive qualities in their therapists.

It is of particular interest for the MOS that patients' perceptions of these qualities in their therapists may relate to outcome. Orlinsky and Howard (1978) reviewed 30 studies of patients' perceptions of their therapists in three areas: the therapists' warmth, accuracy of understanding, and positive valuation of the patient. They concluded that these perceptions are significantly related to the outcomes of treatment. Further, the patients' perception that the therapist was critical or hostile was negatively associated with outcome in two studies reviewed. Orlinsky and Howard also reviewed studies of therapists' self-perceptions of these qualities and found less consistent results.

Studies of the relation of process and outcome have focused on many other components of the physician/patient relationship, such as how much the participants like each other, the nature of specific psychotherapeutic interventions, and the general activity level of the therapist. Of these concepts, the therapist's general activity level may be of some value for the MOS. As reviewed by Orlinsky and Howard (1978), a number of studies have suggested that an active, approving approach is associated with better outcomes than is a reflective approach, (i.e., one that emphasizes clarification and interpretation). However, most of these studies used observer assessments of activity. It is not clear that this association would remain if pretreatment characteristics of patients were controlled.

Frank's (1974) conceptual model of the psychotherapy process may be of value for the MOS. He hypothesized that the beneficial effects of psychotherapy of all schools are due to a set of common factors that include understanding, respect, interest, encouragement, acceptance and forgiveness. In addition, Frank proposed that the sharing of a conceptual scheme of the nature of psychological problems and their treatment enhanced patient morale by providing an explanation and instilling hope. Although this formulation has not been empirically tested, sharing a conceptual scheme is included as a component of the dimension of specific psychotherapeutic activities in Table 7.1.

SUMMARY AND IMPLICATIONS FOR THE MOS

The process of care can be conceptualized as consisting of the diagnostic evaluation, the treatment plan, the general intensity of care, specific psychotherapeutic activities, and the affective quality of the physician/patient relationship and specific therapeutic activities. Existing studies of the link between the process and outcomes of care
for depressive disorders have addressed the efficacy of specific modalities (e.g., cognitive therapy, behavior therapy, psychotropic drugs) administered by trained mental health specialists. As discussed in Chap. 4, these studies suggest that the several treatment modalities are positively related to outcome. One study of the value of specific psychotherapeutic activities in the treatment of depression supports the view that pretreatment characteristics of the patient, rather than specific psychotherapeutic activities, are the major determinants of outcome (Rounsaville, Weissman, and Prusoff, 1981).

Taken as a whole, these studies suggest that, at a minimum, the MOS investigators should examine the provider's choice of treatment and the appropriateness of its administration (e.g., drug dosage). Psychiatrists and general medical providers may differ considerably in such simple, gross aspects of care (Verbrugge, 1984). Further, this dimension of process (the general treatment strategy) seems particularly appropriate for inclusion in the MOS because of the study's focus on the care delivered by general medical providers. In addition, the MOS investigators should consider including measures of the general affective nature of the physician/patient relationship, although this process dimension has not been linked to outcomes while controlling for pretreatment severity of illness.
Chapter 8

CASEFINDING TECHNIQUES

GENERAL STRATEGY

Past studies have used a variety of strategies to identify persons with clinically significant depression, including the clinical diagnosis of a treating physician or researcher/clinician (e.g., Johnson, 1973, 1974), scores on survey measures of mental health (e.g., Beck and Beck, 1972; Zung and King, 1983), structured psychiatric interviews (e.g., Weissman, Myers, and Harding, 1978; Hoeper et al., 1979; Regier et al., 1984), and a variety of laboratory tests, such as the dexamethasone suppression test (DST). For the MOS, we require an instrument for detecting major depression that is independent of the diagnosis of the treating physician, because the detection of depression by the clinician is one of the treatment components to be studied. While self-assessment measures of depression provide an independent assessment of depressive symptoms, they do not provide a specific diagnosis of depressive disorders. Laboratory tests may be used in conjunction with a clinical assessment to enhance confidence in a clinical diagnosis of affective disorder, but they do not provide a definitive diagnosis (Carroll, 1983; Gold et al., 1981). Specific diagnoses are needed in the MOS, however, to control for case-mix and to evaluate the course of treatment. Structured psychiatric interviews are appropriate for this purpose.

Structured interviews have the advantage of standardizing the diagnostic process. As discussed by Endicott and Spitzer (1978), structured interviews reduce information variance and criterion variance, two major sources of low reliability in clinical diagnosis. Information variance occurs when different clinicians have different information available for deciding on a diagnosis. Criterion variance refers to “differences in the inclusion and exclusion criteria that clinicians use to summarize data into psychiatric diagnoses” (Endicott and Spitzer, 1978, p. 837). An additional advantage of structured interviews is that lay persons can be trained to administer them, thus increasing the feasibility of research.

Despite the advantages of a structured psychiatric interview, it would be very costly to administer one to every patient in the full MOS sample. The present author therefore recommends using a two-stage
case-finding procedure, such as that employed by Hoepfer et al. (1979). In a two-stage procedure, each patient completes a self-report measure of depression. Those exceeding a standard cut-off are administered a structured psychiatric interview to establish the specific diagnosis of major depression.

This chapter reviews the two structured psychiatric interviews most widely used in the United States: the Schedule for Affective Disorders and Schizophrenia (SADS) and the Diagnostic Interview Schedule (DIS).

SCHEDULE OF AFFECTIVE DISORDERS AND SCHIZOPHRENIA (SADS)

Endicott and Spitzer (1978) describe the SADS in detail. There are three versions: the regular version, the lifetime version (SADS-L), and a version for measuring change in status (SADS-C).

The regular SADS is divided into two parts. Part 1 obtains detailed information on the patient’s current episode and functioning in the past week. Part 2 obtains information on past psychiatric problems. While the SADS-L also obtains information on both current and past psychiatric problems, it does not obtain a detailed description of the current episode or of functional status in the past week. The SADS includes ratings of symptoms both in the past week and when they were most severe. Thus there is an implied measure of change.

The SADS is designed to be used by mental health clinicians. A complete evaluation takes 90 to 120 minutes. At the end of the interview, the clinician assigns the diagnosis. In addition, items from Part 1 can be aggregated into scales as indicators of eight dimensions of psychopathology (Table 8.1). The inter-rater reliability of the scale scores range from 0.82 to 0.99, the test/retest reliability from 0.49 to 0.93, and internal consistency from 0.47 to 0.87. The correlations among the three “depression” scales (depressive mood and ideation, endogenous features, depression-associated features) are as high (0.72 to 0.91) as their reliabilities (0.79 to 0.87), suggesting that the constructs may not be discrete. However, the correlations of these three scales with suicidal ideation and behavior are more modest (0.40 to 0.52). The summary scales, particularly the four depression scales, are strongly correlated with the HSCL depression factor (0.47 to 0.68).

The reliability of RDC diagnoses assigned by clinicians using the SADS and SADS-L has been reported by Spitzer, Endicott, and Robins (1978). In two studies, diagnoses were assigned independently by one clinician conducting the interview and one who observed. In a third
Table 8.1

DIMENSIONS OF PSYCHOPATHOLOGY: SUMMARY SCALES FROM THE SADS INTERVIEW

<table>
<thead>
<tr>
<th>Scale</th>
<th>Internal Consistency (N = 150)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depressive mood and ideation</td>
<td>0.87</td>
</tr>
<tr>
<td>Endogenous features</td>
<td>0.80</td>
</tr>
<tr>
<td>Depression-associated features</td>
<td>0.79</td>
</tr>
<tr>
<td>Suicidal ideation and behavior</td>
<td>0.80</td>
</tr>
<tr>
<td>Anxiety</td>
<td>0.58</td>
</tr>
<tr>
<td>Delusions/hallucinations</td>
<td>0.87</td>
</tr>
<tr>
<td>Formal thought disorder</td>
<td>0.47</td>
</tr>
</tbody>
</table>

SOURCE: Adapted from Endicott and Spitzer (1978).

study, two interviewers rated the patient at different times. The patients in two studies were newly admitted inpatients at a state hospital, and those in the third were newly admitted inpatients with a depressive or manic syndrome from the Pilot Study of the Psychobiology of Depressive Disorders (N = 150). The kappa statistic (agreement between raters, adjusted for chance agreement) for both current episode and lifetime diagnoses was high (0.75 to 1.00) for joint interviewers. The test/retest kappas were high for present episode (0.65 to 1.00) and more variable for lifetime diagnoses (0.40 to 0.95). The kappas were .88 to .90 for major depressive disorders and 0.81 for minor depression. Across all three studies, the kappa statistics for the nine subtypes of current episodes of major depression ranged from 0.88 to 0.95.

Endicott et al. (1981) developed a technique to use the SADS interview data to derive a measure of severity of depression comparable to the HRS, one of the most widely used observer ratings of severity of depression (described in Chap. 9). The inter-rater reliability of the HRS extracted from the SADS is 0.95, and the correlation of the real HRS with the version extracted from the SADS-C is 0.92. Thus, the SADS can be used to obtain a measure of severity of depression as well as dimensions of psychopathology and specific RDC diagnoses.
THE DIAGNOSTIC INTERVIEW SCHEDULE (DIS)

Robins et al. (1981) provide a detailed description and history of the DIS, a structured interview designed to assign psychiatric diagnoses by the DSM-III, RDC, or Feighner Criteria. The DIS is almost unique among structured interviews in the degree of specification of its questions, probes, and scoring system. As a result of this specificity, it can be administered by specially trained lay interviewers, enhancing its usefulness as a research tool. The full DIS takes 45 to 75 minutes to complete, and diagnoses are assigned by a computer algorithm.

The DIS assigns many more diagnoses than does the SADS. Table 8.2 lists the DSM-III diagnoses that can be assigned with the DIS. A diagnosis of major depressive disorder can be assigned according to the DSM-III, RDC, or Feighner Criteria. By contrast, dysthymic disorder is the only nonmajor depressive disorder that can be assigned. In particular, it is not possible to use the DIS to make an RDC diagnosis of minor or intermittent depression.

In addition to including criteria for more diagnoses, the DIS provides more comprehensive information on psychiatric symptoms than does the SADS. For example, the SADS does not obtain complete information on some disorders if a preemptive diagnosis (e.g., schizophrenia) has already been made. By contrast, the DIS elicits information on all disorders, and the computer program subsequently applies certain exclusion criteria to determine diagnostic priorities.

The reliability of the DIS has been reported for 216 persons who were interviewed on separate administrations of the DIS by lay interviewers and psychiatrists (Robins et al., 1981). The psychiatrists were free to ask other questions if they needed to arrive at a diagnosis. The subjects included 118 psychiatric inpatients, 39 psychiatric outpatients, 10 members of Gamblers Anonymous, 26 ex-patients, and 24 enrollees of a group health plan who had no known psychiatric diagnosis. The indicators of validity in this study included the degree of bias of the lay interviews (McNemar's test), agreement between lay and psychiatric interviews (kappa statistic), and the sensitivity and specificity of lay interviews using psychiatric interviews as the criterion measure.

Bias in diagnoses based on lay-administered interviews included underdiagnosis of alcohol dependence, somatization disorder, panic disorder, and overdiagnosis of drug abuse. The greatest bias was in underdiagnosis of somatization disorder. The main reason for this bias was that lay interviewers were trained to accept each patient's word on the cause of his or her symptoms, while clinicians exercised their own judgment as to the origin of the symptom. This may be of some importance for the MOS, because the DIS would be applied to patients
Table 8.2

<table>
<thead>
<tr>
<th>Code</th>
<th>DSM-III Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>290.xx</td>
<td>Senile and presenile dementias</td>
</tr>
<tr>
<td>295.xx</td>
<td>Schizophrenia</td>
</tr>
<tr>
<td>296.xx</td>
<td>Major depression</td>
</tr>
<tr>
<td>296.3x</td>
<td></td>
</tr>
<tr>
<td>296.5x</td>
<td></td>
</tr>
<tr>
<td>296.6x</td>
<td></td>
</tr>
<tr>
<td>296.4x</td>
<td>Bipolar disorder, manic</td>
</tr>
<tr>
<td>300.03</td>
<td>Panic disorder</td>
</tr>
<tr>
<td>300.21</td>
<td>Agoraphobia</td>
</tr>
<tr>
<td>300.22</td>
<td></td>
</tr>
<tr>
<td>300.29</td>
<td>Simple phobia</td>
</tr>
<tr>
<td>300.30</td>
<td>Obsessive-compulsive disorder</td>
</tr>
<tr>
<td>300.40</td>
<td>Dysthymic disorder</td>
</tr>
<tr>
<td>300.81</td>
<td>Somatization disorders</td>
</tr>
<tr>
<td>301.70</td>
<td>Antisocial personality disorder</td>
</tr>
<tr>
<td>302.7x</td>
<td>Psychosexual dysfunction</td>
</tr>
<tr>
<td>302.00</td>
<td>Ego-dystonic homosexuality</td>
</tr>
<tr>
<td>302.5x</td>
<td>Transsexualism</td>
</tr>
<tr>
<td>305.0x</td>
<td>Alcohol abuse</td>
</tr>
<tr>
<td>303.9x</td>
<td>Alcohol dependence</td>
</tr>
<tr>
<td>305.1x</td>
<td>Tobacco use disorder</td>
</tr>
<tr>
<td>305.2x</td>
<td>Cannabis abuse</td>
</tr>
<tr>
<td>304.2s</td>
<td>Cannabis dependency</td>
</tr>
<tr>
<td>305.3x</td>
<td>Hallucinogen abuse</td>
</tr>
<tr>
<td>305.4x</td>
<td>Barbiturate abuse</td>
</tr>
<tr>
<td>304.1x</td>
<td>Barbiturate dependence</td>
</tr>
<tr>
<td>305.5x</td>
<td>Opioid abuse</td>
</tr>
<tr>
<td>304.0x</td>
<td>Opioid dependence</td>
</tr>
<tr>
<td>305.6x</td>
<td>Cocaine abuse</td>
</tr>
<tr>
<td>305.7x</td>
<td>Amphetamine abuse</td>
</tr>
<tr>
<td>304.4x</td>
<td>Amphetamine dependence</td>
</tr>
<tr>
<td></td>
<td>— Other drug abuse</td>
</tr>
<tr>
<td></td>
<td>— Other drug dependence</td>
</tr>
<tr>
<td>307.10</td>
<td>Anorexia nervosa</td>
</tr>
<tr>
<td>312.31</td>
<td>Pathological gambling</td>
</tr>
</tbody>
</table>

SOURCE: Adapted from Robins et al. (1981).

derived from medical practices, where patients may be more likely to give physical explanations for their symptoms. To date, no study has reported the validity of the DIS in medical settings, so the degree of additional bias that might occur in these settings is unknown.
The DIS is being used in a number of large-scale epidemiologic studies, especially the NIMH Epidemiologic Catchment Area Program (Eaton et al., 1981; Regier et al., 1984). Each research site is conducting studies of the reliability and validity of the DIS, including further comparisons between lay and clinical interviews and among other interview schedules (e.g., the RDC). Two of these studies have been published (Helzer et al., 1985; Anthony et al., 1985).

Helzer et al. (1985) compared lifetime diagnoses derived from lay-administered interviews, using the DIS with clinical diagnoses assigned by psychiatrists. The clinical interview was also based on the DIS format. Overall agreement ranged from 79 percent to 96 percent, depending on diagnosis. There was a significant negative bias for major depression by the lay interview (i.e., underdetection of new cases). The weighted sensitivity for major depression was 59 percent, but specificity was 100 percent. Much of the disagreement over the diagnosis of major depression was due to cases that (by the lay DIS) just missed the diagnostic criteria (i.e., had three rather than four symptoms of depression during depressed periods).

Anthony et al. (1985) compared current (one-month) diagnoses assigned by lay interviewers using the DIS and by psychiatrists using an adaptation of the Present State Examination (PSE). They found a much lower level of agreement than that reported by Helzer et al. (1985) for all diagnostic categories. The Kappa statistic was 0.25 for major depressive disorders. However, the only disorder with a higher Kappa statistic was alcohol use disorder. The authors identify several possible sources of disagreement, including differences in diagnostic criteria and the degree of reliance on self-report information.

Table 8.3 lists the relative advantages and disadvantages of the SADS and DIS. A major limitation of the DIS relative to the RDC is the absence of measures of dimensions of psychopathology, global severity, or changes in disease status. However, the DIS is preferable for the MOS because it can be administered by lay interviewers and because of the flexibility with which diagnoses can be assigned (i.e., by any of three classification schemes).
Table 8.3

COMPARISON OF SADS AND DIS

<table>
<thead>
<tr>
<th>Item</th>
<th>SADS</th>
<th>DIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original purpose</td>
<td>Assign diagnoses by RDC</td>
<td>Assign diagnoses by DSM-III, RDC and Feighner criteria</td>
</tr>
<tr>
<td>Method of administration</td>
<td>Structured interview by clinician</td>
<td>Lay or clinician interview</td>
</tr>
<tr>
<td>Length of time to complete</td>
<td>90 minutes to 2 hours</td>
<td>45 to 75 minutes</td>
</tr>
<tr>
<td>Special requirements</td>
<td>Clinician interviewer</td>
<td></td>
</tr>
<tr>
<td>Products</td>
<td>Current and lifetime RDC diagnoses, global assessment scale, subscales for eight dimensions of psychopathology</td>
<td>Current and lifetime diagnoses, complete description of all syndromes</td>
</tr>
<tr>
<td>Reliability of scales/diagnoses</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>Advantages</td>
<td>Includes measures of change, global severity scale, psychopathology scales</td>
<td>Lay interviewers. Three diagnostic schemes. Large number of DSM-III diagnoses.</td>
</tr>
<tr>
<td>Disadvantages</td>
<td>Does not include all diagnoses. Some criteria (e.g., exclusion criteria for minimally depressive disorders) need updating. Requires clinician/interviewer.</td>
<td>No established symptom dimensions, global assessment, or measure of change.</td>
</tr>
</tbody>
</table>
Chapter 9

OTHER MEASURES OF DEPRESSION

In the MOS, measures of depressive symptoms are needed for the purposes of first-stage screening and to indicate the severity of depression. We are selecting measures of depressive symptoms on the basis of their suitability for these particular research purposes.

To be useful for first-stage screening, a measure of depression should have a very high sensitivity for depressive disorders (e.g., ≥ 90 percent). A high specificity is also desirable to minimize the number of patients who would need to receive a structured interview. However, a moderately low specificity—e.g., 50 to 60 percent—could be tolerated for the sake of high sensitivity.

The second potential use for a depression measure in the MOS is to indicate the severity of the depressive illness. In the MOS, a measure of severity of depression at intake will be used as a covariate, and a measure at exit (and other intervals after intake) will be used as an outcome measure. The severity measure should have several characteristics. It should measure some or all clinically meaningful aspects of the disorder (e.g., mood, endogenous features, social functioning). Specific scale scores indicating different levels of severity should have been previously validated using clinician assessments of severity as the criterion. These scale scores should not only discriminate cases from noncases but should also discriminate among patients with depressions of different severity. Changes in severity for the same patients over time (e.g., as a result of treatment) should be reflected in changes in scale scores. In particular, the measure should discriminate between patients and recovered patients. The measure should not be sensitive to differences in severity of nondepressive (psychiatric or physical) illnesses (after removing the effect of any coexisting depression).

The MOS design requires separate measures of depression for first-stage screening and for determining the severity of depression. If we were to use one measure for both purposes, the observed outcomes at exit for a particular patient group (e.g., patients of general medical providers) would be affected by statistical regression to the mean. While comparisons between patient groups (e.g., patients of fee-for-service providers versus HMOs) would not be so affected, the MOS investiga-
tors are also interested in describing the outcomes obtained for each patient group separately.

This chapter reviews several self-assessment and observer assessment measures of social functioning, and makes recommendations for selecting instruments for first-stage screening and assessing severity of depression.

SELF-ASSESSMENT MEASURES OF DEPRESSION

There are six widely used self-assessment measures of depressive symptoms or psychological distress: the Beck Depression Inventory (BDI), the Center for Epidemiologic Studies Depression Scale (CES-D), the General Health Questionnaire (GHQ), the Hopkins Symptom Checklist (HSCL); the Zung Self-Rating Depression Scale (SDS); and the Mental Health Inventory (MHI). The BDI, CES-D, and SDS were each specifically designed to measure clinically meaningful symptoms of depression, regardless of diagnosis. The GHQ and HSCL were designed to measure general psychopathology (especially anxiety and depression) in patient populations. The MHI was designed specifically to measure mental health status in the Health Insurance Experiment (HIE).

For each measure, the author recommends specific uses of the measure, summarizes evidence for the reliability and validity of scales, and describes the item content and factor structure of the measure. The discussion of the BDI, HSCL, CES-D, and GHQ is based in part on a review by Murphy (1980).

Beck Depression Inventory (BDI)

The BDI was designed to measure the amount or depth of depression, regardless of the patient's diagnosis, in clinical psychiatric settings (Beck, 1967; Beck et al., 1961). While the BDI was originally administered by interview, it is now widely used as a self-administered questionnaire. The original version included 21 "items" or categories that reflected symptoms of depression. A 13-item form is also available.

Each item consists of four or more statements that vary in severity in terms of degree of change from the respondent's usual state, frequency of symptoms, or the intensity of distress. The responden
selects the one statement from each category that is most descriptive of 
his or her state at the time of the test.

Beck (1967) originally proposed three cutoff scores to indicate three 
levels of severity. However, the definition of the cutoff scores was 
unusual: ≥ 17 discriminated between no depression and moderate 
severity, ≥ 21 between no depression and severe depression, and 
≥ 26 between mild and severe depression. Subsequently, cutoffs for 
adjacent severity levels (none, mild, moderate, severe) have been 
developed for patients in general practice (Salkind, 1969).

The BDI includes two factors: a guilt/self-debasement (affective) 
factor and a somatic (physiologic) factor. The internal consistency of 
the total scale is close to 0.90.

The BDI was validated against psychiatrists’ global ratings of the 
depth of depression (none, moderate, severe). The BDI scores and the 
psychiatrists’ scores agreed 56 percent of the time on the same level of 
severity and 97 percent of the time within one level of severity. The 
correlation between the psychiatrists’ ratings and the BDI score was 
0.65 and 0.67 in two samples. The BDI correlates significantly (0.62 to 
0.77) with several other measures of depression (Dahlstrom, Welsh, 
and Dahlstrom, 1972; Nussbaum et al., 1963; Hamilton, 1962; Schwab 
et al., 1965, 1966; Hamilton, 1967). The correlation of the BDI with 
the HRS, a standard observer-scale of severity of depression, is 0.72 to 
0.82. The BDI also appears to be sensitive to changes over time with 
treatment. For example, in one study, 90 percent of patients assessed 
as clinically improved had significantly lower BDI scores after two to 
six weeks of treatment than prior to treatment. Moreover, the BDI 
appears to discriminate between depressed and anxious patients. The 
correlation of the BDI with clinical assessments of depression is 0.59, 
but only 0.44 with assessments of anxiety.

The BDI has been widely used in clinical trials of antidepressant 
drugs and as a measure of depression in psychiatric and medical 
patients. While the BDI may be sensitive to changes in physical 
health status, it appears to be less sensitive to physical health than are 
some other measures of depression. For example, some information is 
available on the response of BDI scores to physical illness. Schwab et 
al. (1965) selected medical inpatients who exceeded cutoffs on both the 
BDI and the Hamilton Depression Scale. Somatic symptoms 
accounted for 29 percent of the variance in BDI scores but 57 percent 
of the variance in the Hamilton ratings.
The Center for Epidemiologic Studies
Depression Scale (CES-D)

The CES-D was designed to identify groups with significant depression in general populations (e.g., for use in epidemiologic studies). Radloff (1977) described the development and psychometric properties of the CES-D. Like the BDI, it detects depression across diagnostic categories. In contrast to the BDI, it has been used primarily to identify probable cases of depression, rather than to measure severity of depression.

The CES-D consists of 20 items. The language used is very simple and direct. Words such as "distressed" or "troubled" are avoided entirely, and it contains no items that directly measure suicide ideation or psychotic symptoms. As a result, it appears to be especially useful for patients in general medical settings. For each item, there are four response categories indicating the frequency with which the symptom has occurred in the past week. Each response category is clearly defined. The item content includes general dysphoric mood and crying, positive affect, and vegetative-motor symptoms (e.g., poor appetite). The CES-D does not include terms that measure vague physical complaints (e.g., aches and pains).

The CES-D appears to be sensitive to variation in both short- and long-term depressed mood (i.e., acute and chronic depression). Its test-retest reliability over periods of 2 to 8 weeks is 0.51 to 0.67, but 0.32 to 0.54 for longer periods (3 to 12 months). Its internal consistency is high (0.85 in the community; 0.90 in psychiatric patient populations).

Factor analyses have demonstrated the presence of three factors: depressed affect, positive affect or well-being, and vegetative-motor signs (Radloff, 1977). All the items in the depressed affect factor are negatively worded, and those in the positive affect factor are positively worded. As a result, it is not clear how much these factors represent a methods effect. However, the factor structure is stable for several ethnic groups.

The CES-D has been validated in patient and community populations (Craig and Van Natta, 1976, 1978; Weissman et al., 1977). In these studies, the correlation of CES-D scores with several other measures of depression was moderately high (0.51 to 0.89). The highest correlation in several studies was with the HSCL anxiety scale (Murphy, 1980).

Studies of the sensitivity and specificity of the CES-D have had mixed results. In one study, the CES-D had a sensitivity of 89 percent and a specificity of 56 percent for identifying acutely depressed
patients, using the Raskin Clinical Rating Scale as the criterion measure. The sensitivity for depressed outpatients in that study was 91 percent. Myers and Weissman (1980) compared CES-D scores to RDC diagnoses assigned by the SADS in a community sample. The sensitivity of the CES-D for detecting major depression was only 64 percent. Of those over a standard cutoff, 33 percent had major depression, 33 percent no diagnoses, and 18 percent another psychiatric diagnosis. Of those under the cutoff, 2 percent had a major depression. Minor alterations in the cutoff score did not alter the sensitivity appreciably. Boyd, Weissman, and Thompson (1982) explored the reasons for discrepancies between the SADS diagnoses and CES-D scores. The main reason for the low sensitivity was the tendency of some subjects to be "nay sayers," i.e., they denied (true) symptoms of depression on the written form. Other subjects received falsely low scores because of difficulty completing the form, possibly because of psychomotor retardation.

The CES-D has been shown to be sensitive to changes in symptoms over time (Weissman et al., 1977). Depressed patient groups who were thought to have recovered by clinical judgments and the Raskin Clinical Rating Scale had a significant drop in their CES-D scores. However, the CES-D scores remained somewhat high, compared with the therapists' own rating of improvement and the Raskin Rating Scale.

The CES-D has recently been widely used in studies of general populations, psychiatric patients, and general medical patients (Husaini, Neff, and Stone, 1979; Husaini, et al., 1979; Craig and van Natta, 1978; Goldberg, Comstock, and Hornstra, 1979; Roberts, 1980). The current consensus is that the CES-D appears to be useful for screening purposes but has only a modest relation to cases of identified depression as defined by specific diagnostic criteria (Myers and Weissman, 1980). Further data on the sensitivity and specificity of the CES-D in identifying DSM-III (DIS) disorders will be available in the near future from the Los Angeles ECA site and other studies.

**Zung Self-Assessment Depression Scale (SDS)**

Like the BDI and the CES-D, the SDS was designed specifically to measure depressive symptoms across diagnostic categories (Zung, 1965; Hedlund and Vieweg, 1979b). The measure consists of 20 items with four response choices for each item to indicate how much of the time the respondent has had the symptom described. The language used in the SDS is very clear.

The reliability (internal consistency) of the SDS is 0.79 (Giambra, 1977). The SDS has a correlation of 0.87 with a semi-structured
interview that assesses depression (Zung, 1974) and between 0.40 and 0.80 with the Hamilton rating scale. However, Downing and Rickels (1972) found that its correlation with global judgments of depression was only 0.45. Wright et al. (1980) obtained a similar finding. These researchers offered a psychiatric evaluation to medical outpatients who had scores >50 on the SDS. Of those exceeding the cutoff, 32 percent were interviewed, and 41.7 percent of these met Feighner criteria for depression. Those with the most serious medical illnesses had the highest scores.

While the SDS has high correlations with other measures of depression in some studies, high correlations are not found consistently (Hedlund and Vieweg, 1979b). For example, in a study of mentally retarded adults with depression, Kazdin, Matson and Senatore (1983) compared the BDI, SDS, MMPI and HRS. The BDI was significantly correlated with all the other measures. No other measure of depression correlated highly with all the other measures.

There is some evidence that the SDS scores are responsive to changes in depression that occur as a result of treatment. Zung and King (1983) identified depressed patients in a primary care setting. Patients who exceeded a cutoff score were randomly assigned to either a trial of alprazolam and notification of their physicians or to their usual care, with their physicians contacted after a four-week waiting period. The experimental group demonstrated a significant decrease in SDS scores. The control also improved, but not as much as the experimental group.

General Health Questionnaire (GHQ)

The GHQ, unlike the three previous measures, was designed both to detect psychiatric disorders in primary care outpatient settings and to serve as a measure of severity of psychopathology. The original GHQ consisted of 60 items. Subsequently, 30-, 28-, 20-, and 12-item versions have been developed (Vieweg and Hedlund, 1983). The content of the original 60-item version included depressed mood and anxiety, social functioning, psychophysiologic symptoms, general health, and vague aches and pains. The depression items refer primarily to severe depression and suicidal ideation. The shorter versions generally exclude items measuring physical health status. The response categories refer to the degree of change from the usual pattern during the last few weeks. The internal consistency of the various versions is generally between 0.80 and 0.90, and the test-retest reliability is 0.90 at six months and 0.75 at one year (for the 60-item version) (Vieweg and
Hedlund, 1983). The correlation of the GHQ with overall clinical assessments of psychopathology ranges from 0.55 to 0.83.

Factor analyses of the 60-item version indicate that the GHQ is comprised of four factors: somatic symptoms (general health), anxiety and insomnia, social impairment, and severe depression. The 60-item version was validated using patients with and without psychiatric disorders (Goldberg and Blackwell, 1970; Goldberg, 1972). The 28-item version was specifically designed as a shorter version that maintains the four-factor structure of the 60-item version. The clinical utility of the separate factor scores has not been established (Cleary, Goldberg, and Kessler, 1982).

The sensitivity and specificity of the GHQ in detecting psychiatric cases is highest for the 60-item version. Hoepner et al. (1979) compared GHQ scores with RDC diagnoses in a primary care population. The sensitivity of the GHQ for detecting major depression was 100 percent. By contrast, the overall sensitivity of the GHQ in this population was 68.1, and the specificity 80.9. Additional information on the sensitivity and specificity of the GHQ in detecting DSM-III (DIS) psychiatric disorders is forthcoming from the John Hopkins ECA site.

**Hopkins Symptom Check List (HSCL)**

The HSCL was designed to measure *improvement* in clinical status of psychiatric patients in response to psychotherapy (Parloff, Kelman, and Frank, 1954; Kelman and Parloff, 1957). The questionnaire consists of a list of symptoms; the respondent indicates how much bother or distress each symptom has caused him or her in the last week.

The longest version (90 items) includes items that measure depression, anxiety, somatic symptoms, and paranoia and psychoticism. The shortest version (25 items) measures only depression and anxiety and is designed for primary care settings. Versions of intermediate length also exist.

Many studies using the HSCL show highly significant changes during or following treatment (Murphy, 1980). It has been used in epidemiologic studies, in clinical trials of psychotherapy and chemotherapy, and in evaluations of medical patients (Murphy, 1980; Craig and Abeloff, 1974; Hesbacher, Rickels, and Goldberg, 1975; Glass et al., 1978; Uhlenhuth et al., 1974). Its factor structure has been evaluated in a number of studies (Derogatis et al., 1970, Dinning and Evans, 1977). The original version includes five factors: anxiety, depression, interpersonal sensitivity, obsessive compulsive traits, and somatizations. The 90-item version also includes factors for hostility, phobic anxiety, paranoid ideation, and psychoticism. The factor structure is stable across gender and social classes.
One advantage of the HSCL is that the total score and the factor scores can be used. For example, depressed patients tend to score higher on the depression factor than on the anxiety factor and anxious patients tend to score higher on the anxiety factor than on the depression factor. The correlation of the total score with psychiatrists' global ratings of severity is between 0.70 and 0.77. Test-retest reliability over six months is 0.72; internal consistency of the total score is 0.95.

At least one study has reported a low correlation of HSCL scores to observer ratings of the severity of depression. Prusoff, Klerman, and Paykel (1972) followed 200 depressed patients (outpatients, day-care patients, emergency treatment service patients, and inpatients) through an acute depressive episode. Scores on a 58-item version of the HSCL and interviewer assessments on the HRS were taken at admission and at ten months after admission. Correlations between measures for item clusters of similar content ranged from 0.63 to 0.11, with assessments of depressed mood correlating only 0.31. However, at recovery, the correlation of total HSCL and HRS scores was 0.81. Prusoff, Klerman, and Paykel found that, during the acute depression, the discrepancies between self- and interviewer assessment were greatest for psychologically depressed patients. Thus, severe depression itself may interfere with the validity of self-reports of symptoms during the acute phase of illness. White, White, and Razani (1984) obtained a similar result in a comparison of scores on the SDS (patient) and the HRS (clinician) for 36 inpatients with major depression; disagreement was highest when depression was rated as severe by the clinician.

Kass et al. (1983) have also concluded that caution is needed in interpreting HSCL scores in psychiatric patient populations. They found that the correlation between patients' HSCL scores and scores obtained from forms filled out by their therapists was only 0.17. Underreporting by patients was caused by fear or paranoia, and overreporting by a demonstrative style of expression. Additional false positives were due to physical illness.

Uhlenhuth et al. (1983) have recently used a brief measure of psychic distress (PSYDIS), an adaptation of the HSCL, to classify persons as having DSM-III-like syndromes on the basis of their factor scores. For example, persons with a "major depressive" syndrome had a high score on the depression factor and positive scores on four of five symptom clusters: decreased energy and interest, impaired cognitive functioning, sleep disturbance, loss of sexual interest or pleasure, and appetite disturbance. Uhlenhuth and his colleagues found that the prevalence of the major depression syndrome in a general population was similar to that reported in other studies for major depressive disorder in the general population.
Mellinger and his colleagues tested the validity of the PSYDIS against structured psychiatric assessments on 287 persons drawn from a household sample (Mellinger et al., 1983). Mellinger and his colleagues categorized persons as high or low in psychic distress and compared these categories with psychiatrists’ ratings of the presence or absence of any DSM-III diagnosis of anxiety or depression, whether treatment was indicated, and the severity of illness. The Kappa statistic, indicating agreement between the PSYDIS category and psychiatric assessment over and above chance, was 0.35 for the presence or absence of a diagnosis, 0.38 for indication for treatment, and 0.45 for severity assessment. Overall levels of concordance ranged from 76 percent to 80 percent. However, the specific DSM-III syndromes reported in Uhlenhuth et al. (1983) have not been validated against assessment of specific DSM-III diagnoses.

The Mental Health Inventory (MHI)

The MHI is a 38-item self-administered questionnaire that was designed specifically to measure mental health in the Health Insurance Experiment (HIE). Unlike some of the other measures discussed in this section, the MHI has not been evaluated against structured clinical assessments of the presence or absence of specific psychiatric diagnoses. Nevertheless, we are considering including the MHI as a measure of mental health status in the MOS because the MOS investigators have extensive experience with the MHI and because a very short (five-item) version is available and would enhance the feasibility of frequent measurement of mental health status.

The MHI was based on the General Well-Being Schedule (Dupuy, 1974) and on items described by Costello and Comrey (1967); Dohrenwend and his colleagues (1980); Beck (1967); and Zung (1965). The MHI defines mental health in terms of two substantially interrelated constructs: psychological distress and psychological well-being. Distress is a measure of the frequency or intensity of symptoms of anxiety and depression as well as a measure of behavioral and emotional control during the past month. Well-being is a measure of general positive affect and interpersonal ties.

The content and structure of MHI items differ from other general population mental health surveys in several important respects. First, the items focus exclusively on psychological states; the items do not involve physiological symptoms, functional status, health worry, health habits, and other general health constructs that were measured by the MHI’s predecessors (Ware et al., 1979). Second, the MHI includes many items measuring positive emotional states. In addition to
discriminating among those who are symptomatic, it seeks to discriminate among those who report no symptoms of psychological distress. Those persons would receive the same score (a perfect score) if positive items were not included.

The MHI has a factor structure that is generalizable across the four major HIE sites (Veit and Ware, 1983). The derived MHI scales reflect primary, secondary, and tertiary factors. Overall mental health status is represented by the total MHI score. Psychological distress and well-being are represented by two scales, which sum to the total score. Five subscales represent the dimensions within psychological distress (anxiety, depression, loss of behavioral or emotional control) and psychological well-being (general positive affect and interpersonal ties). Table 9.1 presents the abbreviated item content.

The total MHI score has a high reliability in all HIE sites. Its internal-consistency reliability is 0.96 in the total HIE sample. (These reliability coefficients also exceeded 0.90 for all income tertiles in the HIE.) The correlation between MHI scores based on administrations one year apart in the HIE sample is 0.64.

The validity of the MHI as a general measure of mental health has been evaluated in terms of a construct validity model. Specifically, hypothesized correlations have been confirmed with: (1) other health status measures (Ware et al., 1980); (2) measures of social contacts and social supports (Donald and Ware, 1982; Williams, Ware, and Donald, 1981); (3) stressful life events (Ware et al., 1979; Williams, Ware, and Donald, 1981); and (4) non-HIE studies with psychiatrist assessments of mental health, and other mental health measures.

The empirical results in Wells et al. (1982) and Ware et al. (1984) indicate that the total MHI score strongly predicts the use of mental health services. For example, the t-statistic for this variable in predicting the probability of any use of ambulatory mental health services is −7.97. The subsample in the bottom third of MHI (total score) uses nearly three times the ambulatory mental health services that the upper third uses.

**OBSERVER ASSESSMENTS**

As noted above, some authors have expressed concern that self-assessments of depression, especially of the severity of depression, may be of limited validity during the acute phase of illness. As a result, an observer rating of the severity of depression will be included in the MOS.
Table 9.1

PSYCHOLOGICAL DISTRESS AND WELL-BEING:
ITEM CONTENT OF MENTAL HEALTH
INVENTORY, ADULT VERSION

<table>
<thead>
<tr>
<th>Subscale/Item Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety</td>
</tr>
<tr>
<td>Very nervous person</td>
</tr>
<tr>
<td>Bothered by nervousness</td>
</tr>
<tr>
<td>Felt tense or high-strung</td>
</tr>
<tr>
<td>Anxious, worried</td>
</tr>
<tr>
<td>Difficulty trying to calm down</td>
</tr>
<tr>
<td>Nervous or jumpy</td>
</tr>
<tr>
<td>Restless, fidgety, impatient</td>
</tr>
<tr>
<td>Rattled, upset, flustered</td>
</tr>
<tr>
<td>Hands shake when doing things</td>
</tr>
<tr>
<td>Can relax without difficulty*</td>
</tr>
<tr>
<td>Depression</td>
</tr>
<tr>
<td>Moody, brooded about things</td>
</tr>
<tr>
<td>Low or very low spirita</td>
</tr>
<tr>
<td>Felt downhearted and blue</td>
</tr>
<tr>
<td>Felt depressed</td>
</tr>
<tr>
<td>Strain, stress, pressure*</td>
</tr>
<tr>
<td>Loss of behavioral/emotional control</td>
</tr>
<tr>
<td>Control behavior, thoughts, feelings</td>
</tr>
<tr>
<td>Concern about losing control of mind</td>
</tr>
<tr>
<td>Felt emotionally stable</td>
</tr>
<tr>
<td>Nothing turns out as wanted</td>
</tr>
<tr>
<td>Felt like crying</td>
</tr>
<tr>
<td>Better off if dead</td>
</tr>
<tr>
<td>Down in the dumps</td>
</tr>
<tr>
<td>Think about taking own life</td>
</tr>
<tr>
<td>Nothing to look forward to</td>
</tr>
<tr>
<td>General positive affect</td>
</tr>
<tr>
<td>Happy person</td>
</tr>
<tr>
<td>Happy, satisfied or pleased</td>
</tr>
<tr>
<td>Daily life interesting</td>
</tr>
<tr>
<td>Felt calm and peaceful</td>
</tr>
<tr>
<td>Felt cheerful, lighthearted</td>
</tr>
<tr>
<td>Generally enjoyed things</td>
</tr>
<tr>
<td>Relaxed and free of tension</td>
</tr>
<tr>
<td>Living a wonderful adventure</td>
</tr>
<tr>
<td>Expect an interesting day</td>
</tr>
<tr>
<td>Wake up fresh, rested</td>
</tr>
<tr>
<td>Future hopeful, promising</td>
</tr>
<tr>
<td>Emotional ties</td>
</tr>
<tr>
<td>Felt loved and wanted</td>
</tr>
<tr>
<td>Love relations full, complete</td>
</tr>
<tr>
<td>Felt lonely*</td>
</tr>
</tbody>
</table>

SOURCE: Adapted from Veit and Ware (1983).

NOTE: All items in this table are positively scored.

*These items were not used in subscales, but were used in the scales and global scale.
There are two ways that observer ratings of depression could be used in the MOS. First, an observer rating could be incorporated into the structured psychiatric interview, either by using an established observer measure or by deriving a severity score directly from the data collected during the administration of the depression section of the DIS. For example, the SADS includes a global measure of impairment, the Global Assessment Scale, as an optional component of the interview protocol. Endicott et al. (1981) developed a technique to extract a severity of depression score comparable to the HRS from the data collected in the SADS interview.

Second, the treating physician in the MOS can be asked to provide ratings of the severity of depression for all patients (in all tracers). This approach would not be valid for patients whose depression is undetected by the treating physician.

Below, we review the Hamilton Depression Rating Scale, the most commonly used observer rating of depression, and the observer rating scales developed by Raskin and used in the NIMH Collaborative Depression Studies (Raskin, 1972). Also reviewed is the Global Assessment Scale, an observer measure of impairment in functioning due to psychiatric illness.

**Global Assessments of Depression**

Many studies have used clinicians' global assessments to measure the severity of depression. While most studies asked the respondent to rate the overall severity of depression as mild, moderate, or severe (Simpson, Angus, and Edwards 1982; Nelson et al., 1982), some obtained global assessments on several specific dimensions. For example, Feighner et al. (1983) compared the response of depressed patients to antidepressants and a placebo. Among their assessment measures were separate global impression scales for physicians and patients. Physicians were asked to rate how mentally ill the patient was, how much the patient changed, and how efficacious the drugs were. In addition, there were criterion measures for the clinical assessment (HRS) and patient assessments (HSCL). The major advantage of unstructured global assessments for the MOS is that they make very little demand on the rater and thus may be appropriate for use by the treating physicians. The major disadvantage of many observer global assessment scales is low inter-rater reliability because of poor specification of response categories.
Hamilton Rating Scale for Depression (HRS)

The HRS is the most widely used observer scale for assessing depression (Hamilton, 1967; 1982; Hedlund and Vieweg, 1979a). The scale was designed specifically to measure severity in patients with known depressive disorders. It was originally intended for use by clinicians following a clinical interview of at least thirty minutes’ duration. The HRS exists in versions that vary the number and content of items (e.g., 6 to 24 items) and response categories. Paykel et al. (1973) have developed a semistructured interview that represents an expanded version of the HRS. The most commonly used version includes 17 items that describe the most common symptoms of depression, including depressed mood, anxiety, and somatic and endogenous symptoms (Table 9.2). For eight items, the observer simply records the presence or absence of the symptom. For the remaining items, the severity of the symptoms is rated on a scale of 0 to 4. The time frame is the last week or two.

The inter-rater reliability of the total score is generally about 0.90 (Hamilton, 1982; Hedlund and Vieweg, 1979a). The correlation between total scale scores and global judgments of severity has been reported at between 0.84 and 0.90. The correlations of HRS scores with self-assessment of depression are generally significant and positive. For example, Hedlund and Vieweg reported median correlation with the BDI of 0.58 for seven studies.

The HRS has been extensively used to evaluate the response of depressed patients to specific psychiatric treatments, especially to antidepressant medication. For example, recent reports indicate that depressed patients who respond to antidepressants show 61 to 67 percent improvement in HRS scores (DeWilde, Mertens, and Wakelin, 1983; Guelfi, Dreyfus, and Pichot, 1983). The HRS has been commonly used as a criterion measure for validating other measures of

<table>
<thead>
<tr>
<th>Table 9.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITEM CONTENT OF THE HAMILTON RATING SCALE FOR DEPRESSION</td>
</tr>
<tr>
<td>Depressed mood</td>
</tr>
<tr>
<td>Somatic symptoms</td>
</tr>
<tr>
<td>(gastrointestinal, general, genital)</td>
</tr>
<tr>
<td>Insight</td>
</tr>
</tbody>
</table>
depression. For example, Weissman et al. (1977) compared CES-D scores with HSR scores and other observer measures.

The HRS has also been used in studies of the outcome of psychotherapy. For example, Blackburn et al. (1981) compared drug therapy and cognitive therapy alone and in combination. They used the BDI to stratify patients by severity and monitored changes in clinical status by the BDI and the HRS. Hedlund and Vieweg (1979a), in a comprehensive review of studies of the HRS, concluded that the HRS and BDI are about equally sensitive to changes in depression in response to treatment.

Carr et al. (1981) developed a self-administered computerized version of the HRS (for patients) and compared scores for the computer version to clinician ratings (using the HRS). They obtained a correlation of 0.78 for depressed outpatients and of 0.72 for depressed inpatients. However, the version of the HRS they used assessed only depressed and anxious affect and excluded endogenous symptoms. As a result, this version seems to have no advantages over the BDI.

The HRS is included in the assessment batteries of several large, ongoing studies of depression, especially the NIMH Collaborative Studies on Depression (see, for example, Katz et al., 1982; Chevron and Rounsaville, 1983; Keller et al., 1983a, 1983b).

Other Observer Ratings of Depression

Raskin (1972) describes several observer rating scales used in the NIMH Collaborative Studies on Depression that have been widely used in clinical trials of antidepressant medication. They include a clinical global assessment scale, a three-area rating scale, the Brief Psychiatric Rating Scale (developed by Overall and Gorham, 1962), the Inventory of Psychic and Somatic Complaints, and additional scales to rate the behavior of patients on inpatient psychiatric wards. The scales were developed for use by clinician- evaluators following a clinical interview.

The level of clinical judgment required for most scales precludes their use as an adjunct to a lay interview. Further, the level of specific psychiatric knowledge required of the evaluator may preclude their usefulness with general medical providers. Nevertheless, one of these scales, the Inventory of Psychic and Somatic Complaints (IPSC), may be useful for the MOS. This scale is a version of the HSCL completed by the clinician. Like the HSCL, the checklist format is straightforward and easy to use. We were unable to find reports of its reliability, but the factor structure has been studied (Raskin, Schulerbrandt, and Reatig, 1967; Raskin et al., 1969; Raskin and McKeon, 1971), and the factors are sensitive to changes in psychopathology with treatment.
While other investigators have obtained therapists’ ratings of their patients’ psychopathology using the HSCL (Kass et al., 1980), we were unable to find studies that used this strategy for general medical providers. It is appealing to consider using similar instruments for patient and clinician ratings of outcomes. This strategy merits further exploration for the MOS depression tracer.

The Global Assessment Scale (GAS)

The GAS was developed by Endicott et al. (1976) for inclusion in the SADS (Table 9.3). The GAS is designed to provide an overall assessment of psychological functioning on a continuum from psychiatric illness to health. In contrast to most other global assessment measures, the GAS gives clear definitions for ten levels of illness severity. It is designed to be completed by clinicians or clinician-researchers. The scale ranges from 1 to 100, with 100 representing perfect health. The scale assesses the worst level of social functioning and/or psychological symptoms in the last week. Written descriptions anchor the scale at ten-point intervals.

Inter-rater reliability was between 0.61 and 0.91 in four studies. Endicott et al. (1976) determined the relation of the GAS to other global assessments of severity, self-reported symptoms, and the likelihood of rehospitalization, as well as the sensitivity of GAS scores to changes in clinical status. GAS ratings by research personnel showed moderate correlations (0.2 to 0.5) in the expected direction with self-assessments of mood and other observer ratings of global impairment. Former psychiatric inpatients with a GAS score of less than 40 had a significantly higher rate of rehospitalization. Moreover, the GAS was a highly sensitive indicator of change in clinical status.

The GAS appears to be most appropriate for evaluations of patients with primary mental disorders who have no functional limitations from medical disease, and for use by raters who have a clear understanding of the symptoms of mental disorders. The measure does not include a description of symptoms of mental illness (although examples of common symptoms are given), and the measure does not specify that impairment in social or personal functioning should be due specifically to the mental disorder. As a result, the GAS would probably have a lower reliability if used by nonpsychiatric physicians in general practice and would probably be sensitive to impairment from physical disease.

The relatively high degree of specification of scale points of the GAS may serve as a model for global assessments by the treating physician in the MOS. However, such assessments would be more useful for the MOS if they elicited information on several specific dimensions of
### Table 9.3

**SOCIAL ADJUSTMENT SCALE SELF-REPORT (SAS-SR): ITEM CONTENT AND ORGANIZATION**

<table>
<thead>
<tr>
<th>Qualitative Category</th>
<th>Role Area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Qualitative Category</td>
</tr>
<tr>
<td>Behavior performance</td>
<td>Time lost</td>
</tr>
<tr>
<td></td>
<td>Impaired performance</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpersonal</td>
<td>Reticence</td>
</tr>
<tr>
<td>Relationships</td>
<td>Hypersensitive behavior</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Friction</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Social discomfort</td>
</tr>
<tr>
<td></td>
<td>Guilt</td>
</tr>
<tr>
<td></td>
<td>Worry</td>
</tr>
<tr>
<td></td>
<td>Resentment</td>
</tr>
<tr>
<td></td>
<td>Lack of interest in heterosexual relationshipsd</td>
</tr>
</tbody>
</table>

**SOURCE:** Adapted from Weisman and Bothwell (1976).

*Includes sections for work outside the home, work at home, and schoolwork for students.
*bIncludes parents, siblings, in-laws, children not living at home.
*cAlso includes persons of opposite sex living together in a permanent relationship.
*dFor single, separated, or divorced persons not living with a person of the opposite sex.
functioning (e.g., severity of depression, impairment in social functioning), rather than combining these dimensions.

IMPAIRMENT IN SOCIAL FUNCTIONING

Measures of social functioning include measures of social support (presumably an independent variable) and social adjustment (a dependent variable). Social support measures typically concern the amount of support received from friends, family. For example, the Social Support Network Inventory used by Flaherty et al. (1983) assesses the availability, emotional support, practical support, and specific-event-related support received from the respondent's closest friends and family members. By contrast, the Social Adjustment Scale Self-Report (SAS-SR) (described below) assesses change in social functioning and the quality of relationships and functioning. In fact, measures of social support and social adjustment contain similar items and measure similar dimensions of social functioning.

While the MOS investigators intend to include generic measures of social well-being, it is important to note that there is a specific history of studying social supports and functioning in relation to mental disorders and depression. Reviews of the subject have been conducted by Weissman (1975) and Weissman, Shalomskas, and John (1981), Beattie and Stevenson (1984) and Kane and Kane (1984) among others. The Social Adjustment Scale-Self-Report (SAS-SR), developed by Weissman and Bothwell (1976), is an example of one of the more carefully studied social adjustment scales. It is a 42-item measure of feelings and behavior in various social settings, including occupational, marital, and family functioning; social and leisure activities; and economic independence. The complete SAS-SR takes 15 to 20 minutes to complete. Factor scores can be used to represent six dimensions of adjustment (work, social and leisure, extended family, marital, parent, family unit; see Table 9.3). The total score reflects overall social adjustment. The correlations between the self-report version and an interviewer-administered version is between 0.40 and 0.76 for the separate factors and 0.72 for the overall adjustment score. The internal consistency is 0.74.

Moreover, the SAS-SR is sensitive to changes in depression. In a study of 76 depressed patients, significant decreases in total and factor scores were found between the acute depression phase and four weeks after treatment. The SAS-SR has been widely used in psychiatric and nonpsychiatric populations (Weissman, Shalomskas, and John, 1981). Because the MOS investigators are developing measures of social
functioning for all tracers in the MOS, these measures are not discussed in detail in this report.

SUMMARY AND IMPLICATIONS FOR MOS

In the MOS, measures of depressive symptoms are required for first-stage screening and as an indicator of severity of depression.

Several alternative self-report and observer measures of depressive symptoms exist. Of the measures reviewed, the CES-D and GHQ appear to be the most suitable for purposes of first-stage screening. In one study, the GHQ had a sensitivity of 100 percent for major depression (RDC). While the CES-D has a lower sensitivity, it is shorter and may be less threatening to medical outpatients. We are also examining the usefulness of the MHI for first stage screening. Regardless of the measure used for screening, the MOS design will include the administration of second-stage screening (the DIS) to a random sample of those persons who are negative on first-stage screening. This will allow an estimate of the false negative rate.

Several investigators have expressed concern about the validity of patient self-reports of depressive symptoms for those who are acutely depressed. As a result, the MOS design will include the HRS, an observer assessment of the severity anxiety and depression. In addition, we will include a self-report measure of several dimensions of psychological distress, including depression. Likely candidates include the BDI (depression) and the MHI. The major depression tracer will also include a measure of social functioning, e.g., the SAS-SR.
Chapter 10

SELECTION OF COVARIATES

Because the MOS is a naturalistic study rather than a randomized trial, the patients in one provider group or institutional setting could have different characteristics from those in another. One objective of the MOS is to explore the degree to which differences in process or outcome between groups may be associated with differences in selected measured characteristics. This can be done either through the sampling procedure (by stratifying on certain variables) or by statistical control in multiple regression analyses. Further, the inclusion of covariates predictive of the dependent variable enhances the precision of multiple regression estimates of the effects of a factor (e.g., system of care) on the process and outcome of care.

The two salient features of useful covariates for the MOS are that (1) the covariates should be predictive of the dependent variable (outcome, process), and (2) the covariates should reflect important differences between patient groups of interest (e.g., patients of fee-for-service or prepaid settings).

This chapter discusses the relevance of nine potential covariates for the depression tracer:

1. Characteristics of the depressive disorder
2. History of treatment for mental disorders
3. Coexisting mental disorders
4. Coexisting physical disease
5. Social supports
6. Psychosocial stressors
7. Attitudes toward mental disorders and treatments
8. Demographic characteristics of patients
9. Health insurance coverage

CHARACTERISTICS OF THE DEPRESSIVE DISORDER

Most of the relevant characteristics of the depressive illnesses have been described in previous sections. These were (a) the severity of the major depressive episode, and (b) other clinical features of the depressive episode: psychotic symptoms, suicidal ideation; “endogenous”
features; anxiety; characteristics of the episode; and preexisting chronic depression.

Severity of Depression

Controlling for the severity of depression is important for two reasons. The patients of mental health specialists may have more severe illnesses than do those of general medical providers (Brown, Regier, and Balter 1979). However, this conclusion has been based largely on studies of differences in the diagnoses assigned to patients by their providers. Such an assessment strategy will lead to a biased estimate if one provider group is more prone to minimize mental disorders than another. As a result, it is not clear whether patients of general medical physicians differ in severity from those of mental health specialists.

Using data from the Rand Health Insurance Study (HIS), Wells et al. (1982) and Ware et al. (1984) found no large differences in general mental health status between outpatient users of mental health specialists and general medical providers, after controlling for physical status, insurance coverage, and demographic characteristics. However, Wells et al. (1982) did not study a patient sample with a specific mental disorder, such as a major depression. Under these conditions, there may be significant differences between provider groups in the severity of illness.

The second reason for controlling for differences in severity of illness is that the initial level of depression has been shown to be significantly predictive of outcome. For example, Rounsaville, Weissman, and Prusoff (1981) found that the most important predictor of the outcome of treatment for depression was a “prognostic index” developed by Auerback, Luborsky, and Johnson (1972). The index combined clinicians’ judgments (after the initial treatment session) of the patient’s aptitude for therapy, emotional freedom, acute depression, general emotional health, and intellectual achievement. Three of these (emotional freedom, acute depression, and general emotional health) appear to represent the pretreatment level of psychological health.

Other Clinical Features of the Depressive Episode

The presence of delusions suggests a poor short-term prognosis and probably a poor long-term prognosis. Fuller detail is provided in Chap. 4. The presence of suicidal ideation is important because depressed patients who verbalize it may be more likely to actually commit suicide. It may be desirable to limit the MOS tracer condition to patients who are not at high risk for suicide (see Chap. 5).
The presence of so-called "endogenous" features (melancholia) suggests a good response to tricyclic antidepressant medication. There is some evidence that symptoms of anxiety, by contrast, predict a poorer response to tricyclics and may be an indication for the use of an MAO inhibitor. (See Chap. 4 for fuller detail.)

Two characteristics of the episode are of prognostic importance: the duration and the number of prior episodes. There is some evidence that the longer treatment is delayed, the poorer is the response to treatment. Further, a history of multiple episodes, especially if there is persistent depressed mood between episodes, indicates a poor prognosis and the need for prophylactic antidepressants or lithium. Frexisting chronic depression suggests a poorer prognosis. Chapter 3 has presented fuller details on these conclusions.

TREATMENT HISTORY

Patients in the tracer sample should vary considerably in the extent of past treatment for the tracer condition or other mental disorders. This history may have implications for the course of illness and the process of care. Providers should tend to choose management strategies that have worked in the past; e.g., medications, alleviation of job stress. Information on treatment history may be available in the database collected with administration of the DIS. However, patients may be less explicitly aware of past mental health care received from general medical providers because such care may be disguised as physical health care (e.g., antidepressants for "stomach problems").

Differences in treatment history for the current episode (identified at enrollment) are especially salient for this study. Many may have already received some or most of their care for this episode. Thus a pretreatment measure (e.g., of anxiety or depression) may not be possible. This makes it difficult to evaluate changes in health status. In particular, the study may be biased toward showing no change in health status with treatment. This may be especially true for patients who receive care from general medical providers. Using HIS data, Wells et al. (1982) demonstrated that patients who visit general medical providers for their mental health care have, on the average, one or two visits a year that include mental health care. By contrast, patients of mental health specialists have an average of ten such visits annually. Thus patients treated for depression by general medical providers may already have received all their mental health care for the year at enrollment to the study. As a result, it may be advisable for the MOS to oversample (or only include) patients who have received very little care
for the tracer condition—that is, to focus on recently diagnosed or undiagnosed cases. Alternatively, the MOS sample could consist of patients who have a current episode of no more than a specified time (e.g., three months), regardless of treatment.

COEXISTING MENTAL DISORDERS

The DSM-III and the RDC both exclude the diagnosis of major depression when it has been preceded by certain mental disorders or symptoms. However, we will not be applying the full set of exclusion criteria as specified in DMS-III. As a result, some of the patients enrolled in the tracer may have a preemptive mental disorder. Further, some patients will have other coexisting mental conditions or will develop another condition during the course of the study. The disorders most likely to coexist with a depressive disorder are an anxiety disorder or a substance-abuse disorder; they are common and are specifically associated with the depressive disorders. Further, these disorders complicate the management of depression. For example, a patient with major depression may develop alcoholism, which has the direct effect of increasing the depression. In the MOS depression tracer, we will include self-report measures of alcohol use and of symptoms of anxiety. In addition, we will use interviewer-administered screeners for anxiety, alcohol-related disorders, and psychotic symptoms based on sections of the DIS.

COEXISTING PHYSICAL DISEASE

Because the MOS will enroll patients in medical outpatient settings, we anticipate that many of the enrollees in the depression tracer may have coexisting physical illnesses. For some patients, a physical illness may be the cause of the depressive episode (e.g., organic brain disorder with major depression, endocrine disorders). For others, the depressive episode may represent a psychological adjustment to a physical disorder. Other patients may have unrelated diseases.

The MOS must control for coterminous physical diseases because the presence of a physical disorder may have implications for the management or course of the depressive episode. For example, psychotropic medications are contraindicated for certain physical problems. Treatment of major depression accompanied by a disabling physical problem may be ineffective unless the physical disorder is also managed effectively. Symptoms of the physical disease may complicate
the physician's assessment of depression, causing a delay in treatment of the depression (or vice versa).

Another reason why the MOS must control for physical diseases is that they may affect patients' decisions about where to receive mental health care. Wells et al. (1982) and Ware et al. (1984) demonstrated that physical disease is an important determinant of the choice of provider for outpatient mental health care. Patient populations of mental health specialists and general medical providers differ in their physical health status. Specifically, other factors being equal, users with physical limitations are more likely to visit mental health specialists.

The MOS will control for physical diseases in several ways. First, we may exclude from the depression tracer patients who have physical disorders that severely limit functioning. Second, we will include a measure of cognitive dysfunction, such as the Mini Mental Status Exam in our enrollment and exit assessments. We can use scores on such a measure to exclude patients with severe cognitive dysfunction and to stratify patients by level of cognitive dysfunction. Third, the health status assessments in the MOS will include comprehensive assessments of physical health, including patient-reported chronic illnesses, symptoms, physical capacities and functioning, and provider-reported diagnoses. We will use this information to derive physical functioning and disease indexes for use as covariates in our analyses. Because these measures are needed for all the MOS tracers, they are not discussed in this report.

The MOS will also develop strategies to assess symptoms of psychological distress and/or mental disorders for patients enrolled in the medical tracers. These strategies are not discussed in this report.

SOCIAL SUPPORTS

Most of the empirical studies have focused on the role of social supports as a risk factor for the development of depression, rather than on its prognostic significance for a given episode. Nevertheless, it is probably important to include a measure of social supports when controlling for case-mix differences in the MOS. (See Chaps. 6 and 9 for fuller detail.)

PSYCHOSOCIAL STRESSORS

A large body of research has addressed the role of psychosocial stressors in the genesis of depression. As discussed in Chap. 1, the presence of a precipitating psychosocial stressor was formerly thought to be
associated with a less severe depression. However, recent studies have not found meaningful differences in depressive disorders that are or are not preceded by such stressors (Paykel, 1978; Brown, Regier, and Belter, 1979). Stressful events, especially the death of a spouse, have been demonstrated to be significantly associated with the development of depression, other psychiatric disorders, and physical disorders (Paykel, 1982b). For example, Paykel (1978) estimated that there is a six-fold increase in the relative risk of developing depression within six months of the most stressful events.

Social supports are hypothesized to mitigate the negative effects of stressful life events (Miller, Ingham, and Davidson, 1976). Others believe that life events and social support have independent effects on health (Andrews et al., 1978).

Little is understood about the effects of life events on the course of illness once a depressive disorder has developed. For example, chronic stressful circumstances, such as chronic physical illness, may both contribute to the development of depression and complicate the treatment of depression (Brown and Harris, 1978). However, this effect of life events has not been rigorously studied. Because of their possible etiologic significance and their relation to both social supports and depression, a measure of life events may be useful as a covariate for the depression tracer. We plan to include a measure of life events, based on that used for the Rand Health Insurance Study.

ATTITUDES TOWARD MENTAL HEALTH CARE AND PROVIDERS

Even if patients of mental health specialists and general medical physicians do not differ in case-mix, in terms of subtypes of depression or severity, these groups may differ in their orientation toward mental health care. Patients of mental health specialists may be more familiar with psychological concepts, more receptive toward psychotherapy, and more likely to accept a diagnosis of a mental disorder and the treatment offered. If such differences occur, then patients of mental health specialists could have better outcomes of treatment because of their more favorable attitudes (by higher motivation) and/or compliance with treatment. We plan to include a measure of attitudes toward mental health treatments and providers, based on a measure used in the Yale-New Haven ECA site.

DEMOGRAPHIC VARIABLES

Wells et al. (1982) found that users who receive their mental health
care from mental health specialists differ, on the average, from those who receive their care from general medical providers in sex and age mix, education, and other demographic characteristics. For example, when they receive mental health care, women and children are more likely than are men to visit a general medical provider (i.e., controlling for other patient characteristics). Schurman, Kramer, and Mitchell (1985) found similar effects of demographic factors in an analysis of visits made to private-practice psychiatrist and nonpsychiatrist physicians. Horgan (1985) found similar results in an analysis of the National Medical Care Expenditure Survey (NMCES).

Some demographic characteristics may also be related to the process and outcome of treatment. However, most studies have not controlled for initial patient characteristics such as health status. For example, in a review of the relation of client variables and outcome of psychotherapy, Garfield (1978) concluded that individuals of lower socioeconomic status had shorter lengths of stay in psychotherapy. However, Lorion (1974) concluded that socioeconomic status was not related to outcome. Garfield (1978) identified four studies that found a significant positive correlation between educational status and outcome of psychotherapy. However, these studies did not consistently control for initial health status.

Studies of the relation of age and sex to psychotherapy outcome have found mixed results (Garfield, 1978). However, there appears to be a significant positive correlation between intellectual achievement and psychotherapy outcome (Luborsky et al., 1971).

It may be especially important for the MOS depression tracer to include study site as a covariate. Wells et al. (1982), and Schurman, Kramer, and Mitchell (1985) found large differences by site in the relative use of mental health specialists and general medical providers for mental health care. Some of these regional differences were explained by differences in the availability of mental health specialists. Because of this marked variation, it may behoove the MOS investigators to choose a fairly large number of sites, to enhance generalizability, or to sample enough patients and providers in a few sites to allow an estimation of specific site effects.

**INSURANCE COVERAGE**

One may expect that those with more generous health insurance coverage will be more likely to have care from the more expensive mental health specialists and to receive more care, given any use. Wells et al. (1982) demonstrated that variation in cost-sharing affects the prob-
ability of having outpatient mental health care, but not the choice of
provider or the intensity of care from mental health specialists. How-
ever, these conclusions applied when coverage was assigned. In the
MOS, patients will have selected their own coverage. Those who
expect to use a mental health specialist may select more generous cov-
erage. As a result of this self-selection, we may see more use of psychi-
atrists and more intensive care, given any use, for those with more gen-
erous health insurance coverage.

One problem with including measures of insurance coverage as
covariates is that plans differ widely in their cost-sharing provisions
(e.g., deductibles, upper limits, coinsurance rates). It is difficult to
reduce variation in generosity to a small set of variables for use in mul-
tivariate analysis. However, specific MOS staff are developing mea-
sures of health insurance coverage for use in the MOS.
Chapter 11

ETHICAL ISSUES FOR THE DEPRESSION TRACER

Two major ethical issues apply to the MOS depression tracer: how to protect the confidentiality of the information obtained from patients, and whether patients and physicians should be informed when the MOS researchers make a diagnosis of major depression. Confidentiality is an important issue for all tracers, and will not be discussed in detail here. It should be noted, however, that patients may be more sensitive about the information that they provide for evaluation of depression than for evaluation of physical disorders. As a result, special efforts may be necessary for the depression tracer to assure patients and providers that data will be analyzed anonymously.

The main ethical issue for discussion in this chapter is whether the MOS should inform either the patient or the physician when a definite diagnosis of depression is made during the study assessments. Most studies of the detection of depression by general medical providers have failed to discuss this issue and have not included information on whether they provided diagnostic feedback to providers or patients (Kass et al., 1978; Raft et al., 1977; Nielsen and Williams, 1980; Tollefson et al., 1983; Cavanaugh, 1983; Schottstaedt, Deckert, and Schneider, 1971). Epidemiologic studies conducted in general populations have not provided such feedback. In part, the need for such specific feedback in epidemiologic studies is reduced by the patient's informed consent and the patient's freedom to seek or not seek treatment during the course of the study.

Even if the MOS were to provide specific feedback to patients or providers, it is not clear that this would have much of an effect on the process or outcome of care.

Previous studies of the effect of providing feedback to providers have yielded conflicting results. Johnstone and Goldberg (1976) and Hooper et al. (1984) found very little effect of feedback in medical outpatient settings. However, Linn and Yager (1980, 1982) found higher rates of diagnoses of depression and higher total medical costs after such feedback.

The issue of providing diagnostic feedback in the MOS is related to an ethical issue faced in the design of randomized controlled clinical
trials. If treatment is effective, it may be unethical to assign persons to control groups that provide little or no treatment. A variety of strategies have been used to address this concern, including using a waiting-list control group, providing some minimal treatment, or considering psychological assessment as minimal treatment. DiMascio et al. (1979b) described the solutions to this ethical dilemma developed for the Boston-New Haven Collaborative Depression Project. First, they excluded patients with suicidal ideation or psychosis from the study. Second, prior to randomization, patients were informed that they might be assigned to a minimal treatment group. Third, their control group was informed of the availability of a therapist on an “as needed” basis. Fourth, they reevaluated the control group at eight weeks. Patients who had not improved or who were worse at this assessment could be referred for specific treatment. Patients were also informed of this safety feature of the control group.

In the MOS, which is a naturalistic study, the ethical conflict appears to fall somewhere between that of an epidemiologic study and a randomized clinical trial. Like the randomized trial, the patient has presented to the health care system for treatment. While the patient may not be specifically seeking mental health evaluation, most patients want to know what is wrong with them and receive appropriate treatment. However, unlike the randomized trial, the study is not assuming control of the treatment. Rather, patients and providers are free to make their own decisions about care. In this respect, the MOS is like an epidemiologic study. In fact, the ethical dilemma for the MOS may be less complex than for an epidemiologic study, because each patient is known to be in contact with a provider who could potentially offer treatment.

Several alternatives could be included in the MOS design to resolve this ethical concern. Because this concern also applies to other tracer conditions, the MOS investigations should carefully consider these alternatives as generic design issues for the MOS. The first solution is to inform patients and providers in advance of the tracer conditions included in the study. The participants could also be informed that any patient included in the final sample will definitely have one of these conditions, but other patients not included may also have one of these conditions. Such an explanation not only fulfills the requirements of informed consent, but protects the patient by cueing providers to evaluate the patients selected for the study.

The second alternative is to evaluate the severity of depression for the depression tracer patients at a specified time, such as three to six months after enrollment. Those patients who are worse or who have
failed to improve could be contacted, or their physicians could be contacted, or both.

The third alternative would be to conduct a randomized trial of diagnostic feedback within the MOS. Depression tracer patients could be randomly assigned to either immediate feedback or no feedback. In this design, the severity of depression would be reassessed at six months after enrollment. If the group that received feedback demonstrated significant improvement relative to the group that received no feedback, all patients and/or their providers would be informed of the diagnosis of depression. A variant of this design would be to assign patients either to immediate feedback or to delayed feedback (e.g., six months after enrollment).

A fourth alternative would be to give patients the option at enrollment of having the diagnostic feedback sent to them and/or their provider. However, this design might compromise the main MOS analysis by introducing an experimental effect (feedback) that may be confounded with patient characteristics, such as severity of illness. The randomization in the third alternative would minimize this possibility.

Finally, the patient group could be restricted to nonsuicidal, nonpsychotic patients, as in the Boston-New Haven Collaborative Depression Project. One advantage of this restriction is that the risk of mortality in the patient sample decreases, lessening the ethical concern. However, this restriction on the range of severity would also affect the generalizability of study results.

In considering these alternatives, it is important to emphasize that the ethical obligations of a naturalistic study, such as the MOS, are not as well understood as those of a randomized trial. The freedom of the patient and provider to choose or refuse treatment, and the fact that all patients will have identified providers, substantially reduces the study's obligation to provide information on diagnosis, especially if the conditions of informed consent are fully met. Nevertheless, the second and third alternatives discussed above should be considered; excluding persons with a significant risk of suicide should also be an option.
REFERENCES


Steinbrueck, S. M., S. E. Maxwell, and G. S. Howard, "A Meta-Analysis of Psychotherapy and Drug Therapy in the Treatment of


