Corbishley, Maureen Ann

REM DENSITY, REM LATENCY AND THE DEXAMETHASONE SUPPRESSION TEST AS PREDICTORS OF TREATMENT RESPONSE IN DEPRESSED OLDER ADULTS

The University of Arizona PH.D. 1987

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark __✓__.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages ✓
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Dissertation contains pages with print at a slant, filmed as received ✓
16. Other _____

_____

_____

University
Microfilms
International

REM DENSITY, REM LATENCY AND THE DEXAMETHASONE SUPPRESSION TEST AS PREDICTORS OF TREATMENT RESPONSE IN DEPRESSED OLDER ADULTS

by

Maureen Ann Corbishley

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF COUNSELING AND GUIDANCE

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 8 7

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Maureen Ann Corbishley

entitled REM density, REM latency, and the dexamethasone suppression test as predictors of treatment response in depressed adults

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Richard L. Spitzer	8 Dec. 1986 Date
L. Eberenthal	12/8/86 Date
Elizabeth B. York	12/8/86 Date
Philip J. Ladner	12/8/86 Date
	Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

L. Eberenthal	12/8/86
Dissertation Director	Date

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under the rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: M.A. Corbishley

ACKNOWLEDGMENTS

Although completing a dissertation is essentially a solo task, it would not be possible without a great deal of assistance, and I am deeply appreciative of the many kinds of support I have received during my years of working on this project.

I could always rely on Dr Phil Lauver for thought-provoking questions about the purpose of my study; on Dr Dick Erickson for concerns about how the parts fit together; on Dr Reed Mencke for a faith in my ability to perform autonomously.

I am grateful to Dr Elizabeth Yost for stepping into the breach at the last moment as a committee member, and for her many valuable comments on the manuscript.

Dr Larry Beutler earns my gratitude for his stimulating and enthusiastic approach to the subject of this dissertation. It was he who provided the original spark and introduced me to a hitherto unknown area of psychology.

Special thanks go to Keith Meredith for his help with statistics. Not only was he generous with his time and unfailingly patient throughout many analyses, but he

also taught me to enjoy the process of managing and understanding data, and of conducting scientific research.

Ultimately, my chief debt of gratitude goes to my family of origin, to whom this dissertation is dedicated: to my parents, whose sacrifices and determination provided me with opportunities to develop skills and confidence, and whose own pursuit of truth has been tireless; to my younger brothers and sisters, who generously took on extra burdens in order to allow me years of higher education; to those members of my extended family who, whatever the hardships of daily life, modeled enjoyment and pride in learning for its own sake.

TABLE OF CONTENTS

TABLE OF CONTENTS...Continued

LIST OF TABLES

ABSTRACT

The purpose of this study was to investigate whether biological variables could predict how older adults would respond to different types of treatment for depression. Fifty-six adults over the age of sixty-five, diagnosed with major depression (DSM III criteria) were assigned to one of four treatment conditions: group Cognitive Behavior Therapy with alprazolam or placebo medication, and minimal support therapy with alprazolam or placebo medication. Before and after treatment, REM latency and REM density were measured by polysomnograph in the sleep laboratory and the Dexamethasone Suppression Test (DST) was administered. Depression was measured weekly by the Hamilton Depression rating Scale and the Beck Depression Inventory. Subjects presented with normal (i.e nondepressed) values on REM latency and REM density. Thirty five percent of subjects were DST nonsuppressors, a similar percentage to that found in other studies of depressed subjects, but mean DST for the whole group was below the selected cutoff of 4 mcg/dl. The expected correlations among the biological variables and between

these variables and baseline depression levels were not found. It was concluded, therefore, that depression in this group of subjects was not characterized by biological abnormalities. Multiple regression analyses of baseline variables and depression scores at mid and end of treatment and at followup indicated that initial depression levels and DST predicted later depression levels for subjects who received Cognitive Behavior Therapy, regardless of medication assignment. Low baseline DST levels were associated with good response to psychotherapy, confirming the findings of previous studies. Sleep variables were not predictive of response to treatment at any time point.

CHAPTER ONE

RATIONALE AND REVIEW OF LITERATURE

This chapter begins with a brief overview of the subject area treated by the current study, in order to provide a context for later discussion. There follows a review of the literature on various aspects of the study, such as prevalence of depression, assessment of depression, the search for biological correlates of depressive disorders and the current status of research in these areas. The chapter concludes with the research questions and hypotheses that were derived from the literature and that guided the present study.

Depression is a highly prevalent disorder with serious consequences if left inadequately treated. In the elderly, the incidence and costs of depression are higher than in any other age group. While several effective treatments for depression exist, none is effective with all patients or with every type of depression. Therefore, clinicians must choose among treatments in order to select the combinations that they feel will best suit a particular patient. The most important prerequisite for this matching of patient to treatment is accurate assessment of the

patient and his or her depression. To date, differential diagnosis based on self reports, clinical judgement and observer ratings has been somewhat crude, and conflicting evidence has arisen concerning the patient characteristics and depressive symptomatology that best predict response to various types of treatment. In the case of older patients, the diagnosis of depression is further complicated by factors associated with the aging process and with concomitant physical ailments, but it is in this group that the most severe effects of untreated depression are often seen, thus increasing the need for accurate diagnosis and treatment selection. Recently, attention has turned to biological variables that might prove to be markers for depression and aid in more precise classification of depressive syndromes and in the assignment of depressed patients to effective treatments. It was the purpose of this study to investigate three biological markers for depression in a population of patients over the age of 65, in order to determine the extent to which these markers might help predict response to pharmacological and psychotherapeutic treatments for depression.

Incidence and costs of depression in the elderly

In the general population, depression is considered to be the most common diagnosis rendered in primary medical practice, followed by hypertension (Katon, 1984). Reports

of the incidence of depression produce widely varying figures, depending on the diagnostic criteria employed, the severity levels and types of depression included in the study, and the population sampled. However, an adequate picture of the overall occurrence of depression can be obtained by looking at results of some of the recent well-designed epidemiological studies in this area. For example, researchers drew random samples from households in three geographical locations (Myers, et al., 1984). The investigators conducted nine thousand interviews over a six month period, using the Schedule for Affective Disorders and Schizophrenia, developed by Endicott and Spitzer in 1978 as a standardized method of applying their Research Diagnostic Criteria (Spitzer, Endicott & Robins, 1978). They found major depressive disorder in 2.2-3.5% of the sample, dysthymia in up to 3.8%, and major depression, explained by participants as due to bereavement, in up to .3%. These figures represent a considerable amount of depression when one takes into account that the samples studied were not part of a patient population. Rates of depression are higher when the subjects are patients. For example, Nielson & Williams (1980), applying Research Diagnostic Criteria to ambulatory patients in a prepaid health plan, concluded that 5.5% of the patients were suffering from RDC major and 12.1% from RDC minor depression, and furthermore, that their physicians failed to detect 50% of these cases of

depression. Quite similar results were obtained by Hoeper, Nyczy, and Cleary (1979). Using Research Diagnostic Criteria, they found a 5.8% rate of depression in primary care patients. Thus depression rates increase from approximately 2.2% in the general population to 5.8% in a patient population, when the same criteria are used for measuring depression.

When depression is measured by self report instruments, such as the Zung Self Rating Depression Scale and the Beck Depression Inventory, rates are much higher (from 26-56%), partly because these measures tap mood disorders from the severe to the very mild, both longstanding and transient (Katon, 1984), and partly because instruments such as the BDI count symptoms of depression, rather than syndromes. Raskin (1979) reports that older adults endorse many more symptoms of depression than do younger people, but without necessarily meeting the requirements for a major depressive disorder. Higher rates are also found among hospitalized medical patients, where depression has been estimated to occur in 22-33% of cases.

Despite the actual prevalence of depression as described by epidemiologists, the disorder is frequently overlooked or misdiagnosed. In a typical study, depression was recognized by medical staff in only 4% of those patients who were actually depressed (Moffic & Paykel, 1975). The difficulty physicians have in diagnosing depression is

reflected in inappropriate treatment. Katon (1984), for example, indicates that chronic pain patients reporting symptoms of depression have been treated with minor tranquilizers or with multiple surgeries, rather than with treatment appropriate to depression.

It is difficult to determine accurate incidence rates of depression in older adults because inconsistency in diagnostic criteria, measurement, and sampling are compounded by factors which affect the presentation of depression in the elderly and make assessment of the problem more complicated than in a younger population. Gurland, Dean, Cross & Golden, (1980) investigated a random sample of people over 65 in London and New York. Using multiple criteria for depression, they found that 13% of the sample suffered from depression serious enough to warrant medical attention. Highest rates were found at age 80 for men, 24% of whom showed signs of depression, and at age 65-69 for women, 21% of whom were depressed. Blazer and Williams (1980) found dysphoria in 14.7% of a group of elderly people, although only 3.7% of them had major depression as diagnosed by DSM III criteria. Okimoto, Barnes & Veith (1982) found much higher rates of serious depression, over 30%, but their sample was biased in that it was composed entirely of male V.A. patients, who might be expected to show more disturbance than a random community sample.

Even more so than is the case with younger adults, depression in the elderly often goes undiagnosed, untreated or inaccurately treated (Miller, 1978). The result is an emotional and financial drain on the resources of the older adults themselves, of their families and of the community at large. In a population rendered vulnerable by social conditions such as poverty and isolation and by increasing physical frailty, depression can be especially devastating. When the depressive syndrome includes self neglect, somatization of depressive symptoms or hypochondriasis, the response of the medical community is often to provide patients with inappropriate treatment, such as drugs and surgery which do not address the depression itself (Chaisson-Stewart, 1985). Others with undiagnosed or incorrectly treated depression, living alone and feeling unable to care for themselves, seek hospitalization or institutionalization that could probably be avoided if their depression were effectively treated. Depression has been found to be the most important cause of excessive alcohol and drug use among those over the age of 65 (Rosin and Glatt, 1980), and untreated depression can complicate physical illness, resulting in more extensive use of medical facilities and greater suffering for the patient (Weissman, 1983). The gravest cost of depression in the elderly is indicated by suicide rates, which are three times higher in those over 65 compared with the general population (National

Center for Health Statistics, 1977). The failure of clinicians to recognize depression in older patients is indicated by the fact that 76% of one sample of successful older suicides had visited a physician within a month of death, but had not been detected as seriously depressed (Miller, 1978).

Among older adults, therefore, a situation exists where the incidence of depression is higher than in the general population, the costs of untreated depression are also higher, and yet the condition is often overlooked or misdiagnosed, with the consequent failure to provide appropriate treatment. In the face of these facts, there is a clear need for improvement both in diagnosis and treatment, with specific reference to depressed patients over the age of 65.

Current treatments: effectiveness and drawbacks.

Once the basic diagnosis of depression has been made, further assessment of depressive subtypes is only relevant, clinically speaking, if there exist different treatment modalities to which patients respond differentially. In contrast to the situation only a few years ago, there now exist several quite separate treatments for depression, each of which provides a satisfactory level of effectiveness, though it is not yet clear which treatment is best suited to which patients. The three most commonly

used and best researched methods of treating the disorder are electroconvulsive therapy (ECT), antidepressant medication and psychotherapy. Since ECT was not a focus of the present study, we shall consider here only the psychotherapy and medication treatment options.

Considerable evidence has accumulated for the efficacy of various types of antidepressants. In a review of nineteen studies, Siris, Alexander and Stetner (1982) found that from 19-74% of depressed patients improved with tricyclic medication. Janicak et al. (1985) performed a meta analysis of rigorously controlled studies using different types of antidepressants, and found efficacy rates as follows: tricyclic antidepressants (TCA's) = 64.3%; monoamineoxidase inhibitors (MAOI's) = 32%; placebo medication = 37%.

Various types of psychotherapy have been applied to the problem of depression. According to a meta analysis conducted by Smith, Glass and Miller (1980), discrete and recognizable therapies such as behavioral, humanistic, and cognitive produce good results compared to placebo treatments or undifferentiated counseling approaches. Steinbrueck, Maxwell and Howard (1983), investigating treatments specifically for depression, concluded that cognitive, behavioral, marital and interpersonal therapies were equally effective, regardless of whether they were delivered to individuals or in groups. Studies comparing

Cognitive Behavior Therapy (CBT) and antidepressants reveal that, in a non-clinic setting, CBT is usually superior to drugs, whereas in clinic populations the most effective treatment is often drugs or a combination of drugs and psychotherapy (Rush, 1983; Miller and Berman, 1983).

Since both types of treatment, that is, psychotherapy and antidepressant medication, have been found effective for older depressed patients (Mintz, Steuer and Jarvik, 1981; Gallagher and Thompson, 1982; Cole, 1983) it might appear that the question of treatments for depression has been largely solved. However, problems still exist, in that neither of these treatments is effective with all patients or with all types of depression and neither is free of side effects or other drawbacks.

The side effects of antidepressant medication can be cardiac rhythm disturbances, ataxia, memory and learning problems, confusion and slowed reaction time (Branconnier and Cole, 1981). Antidepressants can also interact negatively with other medications, especially those regulating cardiac function (Ayuso-Guttierez, 1983). Since older adults are often already concerned about memory and attention changes occurring as part of the normal aging process, and also frequently suffer from cardiac problems and take many different types of medication, these side effects of an incontestably effective treatment can produce

both subjective disturbance and serious medical consequences in those over 65 (Blazer, 1982).

In the light of these difficulties, psychotherapy might seem by contrast to be a wiser choice of treatment for older patients, but it, too, has its drawbacks. Therapists tend to view older people as unsuited to therapy (Ford and Sbordone, 1980) and often have difficulty in conducting therapy with older adults because of private concerns related to their own parents and to thoughts of their own aging and death (Altrocchi, 1980). Therapies specifically developed for the treatment of depression are also relatively new. It can be difficult, therefore, to find therapists who are both trained in techniques for dealing with depression and willing to work with the elderly. A further drawback of psychotherapy derives from the fact that it tends to have a later onset of action than drugs, so that symptomatic relief is delayed and changes in interpersonal functioning might not be apparent for as long as six months (Weissman, 1983). For many older adults, who already feel that time is short, this delay could be a discouraging factor; for those currently suicidal, the delayed action could be fatal. In addition, psychotherapy is expensive and not always supported by medical insurance, a factor that can be decisive for people living on fixed and limited incomes.

In view of the multiplicity of treatments and the drawbacks inherent in each, the clinician is faced with the

need to select for each depressed patient not only the treatment likely to be most effective but also the one with the fewest risks. At all ages, matching of patient to treatment is important to reduce suffering and waste of effort. In the case of older adults, who usually have less time, money, energy and hope than younger people, and who tend to pay a heavier price in terms of side effects for any treatment, the issue of treatment selection becomes more critical. In a few cases, the issue is not difficult to resolve. Kendall (1981), for example, recommends the use of ECT, because of its rapid action, when the patient's life is endangered by starvation, self mutilation or repeated high suicide risk. It is also obvious that psychotherapy would not be attempted with a patient in whom depression had produced such confusion or stupor that attention could not be focussed on the verbal interaction required by psychotherapy. Nor would one administer antidepressants to a person with a known allergic reaction to these medications.

Thus, when clear indications or contraindications exist, or if the patient has a determined bias against a particular treatment, the choice of interventions is less complicated. Many times, however, the clinician is in a position where it is necessary to decide among several available methods for treating depression, knowing that the range of effectiveness could be as wide as from 19-74%, and

that the "wrong" choice could lead to prolonged, even increased suffering. The choice is likely to become more complicated in the future, since promising new treatments under investigation for depression, such as REM deprivation (Vogel, McAbee, Barker & Thurmond, 1977) and alteration in circadian rhythms (Duncan, Gillin, Post, Gerner & Wehr, 1980), will add to the number of choices available. The selection of an effective treatment for an individual patient depends above all on accurate knowledge of the particular symptoms, history and other factors that are acted upon by different treatments. As we shall see in the discussion of the search for treatment predictors, there is broad clinical and research support for the assignment of endogenously depressed patients to biological treatments and exogenously depressed ones to psychotherapy, but knowledge in this area is still quite general, lacking the precision to support individual treatment decisions.

Assessment of depression in the elderly

It can be problematic, even in younger patients, to differentiate depression from other conditions and to discriminate the different subtypes of depression. In older people, diagnosis of the basic depressive syndrome and assignment to a specific diagnostic category can be far more difficult. First, depression is often overlooked in this population because it is mistaken for other conditions. The

symptoms associated with depression are similar to those occurring as a result of both the normal aging process and medical illnesses prevalent in older people. For example, decreased functioning, memory impairment and poor judgement can be features of both depression and organic brain disorders; weight loss and insomnia occur in depression and also in many physical illnesses; fatigue and changes in sleep patterns are common to both depression and normal old age (Gallagher, Thompson & Levy, 1980). Symptoms such as confusion, depressed affect and irritability can arise from a genuine depressive disorder but also from polymedication or unexpected drug interaction (Whitlock, 1982). Since conditions such as brain disorders and polypharmacy occur more frequently in older patients, the chance is increased that depression will be either overlooked or misinterpreted in this group.

Additional problems arise from our lack of understanding of how depression is manifested in older adults. Freedman, Bucci & Elkawitz (1982) suggest that there may be a subtype of depression specific to this age group, characterized by feelings of meaninglessness and reduced metabolic function. As yet, there exist no diagnostic criteria for such a subclassification. A further difficulty arises from the fact that most research has been conducted with younger patients, but results from these studies do not necessarily generalize to older adults.

A third complication in the assessment of depression in the elderly is their own response to depressive symptomatology. Since many of them are not well physically, they can easily and quite genuinely confuse their depression with their medical condition. This confusion is aided by a lack of psychological sophistication which inclines them to view life in physical rather than in psychological terms, focussing on readily identifiable aches and pains rather than on cognitive or emotional distress. Some older adults, perhaps ashamed of being thought mentally ill, perhaps trying to avoid causing problems for family members, deny depression (Barnes, Veith and Raskind, 1981).

Because of these problems in the assessment of depression in older adults, coupled perhaps with the caregiver's unfamiliarity with conditions of life for many older patients, it is easy for clinicians to assign such patients to inappropriate treatment. It is not uncommon, for example, to find that depression in an older population is induced by malnutrition, in which case the treatment should be improvement in diet rather than any of the standard treatments for depression (Chaisson-Stewart, 1985). Sometimes, to give another example, the many physical complaints of elderly depressed patients are interpreted as vegetative signs commonly associated with depression, an interpretation that can lead to an erroneous diagnosis of endogenous depression and the unwarranted use

of ECT or antidepressants. Conversely, an equally erroneous diagnosis of exogenous depression could be made, with subsequent failure to provide biological treatment, if the physician misinterprets the patient's complaint as hypochondriacal, or attributes depressed mood solely to an external event such as loss of a spouse. In the elderly, therefore, a diagnostic problem that is difficult enough in younger patients is even more complicated, and the likelihood of making an appropriate treatment selection is diminished.

It is apparent, therefore, that there is a need for reliable information about which patients, with which types of depression, respond to which treatments, a need, in other words, for accurate assessment and prediction. Unfortunately, it is here that a major difficulty arises: decades of research, mostly based on symptomatology and history of depression, have not yet resulted in accurate methods of predicting treatment response. In fact, intensive and sophisticated efforts have revealed depression as a more complex disorder than was previously thought, and the question of treatment selection has become, if anything, more difficult than before.

The early search for treatment predictors

Historically, clinical diagnosis alone was the basis on which to make treatment choice, a basis that, while it

might produce accurate generalizations, could be very misleading in individual cases. Despite every effort at accuracy of assessment and the reasonable expectation, from previous experience, that the treatment would be successful, there were always patients who responded badly or not at all, and others, not expected to recover, who responded well. It was apparent that more efficient predictors of treatment response were needed, and the search was conducted in two directions. On the one hand, the observation that depression presented in different symptom patterns led to the hope that more accurate classification of subtypes of depression would increase treatment prediction. On the other hand, researchers who noticed that certain biological anomalies seemed to characterize depression undertook a vigorous investigation of the physiological correlates of depression.

Earliest efforts directed at the classification of subtypes resulted in almost immediate disagreements: one school saw a single disease entity that varied only in severity (Mapother, 1926; Lewis, 1938), while another endorsed a bipartite division into an endogenous (sometimes also called psychotic) type and an exogenous (reactive or neurotic) type (Gillespie, 1929). Both schools used symptom patterns to differentiate their categories (Fowles and Gersh, 1979). With the widespread use of ECT and the advent of antidepressant medication in the late 50s and 60s, the

dualist position was strengthened by clinical observations that patients with endogenous depression, believed to be the result of a biological malfunction, responded better to biological interventions when compared with patients with exogenous depression, where etiology was supposedly psychogenic (Kuhn, 1970; Kiloh, Ball and Garside, 1962).

In reviewing studies concerning response to antidepressant medication, Bielski and Friedel (1976) concluded that the value of the endogenous-exogenous distinction for predicting treatment response was upheld, but only in general terms. Many contradictions and inconsistencies exist among these studies. Kiloh et al. (1962) found that many of the features traditionally associated with exogenous depression did indeed predict poor response to TCA`s, with two exceptions. Subjective retardation and depth of depression, formerly considered to be hallmarks of endogenous depression, predicted poor response to antidepressant medication. Several double blind studies (Paykel, 1972; Raskin and Crook, 1976) found no difference in response to antidepressants between the two types of depression. Fowles and Gersh (1979), in reviewing these and other studies conclude tentatively that the majority of studies support the dualist position, but that the endogenous-exogenous distinction is only an "approximation of an adequate classification system" (p.76). Much of the confusion over the two types of

depression may have arisen because many studies did not use homogeneous patient groups and included schizophrenic as well as bipolar disorders in the endogenous samples. Additionally, different sets of symptoms were chosen in different studies to identify patients in the two categories, and studies often lacked placebo controls and double blind administration of medication, or failed to take into account relevant factors such as age, severity of depression or the presence of delusions. It is difficult, therefore, to evaluate these early studies or compare them effectively.

By the mid 1970s the endogenous-exogenous distinction was established as a valid but not very precise diagnostic method for predicting treatment response. The dichotomization was crude and failed to account adequately or consistently for numerous factors affecting treatment response, such as anxiety, agitation, severity and length of illness, depressive cognitions, etc.

Attempts to differentiate patients and types of depression by means of physiological activities parallelled the efforts to predict treatment response on the basis of clinical presentation. Funkenstein, Greenblatt and Solomon (1952) measured blood pressure response of depressed patients to methacholine and adrenaline probes. He reported that 93% of subjects with high basal blood pressure which took a long time to return to baseline after the probe

responded well to ECT. However, intensive investigation of the Funkenstein test over the next decade produced such confusing and contradictory results that it was abandoned (Hamilton, 1982).

In 1958, Shagass and Jones claimed that a test to measure sedation threshold could differentiate endogenous from exogenous depression. However, the reliability of the test was called into question (Nymgaard, 1959) and modifications of the test failed to differentiate schizophrenics from depressives (Perez-Reyes, 1968). The Shagass test and its modifications were ultimately recognized as measures of anxiety rather than of specifically depressive symptoms or subtypes (Hamilton, 1982).

Other physiological correlates of depression were also under investigation during this period, but although they did not lead to the dead ends of the Funkenstein and Shagass tests, they did not yield sufficiently accurate or reliable results as to be useful in predicting treatment response.

Recent developments in diagnostic criteria

Recognizing the many difficulties that had arisen during the previous decades, researchers in the last 15-20 years have focussed efforts on the development of stringent criteria for designating subtypes of depression, the

exploration of promising physiological correlates and the improvement of research methodology. Precisely defined and operationalized criteria for the diagnosis of depression, regardless of presumed etiology, were presented as Research Diagnostic Criteria (RDC) by Spitzer et al. in 1978. According to these criteria, Major Depressive Disorder is characterized as follows:

(A) A distinct period of dysphoric mood or pervasive loss of interest or pleasure.

(B) At least five of the following symptoms:

1. Change in appetite or weight.
2. Changes in patterns or amount of sleep.
3. Loss of energy, tiredness.
4. Objective psychomotor retardation or agitation.
5. Loss of interest or pleasure in usual activities.
6. Self reproach or guilt.
7. Diminished capacity to think or make decisions.
8. Suicidal thoughts or actions.

(C) Duration of at least one week.

(D) Impaired functioning at work, home etc, or sought help or was referred for help with depression.

(E) Has no symptoms suggesting schizophrenia.

(adapted from Spitzer et al, 1978, p. 776).

These criteria were incorporated into the Diagnostic and Statistical Manual of Mental Disorders, third edition (DSM III) in 1980. The DSM III avoids such dichotomies as

neurotic vs. psychotic or endogenous vs. reactive. Full blown depression is classed as a major depressive episode, either single or recurrent, with subclassifications comprising the presence or absence of either psychotic features or melancholia. The authors maintain that symptoms of melancholia, in addition to other features of depression, correspond to the traditional diagnosis of endogenous depression. As can be seen below, the revisions from RDC to DSM III are quite minor, with the exception of the requirement that the dysphoric mood have lasted at least two weeks. This criterion alone makes DSM III more stringent than the earlier RDC.

(A) Dysphoric mood or loss of interest or pleasure in all or almost all usual activities and pastimes. The mood disturbance must be prominent and relatively persistent, but not necessarily the most dominant symptom.

(B) At least four of the following symptoms have been present nearly every day for a period of at least two weeks.

1. Changes in weight (when not dieting) or appetite.
2. Changes in amount or patterns of sleep.
3. Observable psychomotor agitation or retardation.
4. Loss of interest or pleasure in activities or decrease in sexual drive.
5. Fatigue or loss of energy.
6. Feelings of worthlessness or inappropriate guilt.

7. Complaints or evidence of decreased ability to think, concentrate or make decisions.

8. Recurrent suicidal ideation or attempts.

(adapted from DSM III (1980), pp. 213-214).

The use of these two sets of criteria has been widespread in clinical practice and in research, and their value in predicting treatment response has been upheld in some cases. For example, Prusoff, Weissman, Klerman and Rounsaville (1980) found RDC criteria for major depression of the endogenous and situational subtypes reasonably accurate in predicting response to amitriptyline and short term interpersonal therapy. Patients whose depression was classified as endogenous responded poorly to psychotherapy and well to pharmacotherapy, those not meeting criteria for endogenous depression responded best to psychotherapy alone, and patients with both endogenous and situational features to their depression responded best to a combination of psychotherapy and medication. An interesting finding of this study was that endogenously depressed patients responded worse to psychotherapy than those in the control group, underscoring the need for accurate diagnosis and treatment selection.

Apart from the positive results obtained by Prusoff et al. (1980), the overall picture that has emerged from further research is more complex and confusing than before. Far from confirming and clarifying the endogenous-exogenous

distinction, several researchers have cast doubt on criteria previously accepted as positive indicators of one type or the other and have also posited the existence of several additional subtypes of depression. Nelson and Charney (1981), examining the most recent studies on depression, suggested that a renaming of the two basic types as "autonomous" and "responsive" would more accurately reflect the difference between the categories. In autonomous depression, the patient's mood does not change despite changes in the environment, whereas a patient in a responsive depression will experience corollary mood shifts when something happens in the environment that the person construes as pleasant. Furthermore, Nelson and Charney concluded that there might exist two distinct autonomous states, agitated delusional and retarded anhedonic, each characterized by distinctive symptom patterns. The former is thought to respond well to tricyclic antidepressants but worsens with antipsychotic medication, while the latter type is thought to respond poorly to tricyclics alone and better to a combination of antidepressants and antipsychotics.

When they examined depressive symptomatology as described in recent studies, Nelson and Charney found that self report measures were not as useful in distinguishing autonomous from responsive depression as two symptoms observed by others, i.e. lack of reactivity and psychomotor change. They also noted that changes in sleeping, eating

and weight, long accepted as symptoms reflecting endogenous depression, also occur in nonendogenous syndromes. Similarly, self report of early morning awakening, which correlates by tradition with endogeneity, was not confirmed by EEG sleep recordings or by nurses' observations. Thus, their examination of recent literature threw doubt on not only the basic differentiating feature (reactivity to the environment rather than precipitating event) between the two major subtypes, but also on specific symptoms associated with each type.

In a 1984 study, Kupfer and Frank investigated the relationship among dichotomous subtypes diagnosed by three different sets of criteria: RDC endogenous-nonendogenous; DSM III melancholic-nonmelancholic; Nelson-Charney autonomous-nonautonomous. The three systems of classification were not interchangeable, despite considerable overlap, and subjects categorized by the three diagnostic systems produced quite different polysomnogram (PSG) records. For example, REM latency differentiated the RDC but not the DSM III groups, and DSM III melancholic responded poorly to amitriptyline, unlike the RDC endogenous, even though presumably the classification systems were targeting the same people. The use of three systems to classify the same population and the differences in biological measures among groups confirms that as yet no set of clinical criteria exists that can reliably identify

subtypes of depression, especially where treatment prediction is concerned.

In other studies, Nelson, Charney and Quinlan (1981) noted that among inpatients meeting DSM III criteria for major affective disorder, between 39-50% were responsive to psychosocial intervention, thus casting doubt on the assumption that patients with this diagnosis respond best to somatic treatments and poorly if at all to psychotherapy. A similar observation was made by Gallagher and Thompson (1982) who noted that some, but not all, patients with both types of disorders respond to both types of treatments. Further complications in the diagnostic picture have been introduced by researchers who maintain that the core symptom of depression is retardation, previously considered as only one of many possible symptoms and certainly not essential (e.g. Widlocher, 1983). Clouding of the picture has also arisen as a result of efforts to understand the place and significance in depressive subtypes of psychosis and anxiety.

Despite the limitations, the use of specific diagnostic criteria is an improvement over previous imprecise approaches to the categorizing of depression in that these criteria achieve some measure of precision and provide a language by which researchers and clinicians can communicate with a greater degree of accuracy and consensus. However, Carroll, Feinberg and Greden (1980)

maintain that diagnostic criteria, for all their specificity, have not proved superior to clinical judgement in diagnosis of subtypes of depression. Hamilton (1982) summarizes the current position with the rather lukewarm statement that "research on prediction has given some results and also revealed the complexities of the task" (p.25).

Since virtually all of the foregoing research has been conducted with younger patients, one can only assume that diagnostic criteria would perform even less satisfactorily with older people, given the increased number of assessment problems that occur in the latter population. In spite of intensive efforts to improve treatment selection by sharpening diagnosis, the situation is still that clinical factors and symptomatology are not yet a precise enough tool for reliable treatment selection in individual cases.

Biological markers of depression - recent developments.

Since the serendipitous discovery by Kuhn in 1957 of the antidepressant qualities of imipramine, the rapidly growing field of psychopharmacology has stimulated considerable research into possible biological correlates of depression and into the development of diagnostic and prognostic laboratory tests. In the course of this search, almost every bodily system has been investigated, with

results indicating that multiple systems are interactively involved in depressive disorders. Rubin and Marder (1983) have described several of the tests resulting from research in this area. For example, underactivity of the neurotransmitters serotonin and norepinephrine can be measured in the urine, with different levels of transmitter said to distinguish among bipolar, unipolar and schizoaffective depressions. EEG sleep recordings (polysomnographs or PSG's) supposedly can differentiate normal from several subtypes of depressed patients. Similar differentiation has been shown by measurement of the hypersecretion of ACTH and cortisol that often accompanies depression. Whybrow, Akiskal and McKinney (1984) report other biological anomalies that can occur in depression. Compared with normals, for example, depressed patients, while awake and asleep, show increased respiration, heart rate, muscle tension and sodium levels, together with PSG's of increased complexity and desynchronization. Apparently, depression reflects a disturbance in the regulatory mechanisms of the central nervous system, producing a condition of high arousal and disorganization in the brain. This being so, it is reasonable to assume that there may indeed be specific biological markers for depression, that is, measurable indicators of disease, whether those indicators are pathognomic or causal.

Research into biological markers for depression is in its beginnings, and there is no belief that biological tests can replace clinical diagnosis at this time. However, Schildkraut et al. (1983) maintain that the present state of the art is such that laboratory tests can be used as an adjunct to clinical diagnosis, either to confirm the presence of depression in doubtful cases or to specify a subtype of depression. Such adjunct diagnostic aids would increase the likelihood of selecting appropriate treatment, without resorting to trial and error, currently a necessary approach when more precise information is lacking. Paykel (1979) considers that, eventually, biological tests might prove more accurate than clinical diagnosis as treatment predictors, since they can bypass existing doubtful categories and provide a more direct form of assessment. There are also emerging indications that biologically based tests have the potential to determine the effectiveness of treatment within a few days of initiation and can indicate when it is appropriate to discontinue treatment (Rush, 1983).

In order to be considered effective as a diagnostic aid in depression, a biological test or variable might be expected to perform to some level of satisfaction in the following areas: distinguishing the psychiatrically well from the depressed, and those who are depressed from those with other types of psychiatric illness; clarifying the

problem of depressive subtypes, particularly as they relate to treatment; predicting treatment response, evaluating progress, and indicating whether or not to terminate treatment. To date, three biological conditions have provided good performances in many of these areas, although, naturally, there are limitations and exceptions. These variables, believed by some to be markers for depression, are cortisol secretion, as measured by the Dexamethasone Suppression Test (DST), and two sleep variables, namely Rapid Eye Movement (REM) latency and Rapid Eye Movement (REM) density. As the subject of the present study, these correlates of depression will be discussed in detail.

Cortisol secretion and DST

Cortisol levels in the blood follow a circadian rhythm that produces low peak levels in the early hours of sleep and high levels on awakening (Whybrow et al., 1984). In many depressed patients, cortisol secretion is abnormally high throughout the 24 hour period, but especially during sleep. However, 45% of nondepressed people also show elevated cortisol levels, so a direct measure of serum cortisol is not sufficient to differentiate normal from depressed people. An effective method of producing such differentiation is the dexamethasone suppression test (DST). One milligram of dexamethasone, administered orally, imitates the action of cortisol. The body, deceived into

believing that adequate cortisol is circulating in the blood stream, suppresses further production of natural cortisol for at least 24 hours. This is termed "normal suppression". Early research suggested that, in endogenously depressed patients, the body resisted or escaped from this normal suppression between 16 and 24 hours after the dexamethasone was administered, and produced increased quantities of cortisol despite the presence of the imitator. The term for this escape or resistance is "non-suppression", a state indicated by abnormally high cortisol levels (over 4 micro-dliters) within 24 hours of dexamethasone administration.

Research into the DST has been extensive, partly because of early promises that it might prove to be a specific diagnostic test for melancholia and partly because it has the advantages for clinical use of being inexpensive and easy to administer, requiring no washout period or special preparations (Carroll et al., 1981). Although the reliability of the DST in patient populations has received almost no attention, Brown and Qualls (1981) found that 10 out of 11 patients (both suppressors and nonsuppressors) showed the same DST results on two occasions several months apart. Since all the patients were depressed on both occasions, it would seem that test reliability might be high, at least in the depressive state.

In discussing a test's ability to differentiate between clinically meaningful groups, one must consider the sensitivity and specificity of the procedure. High specificity indicates that the test can correctly identify most of those subjects who do not suffer from the condition under examination. High sensitivity indicates the ability to accurately determine those who do suffer from the condition. When the DST is used to differentiate normal from depressed patients, specificity is consistently high in that, on average, 96% of non depressed people show normal dexamethasone suppression. In one sample, Carroll (1982) found only 3 out of 70 (4%) normals with abnormal suppression. In normals aged 23 to 50, only 4.3% failed to respond normally to the DST (Rush, 1983). In a sample of depressed elderly adults compared with age matched healthy controls, specificity was also 96% (Georgotas et al., 1986). Tourigny-Rivard, Raskind and Rivard (1981) report 5% non-suppression in a normal population of elderly subjects. Exceptions to these consistent results are the findings of Amsterdam, Winokur, Caroff & Conn (1982) that 15.1% of normal healthy controls compared to outpatients with major depression evinced failure to suppress cortisol in response to dexamethasone.

Specificity is somewhat limited by the findings (Rush, 1983) that infections, allergies and some medications (e.g. barbiturates and anticonvulsants) can produce abnormal

DST in nondepressed patients, indicating that patients with abnormal DST should be screened for these factors. Although specificity might be increased by the use of standardized laboratory and administration procedures, it is also possible that reduced specificity indicates that the DST reflects a generalized physiological malfunctioning which is not confined to depression. This suspicion is heightened by the prevalence of abnormal DST suppression in patients with severe weight loss, as in bulimia or anorexia nervosa (Roy-Byrne, 1982). Since decreased appetite and weight loss are often associated with endogenous depression, it is possible that the DST correlates more with biological systems related to appetite than with depression itself.

Sensitivity (the ability to identify all those who are in fact depressed) is considerably lower than specificity, ranging from 24% to 100%, with an average of 45%. In other words, about half of all depressed people have normal responses to the DST, with abnormal responses occurring in fewer than half of depressed patients. Recent findings indicate that this number is seasonally affected. Arato, Rihmer and Szadoczky (1986) found that in winter 37% of patients with unipolar depression exhibited DST nonsuppression, whereas in summer this figure was 56%, significantly different from winter rates. Carroll (1982) explains the low sensitivity on the grounds of different test procedures and variations in clinical criteria used to

select samples of depressed patients. His belief is that "the DST is a specific episode-related biological marker of melancholia" (p.293) and that an abnormal DST is an indicator for somatic treatment.

Another possible explanation of low sensitivity is that existing diagnostic criteria are not yet sufficiently refined to identify truly endogenously depressed patients, and therefore studies have inadvertently included nonendogenous depressives, who account for the normal suppression rates. Support for this explanation comes from a recent study by Zimmerman, Coryell and Pfohl (1986), who used DST results in a sample of 187 primary depressives as a dependent variable in the examination of psychosocial correlates of suppression and nonsuppression. There were significant differences between suppressors and nonsuppressors on eight variables typically associated with the endogenous-nonendogenous distinction.

As is the case with most biological markers, many investigations of the DST have been nosological, with the intention of testing the validity of existing categories of endogenous and nonendogenous depression. There is reason to expect that abnormal suppression would indicate endogeneity, since, in general, patients who present with vegetative signs of depression (traditionally indicative of an endogenous mood disorder) also exhibit an abnormal response to DST (Whybrow, et al., 1984). Differentiation among

subtypes has again proved problematic because of the unreliability of the criteria used to categorize patients' depression. Different rates of suppression and nonsuppression occur, depending on the diagnostic system used e.g. the RDC or the DSM III. However, specificity has been high, ranging from 76% to 90% (Coryell, Gaffney, & Burkhardt, 1982; Rush, Giles, Roffwarg & Parker, 1982; Carroll, et al., 1981). The DSM III category of melancholia has provided the highest correlation with abnormal DST results, with 19 out of 20 patients diagnosed as melancholic showing abnormal suppression (Carroll et al., 1980). It seems that, on the whole, DST nonsuppression identifies a common group of severely depressed patients, though there still remain some apparently endogenously depressed patients who exhibit normal suppression and cannot be clinically differentiated from abnormal suppressors (Carroll, 1982).

Despite these positive results regarding endogeneity, Coryell (1984) warns that this group of melancholics is quite small relative to the large number of patients seen in general practice as opposed to research settings which focus on affective disorders. When prevalence of the disease is low, the value of a test as a screening device is considerably reduced (Griner, Mayewshi, Mushland & Greenland, 1981). The DST is not therefore recommended as a screening device in a population where the

likelihood of endogenous depression is low. Its value lies more in its ability to confirm suspicions of melancholia.

In contrast to the generally satisfactory pattern of results from studies reported above are the findings of Shulman (1980) based on a sample of 24 inpatients. In this study, the DST did not differentiate endogenous from nonendogenous depression, and the highest nonsuppression levels of cortisol were found among patients who were both nonendogenously depressed and also responded well to psychotherapy without drug treatment. Shulman also reports that in a sample of 100 inpatients, the DST was unable to discriminate among several subtypes of depression.

In an attempt to avoid problems due to the use of diagnostic systems such as RDC and DSM III, a study was made of chronic pain patients, all of whom had many of the vegetative signs of depression, but not all of whom were, in fact, depressed (Krishnan et al., 1985). All of the subjects completed three symptom scales for depression and anxiety, the researchers wishing to discover whether the DST responded mainly to vegetative symptoms (in which case it would not discriminate depressed and nondepressed pain patients) or whether it did indeed respond to other symptoms characteristic of major depression. Should the DST prove successful at differentiating the two groups, the researchers also wished to discover any symptom patterns associated with both suppression and failure to suppress.

Results showed that, as would be expected from previous studies, nondepressed patients showed normal suppression, and 40% of the depressed subjects failed to suppress normally. The authors then investigated the different symptoms reported by depressed suppressors and nonsuppressors and found that a discriminant function analysis derived from these symptom profiles correctly classified over 92% of all subjects, and 100% of depressed subjects. Symptoms associated with nonsuppression were: reported sadness, inner tension, reduced appetite, suicidal ideation, loss of libido, autonomic symptoms, depressed mood, anxious behavior at interview, hypochondriasis, paranoid symptoms, reduced interest, loss of insight, weight gain, concentration and sleep difficulties. Many of these symptoms are traditionally used as clinical indicators of endogenous depression. The authors suggest that similar studies be conducted to develop symptom profiles that are derived from a measurable biological basis and can be used to enhance diagnosis and treatment selection.

The question of differentiation between depression and other psychiatric conditions on the basis of DST has received a somewhat disappointing response over recent years. Abnormal suppression appears not to be specific to depression, since it has also been found in patients with diagnoses of bulimia (Hudson et al., 1983), schizophrenia (Castro, Lemaire, Toscano-Aguilar, & Herschuelz, 1983) and

alcoholism (Oxenkrug, 1978), among others. However, it has been suggested that presence of abnormal suppression might be accounted for by undiagnosed depression in these patients (Coryell, 1984).

Several studies have investigated the usefulness of DST in predicting treatment response among depressed subjects. In 1980, in contradiction of Shulman's study of that same year, Brown and Shuey found that 82% of endogenously depressed patients with abnormal DST's responded well to tricyclic medication whereas a good response occurred in only 37% of those whose DST was normal, even though their depression was classified, clinically, as endogenous. Brown, Haier and Qualls (1980) hypothesize that, since elevated cortisol levels are related to a deficiency in norepinephrine (NE), DST nonsuppressors should respond to antidepressants whose action is directed at the increase of NE. They further hypothesize that normal suppressors who are depressed are deficient in serotonin, rather than in NE, and should respond, therefore, to antidepressants intended to increase production of the neurotransmitter serotonin. In a study designed to test these hypotheses, nineteen inpatients, aged from 24-61, with a diagnosis of major depressive disorder according to RDC criteria, were randomly assigned to two drugs affecting NE levels and two drugs affecting serotonin levels. Eight of the patients (42%) were DST nonsuppressors, while eleven

(58%) had normal DST reactions. Results of the study confirmed the hypotheses, with nonsuppressors responding to NE but not serotonin drugs, and normal suppressors responding to serotonin but not NE medications.

Unfortunately, efforts by Greden et al., (1981) to replicate these findings failed to do so. They selected 26 inpatients with a DSM III diagnosis of melancholia and with abnormal DST's, 19 of whom were treated with imipramine (a NE precursor) and 7 with amitriptyline (a serotonin precursor). A control group was not used. There were no significant baseline differences between the two groups, and it was hypothesized that the imipramine group would show better response. However, results showed no statistical differences between the groups, although there was a trend for the imipramine group to fare slightly worse than the amitriptyline group (62% versus 69% improvement). It was disappointing to be unable to fine tune prediction of treatment response, but results were still encouraging in that over 65% of all patients responded well to treatment, thus supporting the belief that depressed patients with abnormal DST's can be helped with antidepressant medication.

A small study by Cobbin, Cairncross, Jurd, Veltman and Pohlen, (1981) confirmed the hypotheses of Brown & Qualls (1981) when they found that patients with normal DST responded well to mianserin, a serotonin enhancer, while abnormal suppressors responded poorly.

In a 1983 study, Brown & Brawley assigned depressed patients to two types of antidepressant on the basis of their response to DST and a pharmacological probe found to be highly correlated with DST. At the start of the study, clinical judgement, diagnostic criteria, behavioral observations and standard instruments to measure depression were inadequate to differentiate suppressors and nonsuppressors. Without the DST results, there were, therefore, no grounds for assigning patients to different treatments or for predicting results. The researchers found that dexamethasone nonsuppressors responded well to amitriptyline, while suppressors responded well to imipramine. These results are important in two ways. First, they encourage the hope that the DST may be sensitive enough to predict specific drug assignment; second, they imply that the DST has the potential to outperform traditional methods of assessment and thus aid in treatment selection.

To summarize the current position on DST as an aid to diagnosis and treatment prediction in depression one can conclude on the positive side as follows: on the whole people who are not depressed, who suffer from a psychiatric condition other than depression, or whose depression is not endogenous, will show a normal response to the DST; the test has shown an ability to differentiate types of depression and correctly assign patients to effective treatments; the

DST is a useful adjunct to, and in some cases is more effective than, traditional methods of assessing depression and determining treatment assignment. On the negative side, a few studies directly contradict other findings, and seem to indicate that the phenomenon of cortisol suppression may be involved, in some patients at least, in syndromes other than just depression. Buchsbaum & Haier (1983) suggest that the DST might eventually be used to identify a single group of patients suffering from hypothalamic-pituitary-adrenal axis disease that is characterized by a great variety of clinical symptoms. However, despite limitations and contradictions, the DST has produced enough positive evidence of its potential usefulness to justify further investigation in yet untested areas. Biological measures tested on adult subjects do not always perform as expected in either young or old people. It is important, therefore, to examine the performance of DST in older patients in order to establish its validity as a marker for the diagnosis and treatment of depression in this population.

Sleep and depression

Literature throughout the ages has borne witness to disturbances of sleep in depression, but scientific investigation of the subject was not feasible until the advent of technology to measure sleep by

electroencephalogram (EEG) in 1946. Since Rechtschaffen and Kales (1968) published the standardized manual for sleep recording and scoring, it has been possible to conduct sleep studies with a high level of precision and comparability.

Prior to a consideration of the relationship of sleep variables to depression, some discussion is required of the difference in sleep between normal and depressed people, and the changes in sleep that occur with age. Over a decade ago, Williams, Karacan & Hursch (1974) provided descriptions of and normal values for sleep in healthy individuals at different ages. The accepted standard in the field, these definitions are summarized as follows. As measured by separate and distinct PSG patterns, sleep is divided into two major types, each accompanied by its different physiological characteristics. Non rapid eye movement sleep (NREM) consists of stages 1,2,3,4, the latter two also called delta or deep sleep. During NREM sleep, muscle tone is slightly decreased from waking levels and there are no movements of the eyes. Heart rate and blood pressure are lowered and breathing is slower and fuller. Brain temperature is also reduced as is the flow of blood to the brain. The second type of sleep, rapid eye movement (REM) sleep, has two phases. In tonic REM, blood flow to the brain and brain temperature increase, the body

fails to either heat or cool itself efficiently, breathing is shallow and muscle tone is so low that many muscles are almost paralyzed. Penile erections occur as long as there is no organic impairment. In phasic REM, breathing is irregular and accompanied by apneic events, muscles twitch, eyes move rapidly, blood pressure rises and heart rate becomes variable. In many ways, phasic REM sleep is so active as to resemble wakefulness. People wakened from REM sleep almost always report visual dreams.

REM and NREM sleep alternate throughout the night, in cycles lasting from about 80-110 minutes. In the first cycle, people pass through NREM stages one to four and then move into a short (as little as five minutes) period of REM sleep. This cycle repeats itself usually five times throughout the night, with some modifications: as the night progresses, delta sleep decreases and REM sleep increases. In the normal adult, stage 1 comprises 4%-5% of sleep time, stage 2, 45%-50%, stages 3 and 4, 20%-30% and REM 20%-25%. About 2% of the night is also spent awake.

With age, certain changes in sleep patterns occur naturally, most of them unfortunately in the direction of deterioration. The older one lives, the more reasons there are to view one's sleep as unsatisfactory (Hauri, 1982). For example, although total time asleep varies little from about 7.5 to 6.5 hours throughout adult life, sleep

efficiency (a measure of time in bed that is spent in sleep) drops rapidly after the age of 49, due to an increased number and duration of awake periods during the night. An unbroken night's sleep is unknown after age 50. Advancing age also brings more frequent stage shifts and an increased number of stage 1 sleep occurrences, which can be experienced subjectively as restless sleep. Delta sleep decreases with age until only 25% of people age 70-80 show any stage 3 or 4 sleep at all. Since this is the deepest sleep, this loss also results in subjective feelings that sleep is inadequate for true rest. Sleep in older adults is also likely to be further disturbed by apneas and nocturnal myoclonus, an involuntary twitching of the leg muscles that often results in partial or complete arousal (Quan, Bamford & Beutler, 1984).

Sleep disturbances in depression resemble somewhat the deterioration associated with age. Depressed sleep has been described as "shallow, short, inefficient" (Gillin et al., 1984). In other words, there are more sleep stage changes and awakenings, less total sleep time, more early morning awakenings, and less delta sleep. However, depressed sleep is also characterized by REM disturbances that do not typically occur, or at least not to so great an extent, in the sleep of older adults. In depression, REM sleep begins earlier (shortened REM latency) and in greater

than normal amounts at the start of sleep, and there is increased activity in the phasic stage of REM sleep proportionate to time spent in REM sleep (REM density) (Rubin and Marder, 1983).

The similarities between aging and depressed sleep, and the increased vulnerability of older persons to sleep disturbances, make it particularly difficult in the case of older adults to ascertain whether or not self-reported sleep problems are indicative of depression. The older person may, for example, have unrealistic expectations of sleeping at 70 the way he or she did at 30. Thus, distress at the inability to maintain sleep throughout the night could indicate only lack of knowledge of normal aging sleep patterns, and not the early morning awakening typical of depression in younger people (Quan et al., 1984). In addition, the elderly are particularly subject to sleep disturbances caused by medications, pain, respiratory and other diseases and an environment not conducive to sleep (ibid). The difficulties in assessing the sleep of older adults indicate both the importance of such an objective measure of sleep as EEG recordings and the need for reliable markers of depression in this age group.

REM variables in depression

The search for sleep variables that correlate with depression and could aid in diagnosis or treatment

assignment has focussed on REM latency (RL) and REM density (RD). The assumptions underlying the research have been that RL will be shortened in depression, supporting the hypothesis that depression is related to disturbances in circadian rhythms, and that RD will be elevated, in support of the hypothesis that depression is a state of physiological hyperarousal (Warburton, 1984). As with other biological markers, the questions have been whether or not these two sleep parameters can differentiate between nondepressed and depressed subjects, particularly at different ages, between depression and other psychiatric conditions, and among the various subtypes of depression, and whether or not REM variables can assist in treatment assignment and prediction of response. Because of the large number of studies in this area, we will first consider major research pertaining to younger subjects, that is, those under the age of 60.

In 1976, shortened RL (the number of minutes between sleep onset and the start of the first REM period) was described by Kupfer as a biological marker for primary endogenous depression, since short RL was at that time found in almost all patients with this diagnosis, and rarely in nondepressed controls or in patients with other psychiatric conditions (Coble, Kupfer and Shaw, 1981). In the normal, nondepressed population, the first REM period

consistently begins about one and a half hours after sleep onset, thus giving a normal REM latency of 70-90 minutes (Williams et al., 1974). In contrast, early studies of depressed patients typically revealed REM latencies ranging from 35-56 minutes (Kupfer and Foster, 1972, 1978).

Unfortunately, more recent studies have cast some doubt on these findings. Berger, Doerr, Lund, Bronisch, and von Zerssen (1982), for example, found that RL failed to distinguish healthy controls from depressed patients, and the mean RL for controls was 68 minutes, not very different from mean RL's previously found to characterize depressed subjects. In contrast to these results are the findings of Akiskal, Lemmi, Yerevanian, King and Belluomini (1982) that a cutoff point of 70 minutes for RL successfully discriminated primary depressives from healthy controls. Not only is 70 minutes within the normal RL range, but the results of this study are further weakened by the fact that the depressed patients were a decade older than the controls. Therefore, they could be expected to have somewhat lower RL than normals, since RL decreases with age in depressed patients compared with people who are not depressed (Ulrich, Shaw & Kupfer, 1980). A 1979 study by Gillin, Duncan, Pettigrew, Frankel and Snyder found that, although several sleep variables did allow a differentiation between normal, insomniac and depressed

subjects, RL and RD did not show the expected discriminatory power and were outperformed as discriminators by two other sleep variables, namely, amount of delta sleep and sleep continuity. Finally, Reynolds et al., (1980) found that in a small group of elderly depressed outpatients RL did not show the expected abnormal values. At the present time, therefore, the status of RL as a marker to discriminate healthy from depressed people is somewhat in doubt.

The sensitivity of REM latency as a single marker for depression is marred by the discovery of shortened REM latencies in patients with narcolepsy (Montplaisir, Billiard, & Takahashi, 1978), and also in patients in the process of withdrawal from drugs and alcohol (Watson, Hartmann and Schildkraut, 1972; Rundell, Lester, & Griffiths, 1972). One possible explanation, not investigated at the time, of these exceptional examples of low REM latency is that the narcoleptic and withdrawal patients were, in fact, also suffering from depression. Another explanation is provided by Coble et al. (1981) who discovered that RL can be affected in nondepressed people by conditions such as disruption of sleep and alteration of circadian rhythms, and suggest therefore, that shortened RL may not always be indicative of depression.

Since researchers are well aware of the prevalence of depression both in other psychiatric conditions and in states of medical illness, efforts also have been directed at establishing the ability of RL to differentiate primary from secondary depression. Many results in this area have been promising. Coble, Foster and Kupfer (1976) were able to able to discriminate primary from secondary depression among psychiatric inpatients by using RL and among psychiatric outpatients by using RL and RD. However, there was a significant age difference between the two groups. Foster, Kupfer, Coble, and McPartland (1976) took a group of primary depressives and attempted to discriminate them from a group of equally depressed patients who also suffered from medical illnesses producing central nervous system impairments. On a clinical basis, it was not possible to determine who belonged to which group, since symptomatology was identical. However, both RL and RD correctly selected group members. Among patients with severe primary depression REM latency averaged 42.2 minutes, and REM density was elevated to 2.0, compared with latencies averaging 71.8 and densities of 1.1 in medically depressed patients. Since the mean age of patients in this study was 41 years, it is not known whether similar differentiation would be found in older patients. The importance of this study is that the relatively noninvasive

sleep EEG procedures performed as well as far more extensive and complicated medical tests in classifying patients.

Kupfer, Foster, Coble, McPartland and Ulrich (1978) were also able to correctly classify primary and secondary depression in 77 out of 95 patients, on the basis of RL and RD. RD alone was also found to differentiate the two types of secondary depression (psychiatric versus medical), since RD was significantly higher in depressed subjects who had other psychiatric conditions than in those with concurrent medical illnesses. It should be noted, however, that differentiation might well have been helped by the fact that those in the secondary depression group were an average of ten years younger than those in the primary group, and would therefore be expected to show longer RL and lower RD than the older subjects.

In direct contradiction to the foregoing studies, which universally found short RL and high RD in primary as compared to secondary groups, Rush, Giles, Roffwarg and Parker (1982) found that primary depressives had longer RL and lower RD than secondary depressives. It is possible, however, that this anomaly was due to sample selection.

The question of whether or not REM variables can provide support for the validity of the classification of endogenous versus nonendogenous depression has received

largely positive answers. Rush et al. (1982), for example, found that RL but not RD was able to make the distinction between these two subtypes of depression, using a RL cutoff of 62 minutes. Similarly, Feinberg, Gillin, Carroll, Greden and Zis (1982) found that RL and RD correctly classified 75% of endogenous and nonendogenously depressed patients. More recently, by contrast, Kerkhofs, Hoffman, DeMartelaere, Linkowski and Mendelewicz (1985) found that RL and RD performed weakly in that they could only correctly classify 57% of patients from three groups: those with major depression, those with major depression in remission and those with minor depression. However, differences among the groups in baseline RL and RD were significant, supporting previous findings.

Since the primary purpose of diagnosis in a clinical population is to improve treatment, any biological marker must perform adequately as a predictor of response to treatment. Few studies have been conducted using sleep variables to predict response to treatment, and the only treatments so investigated have been a single antidepressant, a single anxiolytic and REM deprivation. Duncan et al. (1980) found that elevated REM density and reduced REM latency correctly identified depressed patients who responded to REM deprivation, while normal density and latency levels identified non responders. These results

imply that abnormal REM values are an indication of the need for treatment on the biological level. Kupfer et al. (1981) investigated response to amitriptyline in 34 depressed patients. Responders (as measured by the Hamilton Rating Scale for Depression had a baseline RL of 57.4, which was significantly longer than the RL of 39.2 in the nonresponders. Perhaps RL is capable of distinguishing a group of patients who have poor response to tricyclic medication. The same team (Coble et al., 1981) studied 22 inpatients with major depression over an average of 33 PSG recording nights per patient. Although all subjects received placebo medication throughout the study, clinical change did occur, as measured by the HDRS. Results indicated that a consistent RL of fewer than 20 minutes during the first week predicted poor response to placebo drug treatment. These poor responders were largely those subjects who carried a diagnosis of psychotic depression, were older (mean age 53, compared with 38 for the whole sample), and had higher initial depression levels than other members of the group.

In a more recent study, 52 subjects, aged 18-60, were selected on condition that baseline REM latencies were less than 65 minutes. Subjects were randomly assigned to amitriptyline or alprazolam. Results indicated that, the lower the baseline REM latency, the more effective was

amitriptyline compared with alprazolam, an anticipated result since the former is a true tricyclic antidepressant, while the latter is primarily an anxiolytic with presumed antidepressant qualities (Rush et al., 1985).

By contrast, in a study of differential response to amitriptyline in 18 inpatients, mean age 51, with a baseline REM latency of 41.45 and REM density of 2.03, researchers were unable to show that baseline sleep variables differentiated levels of response to treatment. However, by the end of three weeks treatment, the good responders showed significantly increased RL compared with nonresponders, thus confirming the implication of RL in depression (Kupfer, Foster, Reich, Thompson and Weiss, 1976).

Strong evidence for the value of RL in predicting treatment response comes from several studies where, although baseline REM variables did not differentiate eventual responders from nonresponders, the changes in these variables immediately after administration of amitriptyline correlated highly with clinical recovery. For example, after only two nights of medication, the two types of responders could be differentiated by RL and RD (Kupfer et al., 1976). In a different study, changes in RL, though not in RD, after the administration of medication were better predictors of treatment response

than any baseline clinical variables (Kupfer et al., 1981). In some subjects, therefore, baseline RL might appear normal, but there exists a sensitivity to antidepressant medication that can be detected by PSG recordings and is predictive of treatment response. It is not known, however, what differentiates this group of responders from those who demonstrate initially low RL.

In summary, it is apparent that short RL and high RD are intimately related to the phenomenon of depression, occurring most frequently in endogenous depressions unmixed with other medical or psychiatric conditions, normalizing significantly in response to antidepressants, and predicting (either at baseline or after a medication trial) response to TCA's. Reduced RL and increased RD may not always occur in all depression, especially of the nonendogenous type, but they appear to correlate more strongly with the severity of depression among endogenously depressed subjects. The hypothesis of some researchers that short RL and high RD might be a measure of organicity in depressed patients is particularly important when one considers older adults, in whom organic impairment is more likely to occur, making specific diagnosis of depression difficult (Foster et al., 1976). On the negative side, research thus far has shown many borderline results and contradictory findings, especially the fact that reduced RL

is a feature of conditions other than depression. The current status of sleep PSG variables as markers of depression is similar to that of the DST; as research efforts have intensified, doubts have been cast on originally optimistic and perhaps somewhat naive expectations for the biological variables. On the other hand, despite the complexity of the picture, research has strengthened the belief that these particular variables are an integral part of the depressive syndrome, and thus warrant further study.

Sleep variables and older adults

Because the process of aging brings about changes in sleep and increased variability in most biological responses, it is important that older adults be investigated separately from younger subjects. Thus, for example, there is normally a decline in REM activity and in total REM sleep time in the healthy aged compared with younger adults (Williams et al., 1974). In a study of 40 healthy subjects aged 58-82, RL averaged 54 minutes, a slight decline from younger norms. RD in this older group averaged 1.28 (Reynolds, Kupfer, Taska, Hoch, Sewitch & Spiker, 1985). It has also been reported that there is a steady, and much steeper, decline than normal in RL with age in depressed patients, with a range of from 97-37 minutes in early twenties to 46-09 minutes from age 51-60

(Kupfer, Foster et al., 1978). In order to accommodate these age related changes in sleep variables, Kupfer, Reynolds, Ulrich, Shaw and Coble (1982) investigated RL in older depressed adults and formulated the so-called "rule of 90". In older adults, if the sum of RL and age is less than 90, one can assume, with 95% specificity and 67% sensitivity, that the person is depressed, compared with age matched nondepressed subjects. Summarizing their experiences with older adults and RL, Kupfer and Thase (1983) suggest that a cutoff of 30 minutes to identify depression is more appropriate than the higher cutoffs found capable of discrimination in younger subjects.

According to investigations to date, REM sleep variables appear to have special diagnostic utility for older patients. Fifty percent of elderly depressed patients show cognitive impairment as part of their depressive symptomatology, and 10% are sufficiently impaired as to resemble patients with dementia (Folstein & McHugh, 1978). There is thus a likelihood of misdiagnosis of either depression or dementia in older patients. However, REM density has been shown to differentiate depressed older adults from those with similar symptoms who were in fact suffering from dementia. Using a cutoff of 1.60 for REM density and 30 minutes for REM latency, Reynolds, Spiker, Hanin, and Kupfer (1983) correctly

identified 62% of a sample of depressed patients and 78% of those with dementia. When the two variables were combined, sensitivity improved to 89%. These results were elaborated by Vitiello et al., (1984) who demonstrated that in mild dementia, sleep is not impaired and sleep EEG cannot therefore differentiate normal controls from mildly demented patients. These results imply that abnormal sleep parameters found in cognitively impaired patients are likely to indicate depression, not dementia. The results of Reynolds et al. (1983) were confirmed by Reynolds, Kupfer, Taska, Hoch, Spiker et al. (1985) who found that REM latency significantly differentiated depressed elderly (mean age, 69) patients from demented and normal cohorts. In this study, a REM latency cutoff of 30 minutes correctly identified 68% of the depressives. REM density was also investigated in this study, but failed to differentiate among the three groups.

Unfortunately, very few sleep studies have been conducted with subjects over the age of 60, and it is not possible to draw conclusions about the power of RL and RD to differentiate depression subtypes or to predict response to treatment. On the whole, however, the sleep variables appear to perform well as an adjunct to diagnosis in older adults, and even provide discrimination that would be impossible by any other means discovered so far. When

differential diagnosis is so complex in the elderly, it is encouraging to find a tool that apparently can be quite precise. Investigation of the sleep variables RL and RD seems warranted in the area of treatment response.

Biological markers in combination

Brain research over the past decade has indicated that the biological systems involved in psychiatric disorders are not only more complex than was originally thought but are also highly interactive. For this reason, Buchsbaum and Haier (1983) have suggested that a single biological marker is unlikely to be discovered that could perform satisfactorily as a diagnostic tool. Especially in light of the contradictory and limited results obtained with DST alone and RL or RD alone, it seems likely that combinations of markers would be more effective. Accordingly, several researchers have investigated the combined utility of the DST and sleep EEG measures for the purpose of classification and treatment response prediction. Since DST has the apparent power to identify nondepressed subjects and RL and RD performs well in classifying those who are endogenously depressed, one would reasonably expect the combination of the two biological measures to result in higher specificity and sensitivity than could be obtained with either alone.

Rush et al. (1982) obtained both expected and unexpected results when they attempted to classify depressive subtypes among 70 outpatients with a mean age of 36.5. Patients' depressions were classified according to RDC criteria as primary or secondary, endogenous or nonendogenous. Regarding the latter dichotomy, anticipated results were confirmed: DST showed high specificity and low sensitivity for endogenous depression, and RL the reverse. Discrimination was not perfect however, since both tests produced normal values for 8 patients diagnosed as endogenously depressed, and RD was unable to discriminate the two subtypes. In the case of primary and secondary depressions, results were the opposite of previous findings, with short RL and high RD in secondary depressions. Since many of the secondary depressives were also alcoholics, the authors hypothesize that depression secondary to alcoholism might be biologically similar to primary depression. Rush et al. (1982) suggest a series of test administration similar to that suggested by Carroll et al. (1980), that is, DST, followed by sleep EEG if the DST is normal. Data from their study indicated that this sequence of tests would identify 75% of primary depressives. The authors further theorized that, since an abnormal DST is always almost accompanied by changes in EEG parameters, but sleep changes are not always similarly

accompanied by abnormal DST, perhaps there is a biological progression in endogenous depression, with sleep changes occurring at the onset or when depression is relatively mild, and cortisol responses changing as depression worsens.

In a further study comparing alprazolam and amitriptyline (Rush et al., 1985), the researchers required a reduced RL of 65 minutes or less in order to enter the study, thus increasing the likelihood of endogeneity in patients studied. With this requirement, it is surprising that only 37% of the group showed abnormal DST results. Although amitriptyline proved more effective for all patients, abnormal DST results were associated with poorer response to both treatments, indicating perhaps that DST reflects a greater degree or a different type of disturbance. Mendlewicz, Kerkhofs, Hoffman and Linkowski (1984) examined REM latency, REM density and the DST in 39 depressed inpatients with an average age of 39, comparing them to a small group of healthy controls. The anticipated significant correlations were found in several areas: between RL, but not RD, and severity of depression; between abnormal DST and RL less than 60 minutes; between controls and depressed subjects on DST, RL and RD; between RL and age. The DST was 100% specific, that is, it correctly identified all normal subjects and at a cutoff of 50

minutes, RL correctly classified 87% of those who were depressed, thus confirming the expectation of higher specificity for DST and sensitivity for DST. REM density was not related to the DST, indicating that RD apparently measures a different physiological state from that measured by DST, although both states are related directly to depression. The classification performance of RD was not reported, nor were implications for treatment made by the authors of this study.

Feinberg and Carroll (1984) studied the DST and sleep variables in both inpatients and outpatients diagnosed as endogenously or nonendogenously depressed. The DST was significantly related to diagnosis and the PSG variables increased diagnostic accuracy. Since the DST is less expensive to administer than PSG procedures, the authors suggest that, in order to differentiate the two types of depression, the tests be given in series. That is, the DST should be administered first, and an abnormal finding treated as an indication of the need for antidepressant medication. Should the DST prove normal, sleep PSG should be given to provide greater sensitivity of diagnosis. This study thus provides evidence for the potential utility of two biological markers in treatment selection.

A small investigation was conducted by Ansseau et al. (1984) with 12 inpatients, 8 of whom had primary and 4, secondary depression. Using a 50 minute cutoff, RL correctly classified 87% of the primary and 25% of the secondary depressives ($p = < 0.05$); DST identified 62% and 25%, respectively (ns). However, even though the sensitivity of the DST and RL was not significantly different, the two measures did not identify the same individual patients. This was a rather unexpected finding, in that previous studies had implied a common biological pathway (the cholinergic system) followed by these two variables. These authors also found that RL was an unstable variable, in that values fluctuated greatly over several nights of recording. They suggested that at least three consecutive nights of recordings be averaged in order to have confidence in results. This finding of variability in RL provides a possible explanation for the varying results produced by different studies. RD was not investigated in this study.

In general, the combination of DST and RL and RD has produced similar results in several different studies, results that by and large confirm both the conclusions reached when the variables were studied singly and the expectations that the combination of variables would increase the accuracy of diagnosis.

Although the three variables under consideration present a mixed and as yet limited picture of success in discriminating depressive subtypes and in predicting treatment response, researchers in the field consider that they offer sufficient confidence to be worthy of further study (Reynolds et al., 1985). That such study is proceeding is evidenced by the increased numbers of sleep centers across the country and by the increasing research output in this field in the last five years. However, little effort to this point has been directed toward the depressed elderly. Given the special difficulties in the assessment of depression by clinical criteria and the additional risks associated with incorrect treatment in this population, there is a heightened need to investigate the performance of biological markers in older depressed adults.

It was, therefore, the purpose of this study to investigate the extent to which three biological markers for depression would predict response to four treatment conditions administered to depressed patients over the age of 65. In addition, it was intended to add to current knowledge in several areas, such as the relationship among the biological measures and the performance of these markers in a population that differs physiologically in significant ways from younger adults.

Research Questions

Observations drawn from the literature to date led to three research questions, based upon the following conclusions: (1) Among depressed adults over the age of 65, REM latency is significantly below the age and sex related norms for nondepressed patients; dexamethasone nonsuppression is evidenced in approximately 45% of cases; norms for REM density have not been established, so normative comparison is not possible. (2) Short REM latency is positively associated with both cortisol nonsuppression and elevated REM density; there is little evidence of a meaningful correlation between REM density and DST; high baseline depression levels are related to both short REM latency and DST nonsuppression, but not to REM density. The absence of the relationship between DST and REM density is explained on the grounds that REM density, but not DST, is an indicator of the extent of CNS impairment. (3) Prior observations of treatment outcome indicate that "abnormality" of biological variables is probably associated with good response to antidepressant medication rather than to psychotherapy, and biological "normality" is associated with good response to psychotherapy rather than to medication. Furthermore, REM latency has shown, more frequently than DST or REM density, both its responsiveness to treatment (i.e. latency increases in proportion to the

amount of recovery) and the occurrence of short intervals in a large number of depressed patients.

On the basis of the above findings, three questions were asked, concerning the subjects, the variables, and the prediction of treatment outcome.

Question One - Subjects. HOW IS THIS GROUP OF OLDER ADULTS CHARACTERIZED IN TERMS OF REM LATENCY, REM DENSITY, CORTISOL SUPPRESSION AND LEVELS OF DEPRESSION?

This question was asked in an effort to add to normative data previously assembled on biological variables and depression in older adults.

Question Two - Variables. WHAT IS THE RELATIONSHIP OF THE BIOLOGICAL VARIABLES TO EACH OTHER AND TO BASELINE LEVELS OF DEPRESSION?

Answers to this question were expected to aid in confirming or refuting previous findings. Any profiles of biological variables that were derived might be used in answer to research question three below.

Question Three - Treatment outcome. DO THE PRE-TREATMENT VARIABLES DST, REM LATENCY AND REM DENSITY PREDICT THE RESPONSE OF DEPRESSED OLDER ADULTS TO DIFFERENT TREATMENT CONDITIONS?

This question was asked in order to aid in resolving the controversy surrounding the utility of biological

variables for assigning depressed patients to different treatment conditions.

In order to answer these questions, four hypotheses were tested. It was expected that:

1. Data from this population would be within the same range as that previously gathered on subjects of similar age and diagnosis, and would be different from values previously found in a population of nondepressed subjects.

2. It was expected that the relationship among the biological variables would be similar to relationships previously found in groups of depressed subjects.

3. It was expected that the relationship between the biological variables and initial levels of depression would be similar to relationships found in previous studies.

4. The biological variables would predict response to treatment for subjects in the four study conditions.

CHAPTER TWO

METHODOLOGY

This study was part of a larger investigation conducted at the Arizona Health Sciences Center to assess the efficacy of alprazolam as concomitant therapy with Cognitive Behavior Therapy (CBT) in depressed older adults. It was a placebo controlled, double-blind study, with subjects assigned to four conditions, namely, CBT plus placebo medication, CBT plus alprazolam, alprazolam alone and placebo alone. The Statement on Human Subjects and Protocol for the major study are to be found in appendices B and C respectively. This portion of the major study examined how subjects were characterized in terms of REM latency, REM density and Dexamethasone Suppression Tests, and also how these variables performed as predictors of reduction in depression levels for subjects in all four study conditions.

Subjects

Subjects for this study were outpatients recruited through the media, health care providers and social organizations. Entry criteria were as follows: age 65 or older, a DSM III diagnosis of major depression, and a

Hamilton Rating Scale score of 18 or over both at screen and after a two week placebo washout. Exclusion criteria were: current suicidal risk, history of sensitivity to benzodiazepines, history of drug or alcohol abuse, current diagnosis of psychosis, antisocial personality disorder or organic brain syndrome, poorly controlled or life threatening medical illness, prior treatment with CBT or alprazolam, and the inability or unwillingness to discontinue use of psychotropic medication other than alprazolam for the duration of the study.

Treatments

1. Alprazolam, in tablets of 1.0 mg, was administered, with weekly management by project psychiatrists in a 20-30 minute appointment. Alprazolam is a triazolobenzodiazepine derivative which is more potent but less toxic and has fewer side effects than diazepam. Although primarily an anxiolytic, hypnotic and muscle relaxant, it has been hypothesized to act on the neurotransmitter serotonin that is implicated in some forms of depression. In two studies, alprazolam has been shown effective in reducing depression, primarily of the nonendogenous type (Fabre, 1976; Feighner, Aden, Fabre, Rickels & Smith, 1983). The fact that alprazolam generally produces fewer side effects than tricyclic antidepressants

makes it a preferred medication for older patients who tend to be less tolerant of side effects than younger patients.

2. Subjects in the placebo medication condition received an inert substance for oral ingestion that was identical in appearance to alprazolam. Both placebo and active medications were provided in identical containers.

3. The psychological treatment consisted of Cognitive Behavior Therapy, based on Beck's model (Beck, Rush, Shaw & Emery, 1979) but modified somewhat for group administration and for a population of older adults (Yost, Beutler, Corbishley & Allender, 1986). This form of therapy is based on the assumption that dysfunctional thinking is strongly implicated in the etiology, exacerbation and maintenance of depression. CBT has three aims: (1) for clients to experience and understand the impact of thoughts on emotions; (2) for clients to identify their own dysfunctional cognitions together with the schema (life beliefs or attitudes) that underlie the thoughts; (3) for clients to learn how to reduce their dysfunctional thinking, and in consequence their depressed mood, by means of in vivo experiments and rational challenge. The adaptation of this form of therapy for older adults comprised a greater focus on the educational aspects of the model, a slower pace for teaching conceptual material, and

a simplification of recording devices typically used in CBT.

Psychotherapy was conducted in small groups of six to ten participants on a weekly basis for 20 weeks. Following the same protocol in each group, therapists conducted an opening round to assess progress and set an agenda for the session, then proceeded to teach a simple CBT concept, using as far as possible examples from the group members' experiences. Contact work was conducted with one or more individual clients, at the end of which group feedback was solicited. Finally, both group and individually tailored homework assignments were given, to allow members to practice the concepts learned thus far.

Group Therapists and Medical Personnel

Each group was led by two therapists, one male and one female, all of whom were selected for the project on the basis of performance to criterion level in this modality. The training program, based on Beck's CBT model, was developed at the University of Arizona, and consisted of increasingly intense levels of instruction. Trainees were selectively invited to participate in the more advanced levels as a result of videotaped role plays and live performances using the model. The therapists, selected from a pool of 14 trainees, were three doctoral candidates in counseling and rehabilitation, a masters

level social worker and a masters level counselor, all of whom had from three to ten years experience in the field. In order to ensure compliance to the CBT model across groups and to maintain therapist competence levels, the treatment was manualized, and all group sessions were videotaped and then supervised on a weekly basis by two psychologists experienced in both the therapeutic modality and the procedures in the manual (Yost, Beutler, Corbishley & Allender, 1986).

The screening and medical management of patients was conducted by three psychiatrists with from 3-15 years post residency experience. They were trained in the use of the Hamilton Rating Scale for Depression (see Appendix E) to a $\underline{r} > .80$ level of interrater reliability, and were periodically checked and retrained if necessary to prevent rater drift. It was obviously not possible for those involved in the study to be blind to patient assignment to group therapy, but all personnel and subjects remained blind to medication assignment, except under the circumstances discussed below.

Procedures

1. Initial telephone screening by psychology and counseling graduate students established that the proposed participant was depressed in some degree and met none of the more obvious exclusion criteria, such as a prior course

of CBT. The intention at the outset was to gather a pool of 60 patients to assign to the four conditions. From the 125 who responded to media advertising or were recommended by health care providers, 87 subjects (36 males, 51 females) qualified for full screening.

2. Each of these 87 patients underwent a one hour screening interview by project psychiatrists, who administered the Hamilton Rating Scale for Depression in order to quantify depression levels. When necessary to establish patients' suitability for the study, their medical records were also examined and their current physicians and medical specialists were consulted. At this stage, 12 subjects (4 males, 8 females) did not proceed with the study, for a variety of reasons: insufficient depression (n = 1), exclusionary medical condition (n = 5), and concerns about the medication being offered (n = 3). Three potential subjects withdrew without providing a reason. A total of 75 subjects (32 males, 43 females), meeting both inclusion and exclusion criteria, indicated their willingness to proceed with the study, and signed a consent form (Appendix D) acknowledging that they understood the purpose, nature, potential risks and benefits of the study.

3. The 75 selected subjects entered a two week washout period, during which time they took one daily dose

of single blinded placebo medication and were tapered off the use of all other psychotropic medication. Two males and two females did not complete the washout period, because they could not be maintained off their previous medication or showed a placebo response, in the form of either sudden reduction in depression or undesired side effects attributed by the potential subjects to the medication. At the end of the two weeks, the Hamilton Rating Scale for Depression was again administered by project psychiatrists. There was further attrition at this point, due to placebo responsiveness, continued need for other medication, or a decision to leave town. A total of 64 subjects (28 males, 36 females) thus satisfactorily completed the washout period and decided to continue with the study.

4. Before entering treatment, depending on the availability of laboratory time, screened subjects were administered the baseline measures which were the focus of this study (see Screening Measures below). Thus, patients entered the sleep lab for two consecutive nights and polysomnography was performed at the patients' usual bedtimes, as determined through history and sleep logs. Patients were allowed to waken spontaneously. The DST was also administered as part of the sleep procedure.

5. It was not possible to assign subjects randomly to the two conditions of therapy and no-therapy since subjects had to enter the group therapy when (a) the two therapists were available and (b) a large enough group of participants had been recruited. Therefore, the CBT groups were filled first and subsequent applicants assigned to no therapy. However, medication (alprazolam or placebo) was randomly assigned across all subjects, regardless of therapy condition. The block assignment of subjects to CBT posed little or no threat to the assumption of randomization, since entry criteria did not change throughout the study, and the overall length of the study (two and a half years) was probably sufficient to ensure that subjects were not recruited cyclically.

6. All subjects attended twenty weekly medical management interviews of 20-30 minutes with the project psychiatrist, during which time minimal support counseling was conducted. As part of these sessions, both psychiatrists and subjects completed instruments assessing depression levels, global well being and response to medication. Subjects, psychiatrists and group therapists remained blind to medication status until the entire study was completed, except on the few occasions when it became necessary to break the code, for example in the case of severe side effects or worsening depression that threatened

the patients' well-being. Subjects assigned to group therapy attended, in addition to the medical management interview, twenty weekly sessions of CBT, of approximately ninety minutes duration.

7. After 20 weeks treatment, or immediately upon dropping out for those who did not complete the full treatment course, subjects again took the DST and entered the sleep lab for two consecutive nights.

8. For all subjects, there were three followup sessions with the project psychiatrist, at monthly intervals. These sessions were similar to the weekly treatment monitoring interviews, except that subjects were gradually tapered off medication. The medication code was not broken until the last subject had completed follow up.

Screening Measures

Dexamethasone suppression test

This measure is designed to assess levels of cortisol secretion in response to a single dose of dexamethasone, a cortisol imitator (Carroll, 1982). In the current study, patients were administered 1.0 mg. of dexamethasone before sleep on their first night in the sleep laboratory. Twenty four hours later, on their second visit to the laboratory, blood was drawn, and the sample submitted to radioimmunoassay to determine the extent to

which the patient had resisted suppression of natural cortisol production. Results were recorded as micrograms per deciliter, with four micrograms or above considered as an indicator of failure to suppress (Carroll, 1982).

Sleep EEG REM variables, density and latency

The electroencephalogram (EEG) and the electrooculogram (EOG) were recorded on a polygraph at the rate of 1 cm/second. For the EEG, occipital, central and frontal leads were used, and for the EOG, electrodes were applied to the upper and lower canthi. Analysis of the recording was made according to the criteria of Rechtschaffen and Kales (1968) by an experienced reader pretrained to reliability and blind to the treatment category of the patient. Eye movements during REM sleep were measured, scoring only movements visible on both channels of the EOG, by trained raters with an interrater reliability of $r = .98$. Each 60 second interval of REM sleep yields an eye movement score from 0-8, the sum of scores for the whole night being the value for REM activity. The division of the REM activity score by total minutes of REM sleep gives the value for REM density, ranging from 0-8 (Rechtschaffen & Kales, 1968).

REM latency was calculated as the time in minutes from the onset of stage 1 sleep to the start of the first REM period. A second measure of REM latency was also

calculated by subtracting from the REM latency score the number of minutes spent awake between the onset of sleep and the start of REM period one. The resulting calculation is called REM latency minus awake (RLMA). Since the sleep of older adults is interrupted frequently by waketime, the RLMA measure was felt to provide a more realistic picture of this aspect of sleep. Hence, it was decided to use RLMA as the primary method to represent this variable. However, the performance of both methods of measurement, and their relationship to each other, was also investigated.

Outcome measures

Depression was measured in two ways, by clinical interview and by self report.

Hamilton Rating Scale for Depression

This is a 21 item scale of depressive symptomatology, including cognitive, affective, somatic and behavioral items, with variables measured on a three or five point scale. It is administered by a psychiatrist rater, pretrained to criterion levels, during a structured interview lasting approximately 20-30 minutes, and is designed for use only on patients already diagnosed as suffering from some form of depression. The purpose is to quantify the results of the interview and provide a measure of severity of depression (Hamilton, 1960).

The maximum score on the HRSD is 52; scores of 25 or more reflect severe depression, 18-24 moderate depression, 7-17 mild depression and 6 or below nondepressed functioning. The cutoff score of 17 is common in many research studies. After conducting an extensive review of all available research reports from 1967-1979, Hedlund and Vieweg (1979) summarize findings as follows: interrater reliability on the HRSD is consistently high (.84 and above); in terms of validity, the HRDS has differentiated depressives from both normals and nondepressed psychiatric patients; the HRSD correlates with global ratings by physicians and moderately with other commonly used measures of depression, such as the Beck Depression Inventory (the median correlation with the BDI is .58). Shaw, Vallis and McCabe (1985) report that the HDRS has also been shown to be sensitive to change in depression. Its use in psychotherapy research has been so frequent that it has emerged as a standard measure of depression. (Endicott, Cohen, Nee, Fleiss and Sarantakos, 1981).

Some aspects of the HRSD may impose limitations on its value as a measure of depression severity in older adults. The HRSD places heavy emphasis on behavioral and somatic correlates of depression, either as reported by the patient or as observed by the interviewer. Indeed,

behavioral items such as weeping can be recorded as symptomatic of depression even when the patient denies dysphoric mood. Thus, the total behavioral and somatic component of the HRDS accounts for at least 50% and up to 80% of the score (Lambert, Hatch, Kingston & Edwards, 1986). This may prove to be a drawback in the measurement of depression in the elderly, since they typically endorse more behavioral and somatic than affective items, even though some of their physical symptoms may be related to other factors rather than to depression. A further potential limitation of the scale is that it depends for its value on the skill of the rater in eliciting the necessary information. For example, has the patient reduced daily activities as a direct result of depression or as a result of other factors that heavily influence older adults, such as the weather, a minor but indisposing ailment, a period of overexertion?

Beck Depression Inventory

This is a commonly used self report measure of the severity of depression (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). There are two forms of the inventory, one 21 items long, the other a shorter 13 item version. Both scales elicit a response to each item ranging from 0 to 3. The cutoffs for the short form are: 0-4 = none or minimal depression; 5-7 = mild; 8-15 = moderate; 16 + = severe.

Correlation between the two forms is .96 (Beck & Beck, 1972). For this study, the short form was selected as being quicker to complete by patients who were required to complete many other measures (See Appendix E). The BDI taps several dimensions of depression, including the behavioral, vegetative and cognitive, and has been shown to have both construct and concurrent validity (Beck, 1967; Strober, Green & Carlson, 1981). In contrast to the HRDS, somatic and behavioral items constitute only 29% of the full score of the BDI (Lambert et al, 1986).

Studies correlating the BDI and the HRSD show equivocal results, largely because of the samples used. Coefficients range from .54 to .82, but in some cases, not all subjects were depressed, and in others, the range of depression was limited to milder levels (Carroll, Fielding & Blashki, 1973). A metaanalytic comparison of the two scales, using 25 studies where all subjects exhibited some form of depression, led to the conclusion that the HRSD consistently reflected more improvement in symptomatology than did the BDI (Lambert et al., 1986). Furthermore, the difference between the two scales was greater when they were used in tests of psychotherapy rather than of pharmacotherapy. Possible explanations for these results are: the HRSD measures improvement not recognized or acknowledged by patients who are still sufficiently

depressed to rate their symptoms inaccurately; the two scales measure different aspects of depression. In light of these and similar findings, and diagnostic factors typical of older adults, HRSD and the BDI might have been expected to show different results in this study. The decision was made, therefore, to assess first the intercorrelation of these two measures, and then perform separate analyses unless the product moment correlations between them was equal to or greater than .85.

Method of Analysis

The research questions and hypotheses generated for this study required two types of analysis. Descriptive statistics, that is, frequency counts, means, standard deviations and simple Pearson correlation coefficients were used to summarize the characteristics of the sample and the relationship of these characteristics to normative data. These analyses tested the three hypotheses concerning how this sample compares with others previously studied, the interrelatedness of the variables, and their relationship to initial severity of depression.

Hypotheses relating to the prediction of treatment outcome were based on depression change scores, and subjected to analysis by inferential statistics. The multiple linear regression (Kerlinger, 1973) method of statistical calculation, performed with the use of the

Statistical Package for the Social Sciences (Nie, Hull, Steinbrenner, & Bent, 1975) was used to analyze the contributions of the three independent variables (DST, REM density, REM latency) to each criterion variable separately. The significance of the regression equations was assessed statistically, to determine the extent to which the multiple R might be obtained by chance.

Prior to analysis of the data, several decision rules were formulated regarding use of the data. The first concerned what would constitute the first week of data for subjects assigned to group therapy. Because of scheduling difficulties, it was not possible to provide immediate postwashout assignment to the sleep lab for all subjects, but the medication regimen could not begin until after the sleep lab procedures were completed, since these required a drug-free state. However, even those subjects who had not yet entered the sleep lab had to begin group therapy when the next group started. The reverse of this problem was that some subjects were recruited several weeks before a therapy group was able to start, because of the need to wait for a large enough number to form a group, and since it was felt that it would be unethical to maintain them for that period of time in what would amount to a no treatment control condition, these subjects began medication before starting group. For nine subjects, therefore, the start of

group and the first week of medication did not coincide. However, the gap between starting dates for the two types of treatment was only one week for seven of the subjects, two weeks for another subject and three for the ninth person. For the purposes of analysis, a decision was made to consider as week 1 the first week in which any treatment, whether medical or psychological, was administered.

A second question was what constituted a reasonable trial of either form of treatment. It was decided that four weeks was the minimum period necessary to allow for treatment effects to show; consequently, subjects who did not complete at least four weeks of treatment were excluded from the analysis.

A further decision rule concerned the points in time from which change scores on the dependent measures would be derived. Dependent variables were constructed around five time periods. Baseline, or Time 1 was composed of the HRSD or BDI scores obtained at intake, that is, at the end of the second postwashout week for the HRSD and at week 1 for the BDI; Time 2 was the mean of all scores obtainable from weeks 4-6; Time 3 consisted of the mean of scores from weeks 10-12; Time 4 was the mean of scores from the three treatment sessions prior to the final session (the final session score itself was excluded in order to

avoid the potentially deleterious effect of termination on subjects' scores); Time 5 was the mean of all follow up values. Analysis therefore was applied to data from these periods: T2-T1; T3-T1; T4-T1; T5-T1.

As expected, not all subjects completed the full course of treatment, with differential dropout rates by group. It was decided to accommodate these different rates by the use of end-point analysis. This type of analysis is a conservative measure of change, where the mean of the last three recorded scores of a dropout subject is repeated for each subsequent time point in the study, thus producing a complete data set. At T1 and T2, therefore, there were 56 subjects; by T3, seven subjects had dropped out, so the data set consisted of the 49 "completers" and 7 sets of endpoint data; by T4, there were 36 "completers" and 20 endpoint data sets; by T5, there were also 36 "completers" on the HRDS and 34 on the BDI, thus giving 20 and 22 sets of endpoint data respectively.

CHAPTER THREE

RESULTS AND DISCUSSION

The various sections in the chapter will describe the method of analysis, sample demographics and intergroup validity checks, the relationship of sample baseline variables to norms and to each other, the predictive power of the biological variables, and post hoc analyses.

Sample Demographics and Validation Checks

Of the 64 subjects who satisfactorily completed postwashout procedures and agreed to enter the study, eight failed to complete the required four weeks of treatment and were excluded from analysis. A total of 56 subjects remained, with males (n = 25) constituting 44.6% and females (n = 31), 55.4% of the group. The mean age of all subjects was 70.75, with an average of 13.30 years of education.

To ensure that subjects were randomly assigned to treatment conditions, the groups were compared in terms of sex, age, education and initial Beck and Hamilton scores. Means across groups are presented in Tables One through Five. One way analyses of variance yielded no significant

differences on these variables (all p's >.20), thus supporting the effectiveness of the randomization procedure.

Hypothesis One

IT WAS EXPECTED THAT BASELINE DATA FROM THIS POPULATION WOULD BE WITHIN THE SAME RANGE AS THAT PREVIOUSLY GATHERED ON SUBJECTS OF SIMILAR AGE AND DIAGNOSIS AND WOULD BE DIFFERENT FROM VALUES PREVIOUSLY FOUND IN A POPULATION OF NONDEPRESSED SUBJECTS.

This hypothesis was tested with the t-test for independent samples. In view of the limited information available on sleep variables in subjects of this age, and the small samples used even in normative studies, it was decided to compare the current study values with those obtained in several different studies of both normal and depressed patients (Kupfer, Spiker, Coble & Shaw, 1978; Prinz et al., 1982; Reynolds et al, 1980; Reynolds, Kupfer, Taska, Hoch, Spiker et al., 1985; Reynolds et al., 1983; Vitiello et al., 1984; Williams et al., 1974). Since not all studies reported RLMA values, comparisons were made with whichever measure of RL was available. Results of these comparisons are presented in Tables Six A and Six B.

As can be seen from these Tables, subjects in the current study did not differ significantly on RD from healthy subjects in other studies (t = 1.00, df = 78; t =

.38, df = 62); nor did subjects differ significantly in RD from depressed patients in previous studies (t = .61, df = 78; t = .98, df = 62; t = 1.55, df = 62; t = 1.16, df = 71)

Comparisons of the RL and RLMA values in the study sample with those in both healthy and depressed samples presented an interesting picture. The study means of 121.19 for RL and 90.23 for RLMA were significantly longer than all means obtained in depressed groups (t = 5.53, df = 63, p < .001; t = 4.66, df = 63, p < .001; t = 7.14, df = 72, p < .001; t = 6.75, df = 79, p < .001). RL and RLMA were also longer than means in three samples of healthy subjects (t = 3.10, df = 63, p < .01; t = 4.46, df = 79, p < .001). However, there was no significant difference between the study mean RL and either the mean obtained on one sample of healthy controls (t = 1.41, df = 65) or the mean of a normative sample (t = 1.45, df = 75) described by Williams et al. (1974). Since these researchers did not report values for RLMA, it was not possible to make a similar normative comparison on this method of measuring REM latency.

A cutoff of 4.0 mcg/dcl (considered appropriate for outpatients by Carroll, 1982) was used to characterize DST nonsuppression. Of the 51 subjects for whom DST values were obtained, 33 were suppressors and 18 nonsuppressors. By test for the significance of the difference between two

independent proportions, the current study percentage of 35.29 nonsuppression was not significantly different from previous findings that the DST identifies 45% of a depressed sample ($z = 1.08$. ns).

In summary, therefore, the present sample appeared somewhat similar to previous samples of depressed subjects on DST and produced RD values somewhere between previous samples of both depressed and healthy samples on RD. The study sample presented with significantly longer RL and RLMA compared with two healthy samples, but there was no significant difference on these variables when compared to two other samples of healthy older subjects.

For REM latency, the hypothesis was rejected that this sample resembles other depressed groups; for DST, the hypothesis was accepted; for RD, the hypothesis could be neither fully accepted nor rejected.

Hypothesis Two

IT WAS EXPECTED THAT THE RELATIONSHIP AMONG THE BIOLOGICAL VARIABLES WOULD BE SIMILAR TO RELATIONSHIPS PREVIOUSLY FOUND IN GROUPS OF DEPRESSED SUBJECTS.

These relationships were tested with Pearson product moment correlations. Previously, negative relationships have been found between RL and both DST and RD, but no consistent relationship has been found between DST and RD. Table Seven presents the results of the

statistical tests. Although the expected negative correlation between RL and DST occurred, (-.12), it was not significant. The correlation of .21 between RL and RD was neither significant nor in the expected direction, but the correlation of .06 between RD and DST was both positive and nonsignificant, as anticipated. The relationship between the two methods of measuring REM latency (RL and RLMA) was also investigated, and the expected high correlation, .78 ($p < .001$), was obtained. This result is in accord with the experience of Kupfer et al. (1981) who report that the two methods are not essentially different, except that RLMA correlates more highly with severity of depression than does RL.

To summarize, the hypothesis concerning expected significant relationships among the biological variables was rejected for RL and DST, and for RL and RD, but accepted for RD and DST. On the whole, baseline values on DST, RL and RD for subjects in this study did not intercorrelate as they have been found to do in other samples of depressed patients.

Hypothesis Three

IT WAS EXPECTED THAT THE RELATIONSHIPS BETWEEN THE BIOLOGICAL VARIABLES AND INITIAL LEVELS OF DEPRESSION WOULD BE SIMILAR TO RELATIONSHIPS FOUND IN PREVIOUS STUDIES.

The relationships between these variables and level of depression as measured by both the BDI and the HDRS were tested with Pearson product moment correlations and are included in Table Seven. Specifically, researchers have found a negative relationship between severity of depression and RL, a positive one between severity and DST, and a near zero-order relationship between severity and RD. None of these expectations was met. No significant relationship was found between baseline DST scores and severity of depression on either the BDI or the HDRS. The correlation of .33 between RD and initial depression levels on the BDI was significant ($p < .01$), which does not accord with previous findings of no relationship between severity and RD. The correlation of .25 ($p < .03$) between RL and baseline HDRS scores was positive, rather than in the expected negative direction. Therefore, the hypothesis concerning the relationships between the variables and severity of depression was rejected.

Hypothesis Four

IT WAS EXPECTED THAT THE BIOLOGICAL VARIABLES WOULD PREDICT RESPONSE TO TREATMENT FOR SUBJECTS IN THE FOUR STUDY CONDITIONS.

Prior to primary analysis, a validation check of the randomization procedure was undertaken. Comparisons of the RD, RL and DST values among the four groups were

undertaken by a series of analyses of variance. There were no significant differences among groups on any of these variables.

In order to assess the hypotheses, several multiple regression analyses were performed, for three different time points, namely T3 (weeks 10-12), T4 (weeks 17-19) and T5 (followup). In accordance with pre-analysis decisions, the two dependent measures were analyzed separately, since the baseline correlation between the BDI and the HDRS was only .20. There were nine variables in each equation: baseline scores on the dependent measure (BDI1 or HRSD1), all three biological variables (RD, RLMA or DST), Therapy (CBT versus no CBT), Drug (alprazolam versus no alprazolam) and three variables indicating the interaction between therapy and each of the biological variables (DST-inter; RD-inter; RLMA-inter).

Each analysis was in two stages. The purpose of the first stage was to remove the effect of baseline HRSD or BDI scores on scores at later time points, and also to remove the effect of baseline variability in each biological measure on its respective interaction variable. Once this variability was accounted for, the second stage was designed to show the extent to which the Therapy, Drug and Interaction variables predicted scores on the dependent measures.

To achieve these purposes, the procedure for each analysis was as follows: forced entry of BDI1 or HDRS1 and of a single biological variable, followed by stepwise entry of Therapy, Drug, all three Interaction variables and the remaining two biological variables. This procedure was executed three times in all, forcing in each of the biological variables in turn. Analysis ceased at the point where no variable contributed to the equation at the .05 level of significance. A summary of results from the multiple regression analyses is presented in Table Eight.

Hamilton Depression Rating Scale

When baseline HRSD scores were force entered, initial level of depression was predictive of later change in depression, as follows:

T3: R = .34; F (1,48) = 6.2; significance of F = .02

T5: R = .30; F (1,50) = 10.3; significance of F = .03

After the forced entry of baseline scores, a biological variable was force entered. Neither RD nor RLMA significantly predicted HRSD scores at any time point. However, the forced entry of DST increased the significance of the equation at two points in time, as follows:

T3: R = .41; F (2,47) = 4.9; significance of F = .01

T5: R = .40; F (2,47) = 4.5; significance of F = .02

After the entry of the baseline HRSD and the biological variable, the remaining variables were allowed to enter

stepwise, but none reached the .05 level of significance, at any time point. These results indicate that there was a relationship between initial HRSD and DST scores and later level of depression, but neither baseline depression nor the biological variables predicted reponse to either type of treatment.

Beck Depression Inventory

When baseline BDI scores were force entered, initial level of depression was predictive of later change:

T3: R = .46; F (1,48) = 12.8; significance of F = .001.

T4: R = .28; F (2,47) = 4.1; significance of F = .05

T5: R = .32; F (1,48) = 5.5; significance of F = .02.

After the forced entry of baseline scores, a biological variable was force entered. Neither RD nor RLMA significantly predicted BDI scores at any time point. However, the forced entry of DST increased the significance of the equation at all points in time, as follows:

T3: R = .59; F (2,47) = 12.7; significance of F = .000

T4: R = .51; F (2,47) = 8.4; significance of F = .001

T5: R = .54; F (2,47) = 9.8; significance of F = .001

After the entry of baseline BDI and DST scores, the remaining variables were allowed to enter stepwise. Only the variable DST-inter entered the equation at significant levels, at all time points, as follows:

T3: R = .66; F (3,46) = 12.1; significance of F = .000

T4: R = .58; F (3,46) = 8.2; significance of F = .001

T5: R = .66; F (3,46) = 11.9; significance of F = .000

Low baseline levels of DST (i.e. normal suppression) were related to good response to psychotherapy. These results indicate that baseline DST levels heighten the ability to predict response to psychotherapy, but not to the drug alprazolam.

In summary, therefore, baseline values for both measures of depression were related to later values on the same instruments at most points in the study. The two sleep variables, RD and RLMA, did not improve the prediction of depression levels at any point. When depression was measured by self report, both DST and assignment to therapy predicted scores at all time points. DST was also predictive of change when depression was measured by psychiatrist interview. The hypothesis regarding treatment prediction was, therefore, partially accepted, since only one biological variable, DST, predicted response to treatment on only one dependent measure.

Post hoc analyses

Since the biological variables showed a more or less normal pattern at baseline, post hoc analyses were conducted to investigate the post treatment values for these variables, in order to investigate whether or not they

correlated with response to treatment. It has been reported previously that lengthening of RL correlates with recovery from depression (Kupfer et al., 1981; Kupfer et al., 1976) and that DST also normalizes with good response to treatment (Carroll, 1981). Reduced RD has also been related to reduction in depression (Kupfer, 1976; Rush, 1983). These studies, however, used largely endogenously depressed patients. On the other hand, two studies have found that changes in neither RL nor RD were associated with remission (Rush, 1983; Kupfer, 1976).

The results of pre-post t-tests on the variables are presented in Table Nine. As can be seen from this table, no variable changed significantly from baseline to end of treatment, but all variables changed in the direction of normalization. That is, mean DST and RD for the total sample fell, and RL and RLMA increased.

Summary of results

1. The expected values of two of the biological variables (i.e. short RL, high RD) were not found. DST nonsuppression was found in some subjects but the mean for the whole group was below the normal cutoff for nonsuppression.
2. The expected correlations among biological variables were not found. The usual pattern has been for DST to be accompanied by short RL, and very rarely to occur in the absence of abnormal sleep variables (Rush et al., 1982).

The opposite was found here: DST was the most abnormal of all the biological variables and DST nonsuppression was not always accompanied by short RL. RL and RD did not correlate as expected. High DST did not appear to indicate endogeneity.

3. Severity of depression was not related to any of the biological variables.

4. This sample presented with values on all variables normal or close to normal for nondepressed or reactively depressed patients. Since DSM III diagnosis and HDRS and BDI scores indicated that they were depressed rather than nondepressed, indications are that these subjects were suffering from a reactive depression.

5. Of all the biological variables, DST was the only one to predict change in depression scores, on both BDI and HDRS. Prediction was consistent across time periods. DST also was predictive of response to psychotherapy but not to pharmacotherapy. The sleep variables demonstrated no ability to predict response to treatment.

6. There were no significant changes in sleep variables or DST from pre to post treatment.

Discussion and Conclusions

Discussion of the results of this study is organized according to the research questions which guided the investigation. The first two questions are considered

together, since they both relate to baseline variables, and the answers provide information necessary to an understanding of the analysis of how RL, RD and DST predicted response to treatment in this group of subjects.

Research Questions One and Two

The first question asked: how do the biological values found in this group of older adults compare with normative values in populations matched on sex and age? Question two asked: How do the biological variables relate to each other and to baseline levels of depression as measured by the Beck and Hamilton?

Examination of results indicated that the 56 subjects in this study presented a somewhat unusual pattern of values on the baseline variables under consideration, that is, RL(RLMA), RD, DST, HRDS and BDI. In fact, this group appeared to fall between the two groups - bearing complete resemblance to neither depressed nor healthy subjects in previous studies. In an effort to clarify the picture, let us first consider each variable separately.

Since little of this discussion would be relevant if the subjects under study were not in reality depressed, it is appropriate to consider first the two measures of depression, that is, the Hamilton Depression Rating Scale and the Beck Depression Inventory. The mean HDRS in this investigation was 22.30 (sd = 3.1), which falls between

means reported for other studies of between 17.5 (Mendlewicz et al., 1984) and 25.4 (Reynolds et al., 1985). On the HDRS, therefore, the subjects in this study appeared to be moderately depressed, and comparable to other depressed research subjects. Mean BDI for the group was 11.29 (equivalent to a score of 17 on the long form), which reflects a slightly lower level of depression than does the HDRS. The correlation between these two measures of depression at baseline was .20, unusually low for two instruments that have previously correlated between .64 and .80 (Hedlund & Vieweg, 1984). A partial explanation may lie in the timing of the administration of the baseline measures. Unfortunately, the BDI was not administered until the first week of treatment, by which time subjects had spent at least two weeks in placebo washout, had entered the sleep lab and had been accepted into the study. It is reasonable to suppose that by this point the BDI, which is a self report instrument, might reflect a reduction in depression levels brought about by both the increased activity of patients and their expectations of relief. However, neither the short passage of time nor the presumed change in attitude would have been sufficient to produce significant effects on the HDRS, which focusses more on physical symptoms of depression than does the BDI.

That this group was indeed depressed is further

evidenced by the fact that for all subjects the depressive episode had lasted over three months, and for 82.6% it had lasted over a year. In addition, each subject received a diagnosis of major depression according to independently rated DSM III criteria. Since the raters were not asked to give the fifth code, none of the subjects was diagnosed as melancholic, a category generally taken to correspond to endogenous depression.

The values for both measures of REM latency were either longer than or equivalent to values found in several groups of nondepressed older adults, as long, indeed, as values found in younger healthy populations (Williams et al., 1974); furthermore, REM latency was much longer than that measured in all the comparison groups of depressed patients, and RL did not indicate severity of depression, although such a relationship has been reported by several researchers (Giles et al., 1985; Mendlewicz et al., 1984). These findings do not match those in other studies. Kupfer and his team have found short RL as a correlate of depression in virtually every research group (Kupfer et al., 1972, 1976, 1978; Reynolds et al., 1983; Reynolds et al., 1985). This team has, indeed, such confidence in RL as a marker for depression that they formulated the "rule of 90" for older adults (Kupfer et al., 1982). By this rule, if the patient's age plus his or her RL values is less than 90,

one can trust that the short RL is indicative of depression. In the current study, this rule of 90 would apply to very few subjects. It is apparent that, for the present subjects, REM latency was in no way a marker for depression, or an indicator of severity of depression.

Attempts to understand this considerable departure from previous findings must include discussion of both the measurement of the variable and the subjects themselves. Knowles, MacLean and Cairns (1982) have found that RL in depressed patients, calculated according to seven different methods, differed significantly. The scoring method used by the present lab corresponds to the most generous method, that is, counting RL as time between the onset of stage 1 and REM sleep. This generosity may account for the unusual mean length of RL noted in the present subjects. For all the depressed subjects used as comparisons with the current study subjects, the strictest definition of RL was used, counting from the onset of the first stage 2 epoch followed by 8 minutes of continuous sleep in the next ten minutes. While the difference in scoring methods may be responsible for some of the large discrepancy between the compared means, it must be noted that when Knowles and his colleagues scored 40 records by each method, there was at most a difference of 20 minutes between the extremes. A similar reduction in the current study mean would result in a mean

of over 100, still significantly different from the means found in other depressed subjects. It cannot be considered, therefore, that scoring of REM latency accounted for the differences between this group and those in the literature.

A more probable explanation of the longer RL noted in this study derives from the composition of different study samples. There are several ways in which the subjects in this study were probably unlike groups of subjects used in previous studies. First, in many of the studies reported so far, subjects have been routine inpatient admissions, and groups have contained mixed diagnoses of depression (e.g. bipolar, psychotic) within the reactive-endogenous subtypes. By contrast, the rigorous exclusion criteria for the current study resulted in a homogeneous sample that was unlikely to contain many severely disturbed patients, since all were ambulatory, and bipolar and psychotic depressives were screened out, as were people at high risk for suicide. Second, many of the current subjects were recruited through the media specifically for a research project and were willing to accept both the possibility of assignment to a placebo condition and the commitment to a certain amount of tedious paperwork, implying that their depression was probably not endogenous. While endogeneity does not necessarily indicate the severity of the depression, patients with this type of disorder tend not to respond to

environmental change, and would be less likely than reactive depressives to find the motivation to enter a research study, especially one requiring the efforts necessary to respond in a group therapy situation.

Third, physicians who referred patients to the project were also aware of the conditions of the study and would doubtless not have suggested participation to patients they felt to be in need of the often more aggressive treatments typically employed with endogenous inpatients. On several counts, therefore, subjects in the present study could be characterized as suffering from nonendogenous depression. In that case, they would not be expected to present with abnormal sleep values, since the strongest findings so far have indicated that short RL is a marker for endogenous rather than for reactive depression (Feinberg et al., 1982; Rush et al., 1982; Kupfer et al., 1972, 1976, 1978).

This conclusion is supported by the fact that subjects responded to CBT with significant reduction in depression scores (as measured by the BDI), whereas medication was less effective. It has been suggested that people whose RL is greater than 70 minutes are not endogenously depressed and may require and respond to psychotherapy rather than medication (Kupfer, 1983). Thase (1986) found that subjects with RL greater than 60 minutes

responded to CBT, whereas those with RL below 60 responded very poorly.

In several recent studies, REM density has not upheld its early promise as a marker for depression, typically proving less valuable than RL, findings borne out by the present study, where RD was not significantly different from RD in either depressed or healthy age-matched comparison groups. A summary of RD values in 16 studies indicated that the mean RD was 1.9 for subjects with primary and major depression and 1.55 for control groups of both secondary depressives and healthy subjects (Gillin et al., 1984). The present sample mean RD of 1.66 thus fell between these two types of groups, differing little from either, and thus contributing poorly to the understanding of this sleep variable in subjects in the current study. The typical relationship between RL and RD was not confirmed in this study, as RD appeared to act independently of RL. It must be remembered, however, that the correlation between these two sleep variables has been found most in endogenously depressed inpatients. One might conclude that the very independence of RL and RD in this study could be regarded as an indicator that the subjects were not endogenously depressed.

Like RD and RL, DST results were somewhat ambiguous. The mean for the whole sample (3.9) was below the 4.0

cutoff, and the mean for nonsuppressors only 7.9, despite a distribution positively skewed by three scores that were extremely high for this group. Thus, DST scores in this group were, on the whole, normal (Carroll, 1982). Furthermore, DST did not correlate, as previously found in endogenously depressed samples, with RL. Thus, indications are, again, that this sample resembled reactive rather than endogenously depressed patients.

The fact that there were 18 nonsuppressors in this study may have little significance, given the characteristics of this group. Consider the results of a study undertaken on a sample of mildly to moderately depressed outpatients, recruited by self referral and through health professionals (Rabkin, Quitkin, Stewart, McGrath & Puig-Antich, 1983). These researchers found that DST was slightly related to HDRS scores but not to endogenous or reactive subtype. They concluded that DST results obtained on severely depressed patients do not apply to those who are less ill. Furthermore, high levels of nonsuppression have been found in other groups of nonendogenously depressed patients (Shulman, 1980). In particular, it has been found that in older adults the hypothalamic-pituitary axis is more sensitive to stress than in younger people (Jacobs et al., 1984). In a recent study, Baumgartner, Haack and Vescei (1986) found that

nonsuppression occurred more frequently in older than in younger patients with similar diagnoses of depression. These authors discuss the influence on DST results of such intervening variables as time and frequency of blood sampling and nonspecific stress (e.g. hospital admission, psychotherapy) and speculate that DST represents a general biological disturbance common to several psychiatric conditions, an index of illness rather than a marker for depression. This hypothesis would help explain why, in this study, the DST was not related to the sleep variables, which have been linked more consistently to depression, especially of the endogenous type.

In summary, this study provided mixed answers to the two research questions concerning the characteristics of the subjects on baseline variables and the interrelationship of these variables. In terms of RL, this group resembled healthy more than endogenously depressed subjects, which is consistent with a categorization of this group as reactively depressed. DST results were consistent with a description of subjects as either endogenously or reactively depressed, but lacked the correlation with RL that might have confirmed the presence of endogenous depression. RD contributed nothing to the comparison of normal with healthy subjects. Severity of depression was unrelated to the biological variables. This group of subjects, therefore, was unique in

that DST, RD and RLMA did not appear to "fit" each other in the pattern anticipated from previous studies. The lack of "fit" is understandable if one views these subjects as reactively depressed.

Findings from this study tend to support previous indications that biological variables are less reliable in older than in younger subjects. For example, although RL has been related to severity of depression in several of the studies discussed so far, this correlation was not found in an inpatient sample of older adults (Reynolds, 1985), nor was RL found to distinguish demented from depressed patients (Reynolds, 1983). In that same study, there was a significant difference in RD between demented and depressed subjects, but the same variable did not differentiate depressed from demented from healthy control subjects in the author's later study (1985). Different values for RD in older subjects have also been found, depending on patient status, with outpatients showing reduced RD, and inpatients elevated levels, regardless of level of depression (Kupfer, 1982). The current study would tend to confirm the suspicion that biological variables may not perform as consistently in older patients. If this is indeed the case, then the utility, for older adults, of biological variables for diagnosis and treatment assignment is reduced.

It is possible, too, that the biological variables

are not as consistent in any population, regardless of age, as was previously found. For example, one study found that RL was so unstable over 4-6 consecutive nights of recording that the researchers suggested this variable could only be used as a marker for depression if data from at least three consecutive nights were used. (Ansseau et al., 1984). In the case of DST, Sherman, Pfohl and Winokur (1984) have found that a person can show both suppression and nonsuppression within a short time period in the same afternoon. This would suggest that more than one DST needs to be obtained in order to provide more stable results.

Research Question Three

The third research question concerned the power of the biological variables to predict subjects' response to four different treatment conditions. The results of the multiple regression analyses must be explored in their full context, that is, in light of the patterns of biological variables characteristic of the subjects in this study (described above) and in the context of treatment outcome, as measured by the HDRS and the BDI, and reported elsewhere (Beutler et al, under review). The results concerning the effectiveness of treatment can be summarized as follows.

A 2 (alprazolam vs.placebo) by 2 (group therapy vs. no group therapy) by 5 (time of assessment) analysis of variance, with repeated measures on the third factor, was

undertaken to assess the differential effectiveness of treatment. Analysis of the HDRS scores revealed that, although there was a mean reduction for the whole group in depression across time, there was only one significant effect ($p < .05$), for time of assessment from T1 (baseline) to T2 (weeks 4-6). There were no significant differences among groups at any time point.

When the BDI data were similarly examined, two significant effects were found. Depression was reduced across time ($p < .001$) from T1 to T5 (follow-up). In addition, analysis revealed a time by group interaction effect ($p < .01$), with subjects assigned to group therapy (i.e. groups I and II) showing significantly lower levels of depression at all times of assessment. By contrast, subjects not assigned to group therapy (groups III and IV) tended to return to near baseline levels of depression by follow-up.

To set the scene for a discussion of results related to treatment prediction, we can summarize as follows: this study treated older adults who, though undoubtedly suffering from a major depressive disorder, were probably not endogenously depressed, and presented with baseline values of biological variables at normal or near normal levels for nondepressed or reactively depressed adults. As might be expected in such a population, they responded to

psychotherapy but not to drug treatment or to a placebo condition.

Given these findings, the expectations that the biological variables would aid in predicting response to treatment were unlikely to be met. It is, therefore, entirely in keeping with the characteristics of this group of subjects that the multiple regression analyses produced the results they did. In a biologically normal population, the sleep variables RLMA and RD would not be related to recovery from depression, and would not interact with treatment effects. The DST provided the only indication of biological abnormality in this group, and, not surprisingly therefore, was the only variable related to treatment outcome, doubling or tripling the amount of variance in depression scores accounted for at each time point. The finding that low baseline DST levels were related to good response to CBT provides confirmation of previous findings, namely that normal suppression of DST indicates lack of endogeneity and a greater likelihood of response to psychotherapy.

Limitations, and directions for further research

Perhaps the most important limitation derives from the subjects themselves. This was a group of biologically normal depressed patients, and results cannot, therefore, be generalized to patients with depression characterized by

levels of RD or RLMA typical of endogenous depression. Even in the case of the DST, caution must be observed in the use of the results of this study. It would not be accurate to assume that normal DST suppression is a clear indicator of the need for psychotherapy. For one, this study only drew a single blood sample, which may not supply a complete picture of the patients' status on this measure. Since the DST has been shown less reliable in older patients, there is probably a greater need for multiple blood assays in this population before any clinical judgment can be made regarding treatment assignment. For another, normal DST results do not necessarily indicate that treatment with medication would be of no value.

This was also a very homogeneous group of patients, established as a result of stringent exclusion and inclusion criteria. Results from the study cannot, therefore, be applied to patients with current suicidal intent or bipolar or psychotic features, or to inpatients.

A further limitation lies in the treatments used in this study. Clearly, results cannot be generalized to medications other than alprazolam or psychotherapy other than CBT.

There may also be a limitation in the instruments used to measure depression. The difference in results on the HDRS and the BDI, both from analysis of variance and

from multiple regression analysis raises a question about the relationship of the two dependent measures to each other. The low correlation of .20 between HDRS and BDI at baseline has already been discussed as perhaps due in part to the different times at which the measures were taken. One must also take into consideration the possibility that these instruments measure different aspects of depression, in particular, aspects likely to be affected by the age of the subjects. In other words, the HDRS' focus on vegetative signs might produce biased or unreliable results in a group of people who are subject to many physical complaints which may or may not be the result of depression, and who are also inclined to somatize psychological discomfort. For most subjects, the duration of the study was between seven and eight months, during which period of time, the aging process and possible development of new physical ailments might well artificially inflate HDRS scores in the later stages of the study. The BDI, on the other hand, is concerned more with mood than with vegetative symptoms and thus may draw different responses than the HDRS. A further possible consideration is that the BDI is more sensitive to daily fluctuations in mood than the HDRS, and that the single baseline score may not have been representative of a more enduring state. At all time points after baseline, the BDI score used for analysis was a mean of three weeks. A

further possible confounding factor on the HDRS might be rater bias, reflecting expectations that psychotherapy and an anxiolytic medication (as opposed to a tricyclic antidepressant more typically prescribed for depression) might have limited effectiveness, especially in older patients.

The results from this study suggest several directions for further research. Perhaps the most urgent is the need for large normative studies of older adults, to gather an adequate body of data on the DST and on sleep variables, in both depressed and nondepressed samples, and in patients with other psychiatric illnesses. These data would help ascertain, for example, whether the subjects in the current study were a representative sample of depressed older adults. Despite contradictions among studies, results continue to indicate that biological variables are probably related to depression, but refining this relationship will require a great deal more information.

In particular, the relationship of DST to sleep variables and to depression needs further exploration. Since the DST is apparently quite vulnerable to influences other than depressive state, there needs to be more investigation of variability in DST results. A suggestion is that more extensive medical examinations and more stringent medical criteria be employed in future studies on

DST, in order to eliminate influencing variables. The multiple medications taken by older adults might well be found to interact with DST.

The lack of correlation between the HDRS and the BDI suggests a need, especially with older patients, for the use of multiple instruments in the assessment of depression. It would be valuable to explore the possibility of modifying the HDRS for older adults.

Finally, it is evident from the paucity of studies of this nature with older adults that more attention needs to be directed to this population. While it is encouraging that adults over 65 respond to psychotherapy, the finding that a biological variable strongly influences this response indicates a need to attend to the physiological status of patients before assigning them to treatment.

APPENDIX A

TABLES

TABLE I

MEAN AGE BY GROUP

	Group I	Group II	Group III	Group IV	Total
MALE	70.33	71.25	68.86	70.50	70.24
FEMALE	74.71	71.63	67.80	70.09	71.16
Total	72.69	71.44	68.42	70.20	70.75

TABLE II

MEAN YEARS EDUCATION BY GROUP

	Group I	Group II	Group III	Group IV	Total
MALE	14.17	12.89	14.00	16.50	14.08
FEMALE	12.14	12.75	13.00	14.00	13.10
Total	13.08	12.81	13.58	14.66	13.30

TABLE III

GROUP ASSIGNMENT BY SEX

	Group I	Group II	Group III	Group IV	Total
MALE	6	8	7	4	25
FEMALE	7	8	5	11	31
Total	13	16	12	15	56

Group I = CBT plus alprazolam
Group II = CBT plus placebo
Group III = Alprazolam
Group IV = Placebo

TABLE IV

MEAN BASELINE HAMILTON SCORES BY GROUP

	Group I	Group II	Group III	Group IV	Total
MALE	23.83	20.75	18.86	20.25	20.88
FEMALE	21.14	21.50	22.40	21.73	21.65
Total	22.38	21.13	20.33	21.33	21.30

TABLE V

MEAN BASELINE BECK SCORES BY GROUP

	Group I	Group II	Group III	Group IV	Total
MALE	12.73	11.25	9.71	10.12	10.85
FEMALE	10.92	12.75	9.90	12.31	11.47
Total	11.57	12.00	9.79	11.73	11.29

Group I = CBT plus alprazolam
Group II = CBT plus placebo
Group III = Alprazolam
Group IV = Placebo

TABLE VI A

SLEEP VARIABLES IN COMPARATIVE SAMPLES (nondepressed)

	Group 1 n = 56	Group 2 n = 25	Group 3 n = 9	Group 4 n = 21
AGE	70.3	69	65.5	70
PATIENT STATUS	Out-patient	Non-patient	Non-patient	Non-patient
RD				
Mean	1.66	1.35	1.78	
S.D.	.58	.48	.45	
t		1.00	.38	
p		ns	ns	
df		78	62	
RL				
Mean	121.19	67.2	99.4	97.1
S.D	51.73	8.5	72.76	48.43
t		3.10	1.45	1.41
p		<.01	ns	ns
df		63	75	65
RLMA				
Mean	90.23	57.6	53.8	
S.D	34.45	16.9	7.9	
t		4.46	3.16	
p		<.001	<.01	
df		79	63	

Group 1: Current study
Group 2 Reynolds, Kupfer, Taska, Hoch, Spiker, Sewitch, Zimmer, Marin, Nelson, Martin & Morycz, 1985.
Group 3: Vitiello, Bokan, Kukull, Muniz, Smallwood & Prinz, 1984.
Group 4: Williams, Karacan & Hursch, 1974.

RD = REM density; RL = REM latency; RLMA = RL minus awake.

TABLE VI B

SLEEP VARIABLES IN COMPARATIVE SAMPLES (depressed)

	Group 1 n = 56	Group 2 n = 25	Group 3 n = 9	Group 4 n = 9	Group 5 n = 18
AGE	70.3	69	65.5	62	64
PATIENT STATUS	Out-patient	In-patient	In-patient	Out-patient	In-patient
RD					
Mean	1.66	1.57	1.88	1.46	1.82
S.D	.58	.67	.82	.16	.15
t		.61	.98	1.55	1.16
p		ns	ns	ns	ns
df		78	62	62	71
RL					
Mean	121.19	27.5	40.1	33.6	
S.D	51.73	21.8	5.2	5.7	
t		5.53	4.66	7.14	
p		<.001	<.001	<.001	
df		63	63	72	
RLMA					
Mean	90.23	33.3			
S.D	34.45	36.5			
t		6.75			
p		<.001			
df		79			

Group 1: Current study
Group 2: Kupfer, Spiker, Coble & Shaw, 1978.
Group 3: Reynolds, Coble, Black, Holzer, Carroll, & Kupfer, 1980.
Group 4: Reynolds, Spiker, Hanin & Kupfer, 1983.
Group 5: Reynolds, Kupfer, Taska, Hoch, Spiker, Sewitch, Zimmer, Marin, Nelson, Martin & Morycz, 1985.

RD = REM density; RL = REM latency; RLMA = RL minus awake

TABLE VII

INTERCORRELATIONS AMONG BASELINE VARIABLES

	DST	RD	RL	RLMA	HAM	BDI
DST	1.000	.057	-.119	-.238	.064	.061
RD		1.000	.212	.143	.174	.326*
RL			1.000	.781	.247	.209
RLMA				1.000	.136	.120
HAM					1.000	.202
BDI						1.000

* $p < .02$

TABLE VIII

PREDICTION OF TREATMENT RESPONSE

Dependent measure = Hamilton Depression Rating Scale

Predictor	Time	R (R sq)	F (df)	sig F
Enter HRSD 1	T3	.34 (.12)	6.2 (1,48)	.02
Enter DST		.41 (.17)	4.9 (2,47)	.01
Enter HRSD 1	T5	.30 (.09)	4.8 (1,48)	.03
Enter DST		.40 (.16)	4.5 (1,47)	.02

Dependent Measure = Beck Depression Inventory

Predictor	Time	R (R sq)	F (df)	sig F
Enter BDI 1	T3	.46 (.21)	12.8 (1,48)	.001
Enter DST		.59 (.35)	12.7 (2,47)	.000
Step DST-Inter		.66 (.44)	12.1 (3,46)	.000
Enter BDI 1	T4	.28 (.08)	4.1 (1,48)	.05
Enter DST		.51 (.26)	8.4 (2,47)	.000
Step DST-Inter		.58 (.35)	8.2 (3,46)	.000
Enter BDI 1	T5	.32 (.10)	5.5 (1,48)	.02
Enter DST		.54 (.29)	9.8 (2,47)	.000
Step DST-Inter		.66 (.44)	11.9 (3,46)	.000

TABLE IX

PRE POST CHANGE IN BIOLOGICAL VARIABLES

	Pre	Post	Change	t (df)	Signif.
RL	121.19	145.45	24.26	.87 (54)	ns
RLMA	90.20	111.87	21.67	1.03 (54)	ns
RD	1.66	1.41	-.25	1.24 (53)	ns
DST	3.90	3.44	-.46	.71 (49)	ns

APPENDIX B

STATEMENT ON HUMAN SUBJECTS

The broader study of which this dissertation is a part has been approved by the committee on human subjects research. Since the data used here were already collected as part of the treatment process and this was approved at that time, no further approval is necessary.

APPENDIX C

PROTOCOL

THE UPJOHN COMPANY
DIVISION OF MEDICAL AFFAIRS

PSYCHOPHARMACOLOGY

Protocol No. 4603

XANAX® (Alprazolam) - IND No. 7111 - Phase IV

"Enhancement of Psychotherapeutic Effects in the Depressed Elderly Using XANAX® and Cognitive Behavior Therapy"

I. RATIONALE

Research on XANAX® (Alprazolam) has primarily dealt with various forms of nonpsychotic, unipolar depressive disorders among the young and middle adult populations. Unanswered by current research are questions having to do with the degree to which XANAX facilitates psychotherapeutic involvement and the degree to which it might be applicable to the specific problems of depression among the elderly. The current project is designed to attend to both of these issues.

The elderly population represents a unique group insofar as depression is concerned. Not only are they subject to an increasing number of personal losses as age increases, but they also confront other major depressogenic issues having to do with retirement, relocation, disability, and death. Hence, depressive reactions tend to increase with age and a common approach to treatment is to prescribe increasing doses of various medications for specific symptoms. Many elderly depressed patients, in spite of the significance of their depressive symptoms, do not come under the purvue of the mental health system. Instead, they are treated symptomatically by family physicians and, thereby, encounter the multiple problems of polypharmacy. Since depression in the elderly is often accompanied by sleep disturbance and anxiety, the use of antianxiety medication as well as hypnotics and antidepressants increases exponentially. The elderly are often subject to exacerbations of the usual negative side effects of such medications, however. As a consequence, increasing attention has been turned to the supplementation of medication with psychotherapy.

The most specific and possibly powerful psychotherapeutic strategy available involves the use of cognitive behavior therapy (CBT) as formulated by Beck and his colleagues at the University of Pennsylvania. While this is a reasonably powerful intervention, there are certain problems with it among those elderly individuals who may require antidepressant medication as well. Some writers suggest that rather than enhancing the effects of CBT, tricyclic antidepressants may actually impede these effects in certain patients, resulting in perpetuation of depression. Such a concern may be particularly critical for the depressed elderly. It is important to investigation, therefore, whether XANAX may represent a more appropriate alternative to Imipramine as a concomitant treatment with cognitive behavior therapy in this age group. The degree to which XANAX alone and in combination with CBT, effects depression and its correlated symptoms (e.g., those related to sleep and sexual disturbance) bears investigation.

Protocol 4603

II. PURPOSE OF STUDY

To assess the efficacy of XANAX as concomitant therapy with cognitive behavior therapy in depressed, elderly patients.

III. SUMMARY OF STUDY PLAN

This will be a placebo controlled, double-blind study of 8 months duration. The initial 20 weeks will include psychological intervention and drug therapy; and the remaining 12 week period will involve medication alone.

IV. METHODS

A. Entry Criteria

1. Age 65 or older (males and females).

2. DSM-III diagnosis of Major Depressive Disorder.

3. Hamilton Depression Scale of 18 or over (at intake and post-placebo washout).

B. Exclusion Criteria

1. Current suicidal risk.

2. History of sensitivity to benzodiazepines.

3. History of drug or alcohol abuse.

4. Patients who are psychotic or who are diagnosed as having Antisocial Personality Disorder or Organic Brain Syndrome.

5. Patients with poorly controlled or life-threatening medical illness.

6. Patients requiring other psychotropic medication(s) during the course of the study.

7. Patients who have had prior course of cognitive behavior therapy.

V. PROCEDURES

A. Participation in this study is voluntary. The nature of the study will be fully explained to the patient and all questions regarding the study will be answered. Upon approval, an informed consent form will be signed by the patient.

B. After having signed the informed consent and satisfied entry/exclusion criteria, the patients will be entered into the study.

C. A two week placebo washout will then begin during which time single blinded placebo will be given twice daily. Upon completion of placebo washout, the Hamilton Depression Scale will again be given and placebo responders will be omitted from the study.

Protocol 4603

D. Subjects will now be randomized into one of four cohorts, fifteen per group.

The four cohorts:

1. Cognitive behavior therapy plus XANAX.

2. Cognitive behavior therapy plus placebo.

3. XANAX alone.

4. Placebo alone.

E. Medication

1. XANAX (alprazolam) 1.0 mg tablets

2. Matching placebo tablets

Medication will be provided to subjects during a weekly twenty minute clinical management session with a psychiatrist who is blinded to treatments.

3. Medication regimen:

a. Treatment day 1 and 2: 1/2 tablet before sleep
b. Treatment day 3 and 4: 1/2 tablet at noon, 1/2 before sleep
c. Treatment day 5 and 6: 1 tablet at noon, 2 before sleep

Thereafter, dosage is flexible and dependent upon clinical response and/or emergence of side effects.

4. The maximal dosage will be 8 mg of XANAX per day (8 tablets).

5. The cognitive behavior therapy program will be conducted in small groups, led by two co-therapists skilled in this modality. Treatment will continue on a weekly basis for twenty weeks, supplemented by the twenty minute, weekly medical management sessions.

F. Activities Rating Scales (See appendix)

1. At Intake Screen:

a. History and Physical Examination
b. EKG
c. Laboratory evaluation
d. Hamilton Depression Scale
e. Diagnostic Interview
f. Informed Consent

2. Post Washout:

a. Hamilton Depression Scale (to spot placebo responders)
b. Cornell Medical Index

c. Serum Cortisol Levels
d. Dexamethasone Suppression Test
e. Sexual Functioning Inventory
f. SCL-90R
g. Personal Reactions Inventory
h. Polysomnography
i. Beck Depression Scale

3. At Weekly Intervals:

a. Hamilton Depression Scale
b. Beck Depression Scale
c. Physician's Global Impression
d. Patient's Global Impression

4. At 10 Weeks:

a. Cornell Medical Index
b. Hamilton Depression Scale
c. SCL-90R
d. Physician's and Patient's Global Impressions

5. At Treatment End (Week 22):

a. Cornell Medical Index
b. Serum Cortisol Levels
c. Dexamethasone Suppression Test
d. Hamilton Depression Scale
e. Sexual Functioning Inventory
f. Beck Depression Scale
g. SCL-90R
h. Personal Reactions Inventory
i. Physician's Global Impression
j. Patient's Global Impression

G. Sleep Laboratory

Each subject will spend two nights in the sleep laboratory; one night at the end of placebo washout, one night at the end of active treatment. The records will be assessed for:

1. Sleep latency

2. Sleep efficiency

3. REM latency

4. REM density

5. Slow wave sleep suppression

H. Subjects who fail to respond to treatment and incur the need for immediate, alternative treatment will be appropriately referred.

I. Evaluability

Subjects who complete a minimum of 8 weeks (2 weeks washout, 6 weeks therapy) will be considered evaluable for efficacy if they drop out after this period of time. Patients who drop out at any time due to side effects will be considered evaluable for safety parameters.

VI. ANALYSIS

The first task of analysis will be to collapse the multiple sources of data into manageable units. Medical and laboratory data will be analysed separately but all psychological and cognitive data will be reduced through a principal components analysis. The factors resulting from such an analysis will reduce the degree of evaluation overlap and improve the reliability of the resultant change measures. Change will be assessed through the application of residual gain scores, first by a one way analysis of variance and then by a series of multiple regression analyses in which patient variables will be assessed for predictive efficiency. This procedure will allow assessment both of the degree to which treatment influences depression and the degree to which a broader range of psychiatric symptoms are altered as well.

VII. PROTOCOL EXCLUSIONS

A. Drug Accountability - -

At the completion of the study, the investigator will fill out an investigational Drug Distribution Form, and all unused drug is to be returned to The Upjohn Company, Returned Goods, 7862-41-0, Kalamazoo, Michigan, 49001.

B. Monitoring of Study

The investigator will permit a representative of the sponsor's monitor team to inspect all Case Report Forms at regular intervals throughout the study. These inspections are for the purpose of verifying the adherence to the protocol, and the completeness and exactness of the data being entered on the Case Report Forms.

If corrections or completions on the Case Report Forms are needed, a photocopy of the original Case Report Form will be returned to the investigator for appropriate action. He will initial the change and return the form to the monitor.

On-site monitor visits will be made as often as is deemed necessary by the sponsor and/or the investigator.

C. Breaking of Code

Each bottle of medication has a "tear off" label so that the contents can be identified, if necessary. In the event that the code is broken, regardless of reason, the Upjohn monitor must be notified immediately by the

investigator. All "tear off" labels are to be mailed back to Upjohn stapled to the patient's report forms.

D. Study Conditions and Precautions

Patients will be allowed their regular diets. Smoking is permitted. Patients should be warned against the use of alcohol in any form during the study. They are not to receive any other drugs except in the case of an emergency. The investigator will warn the patients about the possibility of impairment of faculties such as drowsiness or impairment of coordination or the risk of operating a motor vehicle and other similar activities while taking the medication.

E. Adverse Reactions

A record must be kept of all adverse reactions or accidents. An Adverse Reaction and Concomitant Therapy form relating to FDA Form 1639 will be submitted to the monitor promptly if side effects occur. Unexpected or unusually severe adverse reactions shall be reported to the sponsor immediately by telephone.

F. Report by Investigator

At the completion of the study, the investigator will submit a report on all results and evaluations generated by the study.

VIII. APPROVAL

Alan I. Levenson — June 7, 1982

CHIEF INVESTIGATOR
Alan I. Levenson, M.D.

DATE

L. E. Beutler — June 7, 1982

CO-INVESTIGATOR
Larry E. Beutler, Ph.D.

DATE

Robert B. Raskin — 3/25/82

MONITOR
Robert B. Raskin, M.D.

DATE

APPENDIX D

SUBJECT CONSENT FORM

SUBJECT CONSENT

(Revised 3/23/83)

Psychotherapeutic Effects in the Treatment of Depression

You are being asked to participate in a project entitled "Enhancement of Psychotherapeutic Effects in the Depressed Elderly Using XANAX and Cognitive Behavior Therapy". The purposes of this study are to understand the relative benefits of an established form of psychotherapy and a newly developed drug for treating depression. Previous research has demonstrated the value of this form of group psychotherapy in alleviating depression. Likewise, research has suggested the value of the current medication in effecting anxiety. The drug has not been fully studied, however, in its ability to alleviate depression. We wish to find out if the drug is as effective as this particular form of group psychotherapy.

You have been selected for this study because of your depressive and unhappy feelings and because we particularly want to study the effectiveness of treatment with individuals who are 65 years of age and older. Before treatment is started we will also make sure that your medical status is such that you can tolerate either of the treatments provided. This will be done through a physical examination as well as interviews with our professional staff (psychiatrists and clinical psychologists).

If you chose to participate and if your physical and psychological condition warrants it, you will be assigned to one of four treatments. Some treatments will combine the use of group psychotherapy with medication. Other subjects will be treated with medication alone. Half of those who are treated with medication either alone or in combination with group psychotherapy, will receive an inactive form of the drug, rather than the active chemical compound. This will allow us to assess the particular effects of the drug. The treatment period will last approximately 20 weeks. There will be an additional 12 weeks of follow-up, following the treatment phase. During the 12 week follow-up period, participants will continue to meet with our doctor monthly so that the effects of the medication can be assessed.

All of those who participate in this study will be followed very closely. If there is any reason to think that the treatments are harming you, we will make arrangements to provide a referral so that appropriate treatment can be supplied. In an effort to do this, we will meet with you weekly to review your response to treatment and to make whatever changes are needed in the dosages of medication or additional treatment. Periodically throughout this time, we will ask you to complete a variety of short questionnaires which will help us to determine how effective the treatments are. At the conclusion of the study, we will also be happy to answer any questions you have about the treatment. In order to help us see the effects of the medication, a doctor will take small samples of your blood

(4 teaspoons) at the beginning of the study, at the end of the treatment phase (20 weeks), and at the end of the follow-up phase (an additional 12 weeks). Prior to beginning treatment, we will need to withdraw you from all nonessential medications which might interfere with the new medication. This will take approximately two weeks. We will maintain close contact with you throughout this time and following this period so that any side effects of the medication can be appropriately treated.

For those individuals who are asked to participate in group counseling, the weekly contacts with their physician will be supplemented by weekly sessions of approximately one and a half hours of group psychotherapy. You will meet with a psychotherapist who is especially trained for this project along with a small group of other participants, like yourself. These sessions will be designed to help you understand the nature of depression and to establish some control over these feelings through a process called "cognitive modification". The procedure essentially focuses on the relationship between how one feels and what one thinks and then assists individuals to change their thought patterns in a way that makes them feel less depressed.

Both at the time you enter this study and again at the time the treatment phase ends, we will want to study how efficiently and well you sleep. In order to do this, we will ask that you come to University Hospital to sleep for two nights. You will not need to check into the hospital, and this will not affect your daytime activity. You will also be asked to come to University Hospital for weekly medication checks and some individuals will come to University Hospital for group therapy or counseling sessions. Other of the group counseling sessions can be arranged within the community, a little more convenient to your place of residence.

There will be no charges incurred for any of the treatment, evaluation, or follow-up procedures. You will receive free treatment, and free assessments of sleeping pattern and ability.

Individuals treated with the medication to be used in this study sometimes experience mild side effects. While most of these side effects are more mild and temporary than those of other medications used for depression, they may, nonetheless, occur. Dry mouth, urinary retention, and maybe some sleep disturbances will be noted. You may also experience constipation and at times some drowsiness. If these side effects occur, we will want you to report it to the physician who examines you each week so that he might make appropriate modifications. You will be given the names of two doctors who will be available to you for consultation, should any of the negative side effects of treatment become apparent. Since some people who participate in this study will receive an inactive drug and no group counseling, it is possible that your depression will continue. Moreover, occasionally people, even on an otherwise effective treatment program, experience a period of increased depression before they get better. For this reason, we will want to monitor your depression and your feelings very closely throughout the course of the treatment and through the follow-up period. However, if you become very seriously depressed during the course of this study we will want you to tell the doctor who sees you each week and arrangements will be made to find some other form of treatment which may be more helpful. At the conclusion of this study, if you still have depression and want to seek additional treatment, we would be happy to help you find appropriate treatment resources in the community.

Participation in this study entails a degree of inconvenience and time. While most of the time involvement will probably help you to feel better, some significant amount of time will be devoted simply to the interest of research. You will need to be able to travel to the University and spend the total of four nights in the EEG laboratory. There is also a slight possibility of incurring some infection or minor pain during the course of obtaining blood samples. Finally, some individuals find the process of participating in a study such as this stress producing. In deciding to participate, you must decide whether the inconvenience and time investment given to this research may be worth the possibility of being helped by the treatment provided. Most people who undergo any of the treatments to be used in this study experience a significant alleviation of anxiety and/or depression. By participating in this study you might also help us understand the benefits of a new treatment procedure and the value of a new medication. We are especially hopeful that these treatments will be useful for individuals who are 65 years of age and older. It is in this group of individuals that medication usage is often a difficult problem and alternative treatments are necessary.

Once this study is completed, we hope to publish the results in a professional journal. Information obtained from you will be coded, however, and put together with many other individuals so that there will be no chance that information provided will be identifiable in any published works. Indeed, at the conclusion of this project, all identifying information will be removed from the records in order to protect your right to confidentiality.

Finally, we would be happy to answer any questions that you have about participating in this study before you sign this document. You should also understand that you are free to withdraw from the study at any point without ill will or risk to receiving other types of treatment through the University of Arizona. If you do decide to participate, both your participating and the nature of your worries and concerns will be kept confidential. Only those directly related to this study will have information about the types of problems which you discuss in the course of your treatment. All questionnaires and forms which you complete will be given a number code so that your identity will remain anonymous. Some of the group therapy sessions will be audio or video taped and these tapes will be stored in a locked cabinet under the supervision of Dr. Larry E. Beutler. Once the information from the tapes are appropriately coded, the tapes themselves will be erased.

I have read the above "Subject Consent". The nature, demands, and benefits of the project have been explained to me. I understand that I may ask questions and that I am free to withdraw from the project at any time without incurring ill will or affecting my medical care. I also understand that this consent from will be filed in an area designated by the Human Subjects Committee with access restricted to the principal investigator or authorized representatives of the particular department. A copy of this consent is available to me upon request. I also understand that should I incur physical or emotional injury as a result of these procedures, there is no compensation available but I will be referred for further care.

APPENDIX E

INSTRUMENTS

HAMILTON RATING SCALE FOR DEPRESSION

1. DEPRESSED MOOD (sadness, hopeless, helpless, worthless)
0 Absent
1 Feelings indicated only on questioning
2 Feelings spontaneously reported verbally
3 communicates feelings nonverbally - i.e. though facial expression, posture, voice, tendency to weep
4 Patient reports ONLY these feelings in spontaneous verbal and nonverbal communication

2. FEELINGS OF GUILT
0 Absent
1 Self reproach, feels he has let people down
2 Guilt or rumination over past errors or sinful deeds
3 Present illness is a punishment. Delusions of guilt
4 Hears accusatory or denunciatory voices and/or experiences threatening visual hallucinations

3. SUICIDE
0 Absent
1 Feels life is not worth living
2 Wishes he were dead, or any thoughts of death to self
3 Suicide ideas or gestures
4 Attempts at suicide (only serious attempt rates 4)

4. INSOMNIA EARLY
0 No difficulty falling asleep
1 Occasional difficulty falling asleep - i.e. more than half an hour
2 Nightly difficulty falling asleep

5. INSOMNIA MIDDLE
0 No difficulty
1 Restless and disturbed during the night
2 Waking during the night - any getting out of bed rates 2 (except for purposes of voiding)

6. INSOMNIA LATE
0 No difficulty
1 Wakes in early hours of morning but goes back to sleep
2 Unable to fall asleep again if gets out of bed

7. WORK AND ACTIVITIES
0 No difficulty
1 Thoughts and feelings of incapacity, fatigue or weakness related to activities, work or hobbies
2 Loss of interest in activity, hobbies or work - either directly reported by patient or indirect in

listlessness, indecision and vacillation (feels he has to push self to work or activities

3 Decrease in actual time spent in activities or decrease in productivity. In hospital, rate 3 if patient does not spend at least three hours a day in activities (hospital job or hobbies) exclusive of ward chores

4 Stopped working because of present illness. in hospital, rate 4 if patient engages in no activities except ward chores or fails to perform ward chores unassisted

8. RETARDATION (slowness of speech or thought; impaired ability to concentrate, decreased motor activity)

0 Normal speech and thought
1 Slight retardation at interview
2 Obvious retardation at interview
3 Interview difficult
4 Complete stupor

9. AGITATION

0 None
1 "Playing with" hands, hair, etc
2 Hand-wringing, nail-biting, hair-pulling, biting of lips

10 ANXIETY PSYCHIC

0 No difficulty
1 Subjective tension and irritability
2 Worrying about minor matters
3 Apprehensive attitude apparent in face or speech
4 Fears expressed without questioning

11 ANXIETY SOMATIC

0	None	Physiological concomitants of anxiety:
1	Mild	Gastrointestinal: dry mouth, diarrhea,
2	Moderate	wind, cramps, belching, indigestion.
3	Severe	Cardiovascular: palpitations, headaches
4	Incapacitating	Respiratory: hyperventilation, sighing
		Urinary frequency; sweating

12 SOMATIC SYMPTOMS - GASTRO INTESTINAL

0 None
1 Loss of appetite but eating without encouragement. Heavy feelings in abdomen
2 Difficulty eating without staff urging. Requests or requires laxatives or medication for bowels, or medication for G.I symptoms

13 SOMATIC SYMPTOMS - GENERAL

0 None
1 Heaviness or aches in limbs, back or head. Loss or energy or fatigability

2 Any clearcut symptom rates 2

14 GENITAL SYMPTOMS
0 Absent — Symptoms such as loss of libido,
1 Mild — menstrual disturbances
2 Severe

15 HYPOCHONDRIASIS
0 Not present
1 Self absorption (bodily)
2 Preoccupation with health
3 Frequent complaints, requests for help, etc
4 Hypochondriacal delusions

16 LOSS OF WEIGHT (rating by history)
0 No weight loss
1 Probable weight loss associated with present illness
2 Definite (according to patient) weight loss

17 INSIGHT
0 Acknowledges being depressed and ill
1 Attributes illness to bad food, climate, overwork, virus, need for rest, etc
2 Denies being ill at all

18 DIURNAL VARIATION
0 No variation
1 Worse in a.m.
2 Worse in p.m.
If variation present: 1 Mild 2 Severe

19 DEPERSONALIZATION
0 Absent — Such as: feelings of unreality,
1 Mild — nihilistic ideas
2 Moderate
3 Severe
4 Incapacitating

20 PARANOID SYMPTOMS
O None
1 Suspicious
2 Ideas of reference
3 Delusions of reference and persecution

21 OBSESSIONAL AND COMPULSIVE SYMPTOMS
0 Absent
1 Mild
2 Severe

BECK DEPRESSION INVENTORY

This Appendix was removed due to a copyright claim.

This Appendix was removed due to a copyright claim.

REFERENCES

Akiskal, H. S., Lemmi, H., Yerevanian, B., King, D., & Belluomini, J. (1982). The utility of REM latency in psychiatric diagnosis: A study of 81 depressed outpatients. Psychiatric Research, 7, 101-110.

Altrocchi, J. (1980). Abnormal Behavior. New York: Harcourt Brace Jovanovich.

American Psychiatric Association. (1980). Diagnostic and statistical manual of mental disorders (3rd ed.). Washington, DC: Author.

Amsterdam, J. D., Winokur, A., Caroff, S. M., & Conn, J. (1982). The dexamethsaone suppression test in outpatients with primary affective disorder and healthy control subjects. American Journal of Psychiatry, 139, 287-291.

Ansseau, M., Scheyvaerts, M., Doumont, A., Poirrier, R.,Legros, J. J., & Franck, G. (1984). Concurrent use of REM latency, dexamethasone suppression, clonidine, and apomorphine tests as biological markers of endogenous depression: A pilot study. Psychiatry Research, 12, 261-272.

Arato, M., Rihmer, Z., & Szadoczky, E (1986). Seasonal influence on the DST results in unipolar depression. Archives of General Psychiatry, 43, 813.

Ayuso-Guttierez, J. L. (1983). Later life depression: Clinical and therapeutic aspects. In Davis, J. M. & J. W. Maas (Eds.), The affective disorders (pp. 203-210). Washington, DC: American Psychiatric Press.

Barnes, R., Veith, R. C., & Raskind, M. A. (1981). Depression in older patients: Diagnosis and management. Western Journal of Medicine, 135, 463.

Baumgartner, A., Haack, D., & Vecsei, P. (1986). Serial dexamethasone suppression tests in psychiatric illness: Part III. The influence of intervening variables. Psychiatry Research, 18, 45-64.

Beck, A. T. (1967). Depression: Clinical, experimental and theoretical aspects. New York: Harper & Row.

Beck, A. T., & Beck, R. W. (1972). Screening depressed patients in family practice. Postgraduate Medicine, 54, 81-85.

Beck, A. T., Rush, A. J., Shaw, B., & Emery, G. (1979). Cognitive therapy of depression. New York: Guilford.

Beck, A. T., Ward, C., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. Archives of General Psychiatry, 4, 53-63.

Berger, M., Doerr, P., Lund, R., Bronisch, T., & Zerssen, D. von. (1982). Neuroendocrinological neurophysiological studies in major depressive disorders: Are there biological markers for endogenous subtypes? Biological Psychiatry, 17, 1217.

Bielski, R. J., & Friedel, R. O. (1976). Prediction of tricyclic antidepressant response: A critical review. Archives of General Psychiatry, 33, 1479-1489.

Blazer, D. (1982a). Pharmacologic treatment. In D. Blazer (Ed.), Depression in late life (pp. 236-254). St. Louis: The C. V. Mosby Company.

Blazer, D. (1982b). Psychobiology. In D. Blazer (Ed.), Depression in late life (pp. 49-66). St. Louis: The C. V. Mosby Company.

Blazer, D., & Williams, C. (1980). Epidemiology of dysphoria and depression in an elderly population. American Journal of Psychiatry, 137, 439-443.

Branconnier, R., & Cole, J. (1981). Effects of acute administration of antidepressants on cognition, cardiovascular functions and salivation in the normal geriatric: A comparison of trazodone and amitriptyline. Journal of Clinical Psychopharmacology, 1, 825-885.

Brown, P., & Brawley, P. (1983). Dexamethasone suppression test and mood response to methylphenidate in primary depression. American Journal of Psychiatry, 140, 990-993.

Brown, W. A., Haier, R. J., & Qualls, C. B. (1980). Dexamethsone suppression test identifies subtypes of depression which respond to different antidepressants. Lancet, i, 928-929.

Brown, W. A., & Qualls, C. B. (1981). Pituitary-adrenal disinhibition in depression: Marker of a sybtype with characteristic clinical features and response to treatment? Psychiatric Research, 4, 115-128.

Brown, W. A., & Shuey, I. (1980). Response to dexamethasone and subtype of depression. Archives of General Psychiatry, 37, 747-751.

Buchsbaum, M. S., & Haier, R. J. (1983). Psychopathology: Biological approaches. Annual Review of Psychology, 34, 401-430.

Carroll, B. J. (1982). The dexamethsone suppression test for melancholia. British Journal of Psychiatry, 140, 292-304.

Carroll, B. J., Feinberg, M., & Greden, J. F. (1980). Diagnosis of endogenous depression: Comparison of clinical, research and neuroendocrine criteria. Journal of Affective Disorders, 2, 177-194.

Carroll, B. J., Feinberg, M., Greden, J. F., Tarika, J., Albala, A. A., Haskett, R. F., James, N. M., Kronfol, Z., Lohr, M., Steiner, M., DeVigne, J. P., & Young, E. (1981). A specific laboratory test for the diagnosis of melancholia. Archives of General Psychiatry, 38, 15-22.

Castro, P., Lemaire, M., Toscano-Aguilar, M., & Herchuelz, A. (1983). Abnormal DST results in patients with chronic schizophrenia. American Journal of Psychiatry, 140, 1261.

Chaisson-Stewart, G. M. (1985). Tragedies of inappropriate or inadequate treatment. In G. M. Chaisson-Stewart (Ed.), Depression in the elderly: An interdisciplinary approach. New York: John Wiley & Sons.

Coble, P., Kupfer, D. G., & Shaw, D. H. (1981). Distribution of REM latency in depression. Biological Psychiatry, 16, 453-465.

Coble, P., Foster, G., & Kupfer, D. G. (1976). Electroencephalographic sleep diagnosis of primary depression. Archives of general psychiatry, 33, 1124-1127.

Cobbin, D. M., Cairncross, K. D., Jurd, S., Veltman, D. G., & Pohlen, G. H. (1981). Urinary MHPG levels and the dexamethsone suppression test predict clinical response to the antidepressant drug mianserin. Neoroendocrinological Letters, 3, 133-138.

Cole, J. O. (1983). Antidepressant drug therapy in the elderly. In Davis, J. M. & J. W. Maas (Eds.), The affective disorders (pp. 385-391). Washington, DC: American Psychiatric Press.

Coryell, W. (1984). The use of laboratory tests in psychiatric diagnosis: the DST as an example. Psychiatric Developments, 3, 139-159.

Coryell, W. H., Gaffney, G., & Burkhardt, P. E. (1982). DSM-III melancholia and the primary-secondary distinction: A comparison of diagnostic validity using the dexamethasone suppression test. American Journal of Psychiatry, 139, 120-122.

Duncan, W. C., Gillin, J. C., Post, R. M., Gerner, R. H., & Wehr, T. A. (1980). Relationship between EEG sleep patterns and clinical improvement in depressed patients treated with sleep deprivation. Biological Psychiatry, 15, 879-889.

Endicott, J., Cohen, J., Nee, J., Fleiss, J., Sarantakos, S. (1981). Hamilton Depression rating Scale: Extracted from regular and change versions of the schedule for affective disorders and schizophrenia. Archives of General psychiatry, 38, 98-103.

Fabre, L. F. (1976). Pilot open label study with alprazolam (U31,889) in outpatients with neurotic depression. Current Therapy Research, 19, 661-668.

Feighner, J. P., Aden, G. C., Fabre, L. F., Rickels, K., & Smith, W. T. (1983). Comparison of alprazolam, imipramine and placebo in the treatment of depression. Journal of the American Medical Association, 249, 3057-3064.

Feinberg, M., & Carroll, B. J. (1984). Biological markers for endogenous depression in series and parallel. Biological Psychiatry, 19, 3-11.

Feinberg, M., Gillin, J. C., Carroll, B. J., Greden, J. F., & Zis, A. P. (1982). EEG studies of sleep in the diagnosis of depression. Biological Psychiatry, 17, 305-316.

Folstein, M. F., & McHugh, P. R. (1978). Dementia syndrome of depression. In R. Katzman, R. D. Terry & K. Bick (Eds.), Alzheimer's Disease: Senile dementia and related disorders (pp. 87-93). New York: Raven.

Ford, C. V., & Sbordone, R. J. (1980). Psychiatrists' attitudes toward the elderly. American Journal of Psychiatry, 137, 571-575.

Foster, D. G., Kupfer, D. J., Coble, P. A., & McPartland, R. J. (1976). Rapid eye movement sleep density. Archives of General Psychiatry, 33, 1119-1123.

Fowles, D. C, & Gersh, F. (1979). Neurotic depression: The endogenous-neurotic distinction. In R. A. Depue (Ed.), The psychobiology of the depressive disorders (pp. 55-79). New York: Academic Press.

Freedman, N., Bucci, W., & Elkawitz, E. (1982). Depression in a family practice elderly population. Journal of the American Geriatric Society, 30, 373-377.

Funkenstein, D. H., Greenblatt, M., & Solomon, H. G. (1952). An autonomic nervous system test of prognostic significance in relation to electroshock treatment. Psychosomatic Medicine, 14, 347-362.

Gallagher, D., & Thompson, L. (1982). Treatment of major depressive disorder in older outpatients with brief psychotherapies. Psychotherapy: Theory, Research and Practice, 19, 482-490.

Gallagher, D., Thompson, L., & Levy, S. M. (1980). Clinical assessment of older adults. In L. W. Poon (Ed.), Aging in the 80's: Psychological issues (pp. 19-40). Washington, DC: American Psychological Association.

Georgotas, A., Stokes, P. E., Hapworth, W. E., Kim, O. M., Fanelli, C., Stoll, P. M., Sinaiko, E., & McCue, R. E. (1986). The relationship of the dexamethasone suppression test to subtypes of depression and to symptomatic severity in the elderly. Journal of Affective Disorders, 10, 51-57.

Giles, D. E., Roffwarg, H. P., Sclesser, M. A., & Rush, J. A. (1986). Which endogenous symptoms relate to REM latency? Biological Psychiatry, 21, 473-482.

Gillespie, R. D. (1929). The clinical differentiation of types of depression. Guy's Hospital Reports, 79, 306-344.

Gillin, J. C., Duncan, W., Pettigrew, K. D., Frankel, B. L., & Snyder, F. (1979). Successful separation of depressed, normal and insomniac subjects by EEG sleep data. Archives of General Psychiatry, 36, 85.

Gillin, J. C., Sitaram, N., Wehr, T., Duncan, W., Post, R., Murphy, D. L., Mendelson, W. B., Wyatt, R. J., & Bunney, W. E. (1984). Sleep and affective illness. In R. M. Post & J. C. Ballenger (Eds.), Neurobiology of Mood Disorders (pp. 157-188). London: Williams & Wilkins.

Greden, J. F., Kronfol, Z., Gardner, R., Feinberg, M., Mukhopadhyayi, S., Albala, A. A., & Carroll, B. J. (1981). Dexamethasone suppression test and selection of antidepressant medications. Journal of Affective Disorders, 3, 389-396.

Griner, P. F., Mayewshi, P. J., Mushlin, A. I., & Greenland, P. Selection and interpretation of diagnostic tests and procedures. (1981). Annals of Internal medicine, 94, 553-570.

Gurland, B., Dean, L., Cross, P., & Golden, R. (1980). \the epidemiology of depression and dementia in the elderly: The use of multiple indicators of these conditions. In J. O. Cole & J. E. Barrett (Eds.), Psychopathology in the aged (pp. 37-62). New York: Raven Press.

Hamilton, M. (1960). A rating scale for depression. Journal of Neurological Neurosurgical Psychiatry, 23, 56-62.

Hamilton, M. (1982) Prediction of the response of depressions to ECT. In R. Abrams & W. B. Essman (Eds.), ECT: Biological foundations and clinical applications (pp. 113-127). New York: Spectrum Publications.

Hauri, P. (1982). The sleep disorders. Kalamazoo, Michigan: The Upjohn Company.

Hedlund, J., & Vieweg, B. (1979). The Hamilton rating scale for depression: A comprehensive review. Journal of Operational Psychiatry, 10, 149-162.

Hoeper, E. W., Nyczy, G. R., & Cleary, P. D. (1979). Estimated prevalence of RDC mental disorder in primary care. International Journal of Mental health, 8, 6-15.

Hudson, J. I., Pope, H. G., Jonas, J. M., Laffer, P. S., Hudson, M. S., & Melby, J. C. (1983). Hypothalamic-pituitary-adrenal-axis hyperactivity in bulemia. Psychiatric Research, 8, 111-117.

Jacobs, S., Mason, J., Kosten, T., Brown, S., & Ostfeld, A. (1984). Urinary-free cortisol excretion in relation to age in acutely stressed persons with depressive symptoms. Psychosomatic Medecine, 46, 213.

Janicak, P. G., Davis, J. M., Gibbons, R. D., Ericksen, S., Chang, S., & Gallagher, P. (1985). Efficacy of ECT: A meta-analysis. American Journal of Psychiatry, 142, 297-302.

Katon, W. (1984). Depression: Relationship to somatization and chronic medical illness. Journal of Clinical Psychiatry, 45, 4-11.

Kendall, R. E. (1981). The contribution of ECT to the treatment of affective disorders. In R. L. Palmer (Ed.), Electroconvulsive therapy: An appraisal (pp. 28-36). Oxford: Oxford University Press.

Kerkhofs, M., Hoffmann, G., De Martelaere, V., Linkowski, P., & Mendlewicz, J. (1985). Sleep EEG recordings in depressive disorders. Journal of Affective Disorders, 9, 47-53.

Kerlinger, F. N. (1973). Foundations of behavioral research. New York: Holt, Rinehart & Winston.

Kiloh, L. G., Ball, J. R. B., & Garside, R. F. (1962). Prognostic factors in treatment of depressive states with imipramine. British Medical Journal, 1, 1225-1227.

Knowles, J. B., MacLean, A. W., & Cairns, J. (1982). Definitions of REM latency: Some comparisons with particular reference to depression. Biological Psychiatry, 17, 993-1002.

Krishnan, K. R. R., France, R. D., Pelton, S., McCann, U. D., Manepalli, A. N., & Davidson, J. R. T. (1985). What does the dexamethsone suppression test identify? Biological Psychiatry, 20, 957-964.

Kuhn, R. (1970). The imipramine story. In F. J. Ayd & B. Blackwell (Eds.), Discoveries in biological psychiatry (pp. 205-217). Philadelphia: J. B. Lippincott.

Kupfer, D. J., & Foster, F. G. (1978). EEG sleep and depression. In R. I. Williams & I. Karacan (Eds.), Sleep disorders: Diagnosis and treatment (pp. 163-204). New York: Wiley & Sons.

Kupfer, D. J., & Foster, F. G. (1972). Interval between onset of sleep and rapid eye movement sleep as an indicator of depression. Lancet, 2, 684-686.

Kupfer, D. J., Foster, F. G., Coble, P. A., McPartland, R. J., & Ulrich, R. F. (1978). The application of EEG sleep for differential diagnosis of affective disorders. American Journal of Psychiatry, 135, 69-74.

Kupfer, D. J., Foster, F. G., Reich, L., Thompson, K. S., & Weiss, B. (1976). EEG sleep changes as predictors in depression. American Journal of Psychiatry, 133, 622-625.

Kupfer, D. J., & Frank, E. (1984). Relationship of EEG sleep to vital depression. Journal of Affective Disorders, 7, 249-263.

Kupfer, D. J., Reynolds, C. F., Ulrich, R. F., Shaw, D. H., & Coble, P.A. (1982). EEG sleep, depression and aging. Neurobiology of aging, 3, 351-360.

Kupfer, D. J., Spiker, D. G., Coble, P. A., Neil, J. F., Ulrich, R., & Shaw, D. H. (1981). Sleep and treatment prediction in endogenous depression. American Journal of Psychiatry, 138, 429-434.

Kupfer, D. J., Spiker, D. G., Coble, P., & Shaw, D. H. (1978). EEG sleep recordings and depression in the elderly. Journal of the American Geriatric Society, 26, 53-57.

Kupfer, D. J., & Thase, M. E. (1983). The use of the sleep laboratory in the diagnosis of affective disorders. Psychiatric Clinics of North America, 6, 3-25.

Lambert, M. J., Hatch, D. R., Kingston, M. D., & Edwards, B. C. (1986). Zung, Beck and Hamilton rating scales as measures of treatment outcome: A meta-analytic comparison. Journal of Consulting and Clinical Psychology, 54, 54-59.

Lewis, A. J. (1938). States of depression: Their clinical and aetiological differentiation. British Medical Journal, 2, 875-878.

Mapother, E. (1926). Discussion on manic-depressive psychosis. British Medical Journal, 2, 872-879.

Mendlewicz, J., Kerkhofs, M., Hoffman, G., & Linkowski, P. (1984). Dexamethasone suppression test and REM sleep in patients with major deoressive disorder. British Journal of Psychiatry, 145, 383-388.

Miller, M. (1978). Geriatric suicide: The Arizona study. The Gerontologist, 18, 489-495.

Miller, R. C. & Berman, J. S. (1983). The efficacy of cognitive behavior therapies: A quantitative review of the research evidence. Psychological Bulletin, 94, 39-53.

Mintz, J., Steuer, J., & Jarvik, L. (1981). Psychotherapy with depressed elderly patients: Research considerations. Journal of Consulting and Clinical Psychology, 49, 542-548.

Moffic, H. S., & Paykel, E. S. (1975), Depression in medical inpatients. British Journal of Psychiatry, 126, 346-353.

Montplaisir, J., Billiard, M. & Takahashi, S. (1978). 24 hour recording in REM narcoleptics with special reference to nocturnal sleep disruption. Biological Psychiatry, 13, 73-89.

Myers, J. K., Weissman, M. M., Tischler, G. L., Holzer, C. E., Leaf, P. J., Orvaschel, H., Anthony, J. C., Boyd, J. H., Burke, J. D., Kramer, M., & Stoltzman, R. (1984). Six month prevalence of psychiatric disorders in three communities. Archives of General Psychiatry, 41, 959-967.

National Center for Health Statistics (1977). Advance report, final mortality statistics, 1975. February 11.

Nelson, C., & Charney, D. S. (1981). The symptoms of major depressive illness. American Journal of Psychiatry, 138, 1-13.

Nelson, C., Charney, D. S., & Quinlan, D. M. (1981). Evaluation of the DSM III criteria for melancholia. Archives of General Psychiatry, 38, 555-559.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. (1975). Statistical package for the social sciences (2nd ed). New York: McGraw Hill.

Nielson, A. C., & Williams, T. A. (1980). Depression in ambulatory medical patients. Archives of General Psychiatry, 37, 999-1004.

Nymgaard, K. (1959). Studies in the sedation threshold. Archives of General Psychiatry, 1, 530-536.

Okimoto, J. T., Barnes, R. F., & Veith, R. C. (1982). Screening for depression in geriatric medical patients. American Journal of Psychiatry, 139, 799-802

Oxenkrug, G. F. (1978). Dexamethasone test on alcoholics. Lancet, 1, 795.

Paykel, E. S. (1972). Depressive typologies and response to amitriptyline. British Journal of Psychiatry, 120, 147-156.

Paykel, E. S. (1979). Predictors of treatment response. In E. S. Paykel & A. Coppen (Eds.), Psychopharmacotherapy of Affective Disorders (pp. 193-220). Oxford: Oxford University Press.

Perez-Reyes, M. (1968). Differences in sedative susceptibility between types of depression. Archives of General Psychiatry, 19, 64-71.

Prinz, P. N., Peskind, E. R., Vitaliano, P. P., Raskind, M. A., Eisdorfer, C., Zemcuznikov, N., & Gerber, C. J. (1982). Changes in the sleep and waking EEG's of nondemented and demented elderly subjects. Journal fo the American Geriatrics Society, 30, 86-93.

Prusoff, B. A., Weissman, M. M., Klerman, G. L., & Rounsaville, B. J. (1980). Research diagnostic criteria subtypes of depression: Their role as predictors of differential response to psychotherapy and drug treatment. Archives of General Psychiatry, 37, 796-802.

Quan, S. F., Bamford, C. R., & Beutler, L. E. (1984). Sleep disturbances in the elderly. Geriatrics, 39, 42-47.

Rabkin, J. G., Quitkin, F. M., Stewart, J. W., McGrath, P. J., & Puid-Antich, J. (1983). The dexamethasone suppression test with mildly to moderately depressed outpatients. American Journal of Psychiatry, 140, 926-927.

Raskin, A. (1979_. Signs and symptoms of psychopathology in the elderly. In A. Raskin & L Jarvik (Eds.), Psychiatric symptoms and cognitive loss in the elderly (pp. 75-83).

Raskin, A., & Crook, T. H. (1976). The endogenous-neurotic distinction as predictor of response to antidepressant drugs. Psychological Medicine, 6, 59-70.

Rechtschaffen, A., & Kales, A. (Eds.) (1968). A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Los Angeles: Brain Information Service/Brain Research Institute, UCLA.

Reynolds, C. F., Coble, P. A., Black, R. S., Holzer, B. C., Carroll, R., & Kupfer, D. J. (1980). Sleep disturbances in a series of elderly patients: Polysomnograph findings. Journal of American Geriatric Sociology, 28, 164-170.

Reynolds, C. F., Kupfer, D. J., Taska, L. S., Hoch, C. C., Sewitch, D.E., & Spiker, D. G. (1985). Sleep of healthy seniors: A revisit. Sleep, 8, 20-29.

Reynolds, C. F., Kupfer, D. J., Taska, L. S., Hoch, C. C., Spiker, D. G., Sewitch, D. E., Zimmer, B., Marin, R. S., Nelson, J. P., Martin, D., & Morycz, R. (1985). EEG sleep in elderly depressed, demented, and healthy subjects. Biological Psychiatry, 20, 431-442.

Reynolds, C. F., Spiker, D. G., Hanin, I., & Kupfer, D. J. (1983). Electroencephalographic sleep, aging, and psychopathology: New data and state of the art. Biological Psychiatry, 18, 139-155.

Rosin, A., & Glatt, M. (1980). Alcohol excess in the elderly. Quarterly Journal of Studies on Alcohol, 191, 53-59.

Roy-Byrne, P. (1982). Neuroendocrine tests in bulimia. Presented at American Psychiatric Association Annual Meeting, Toronto, Canada.

Rubin, R. T., & Marder, S. R. (1983). Biological markers in affective and schizophrenic disorders: A review of contemporary research. In M. R. Zales (Ed.), Affective and schizophrenic disorders: New approaches to diagnosis and treatment (pp. 53-100). New York: Brunner/Mazel.

Rundell, O. H., Lester, B. K., & Griffiths, W. J. (1972). Alcohol and sleep in young adults. Psychopharmacologica, 26, 201-218.

Rush, J. A. (1983). Clinical and etiologic implications of biologic derangements in major depressions. In Davis, J. M. & J. W. Maas (Eds.), The affective disorders (pp. 53-71). Washington, DC: American Psychiatric Press.

Rush, A. J., Giles, D. E., Roffwarg, H. P., & Parker, C. R. (1982). Sleep EEG and dexamethasone suppression test findings in outpatients with unipolar major depressive disorders. Biological Psychiatry, 17, 327-341.

Rush, A. J., Hay, B. J., Erman, M. K., Schlesser, M. A., Roffwarg, H. P., Vavasada, N., Khatami, M., Fairchild, C., & Giles, D. E. (1985). Alprazolam versus amitriptyline in depressions with reduced REM latencies. Archives of General Psychiatry, 42, 1154-1159.

Schildkraut, J. J., Schatzberg, A. F., Orsulak, P. J., Mooney, J. J., Rosenbaum, A. H., & Gudeman, J. E. (1983). Biological discrimination of sybtypes. In Davis, J. M. & J. W. Maas (Eds.), The affective disorders (pp. 31-46). Washington, DC: American Psychiatric Press.

Shagass, C., & Jones, A. L. (1958). A neurophysiological test for psychiatric diagnosis: Results in 750 patients. American Journal of Psychiatry, 114, 1002-1010.

Shaw, B. F., Vallis, T. M., & McCabe, S. B. (1985). The assessment of the severity and symptom patterns in depression. In E. E. Beckham & W. R. Leber (Eds.), Handbook of depression: Treatment, assessment and research (pp.372-408). Himewood, IL: The Dorsey Press.

Sherman, B., Pfohl, B., & Winokur, G. (1984). Circadian analysis of plasma cortisol levels before and after dexamethasone administration in depressed patients. Archives of General Psychiatry, 41, 271.

Shulman, R. (1980). The dexamethasone test and depression. Lancet, 2, 1085.

Siris, S. G., Alexander, H. G., & Stetner, F. (1892). ECT and psychotropic medication in the treatment of depression and schizophrenia. In R. Abrams & W. B. Essman (Eds.), ECT: Biological foundations and clinical applications (pp. 91-111). New York: Spectrum Publications.

Smith, M. L., Glass, G. V., & Miller, T. I. (1980). The benefits of psychotherapy. Baltimore: The Johns Hopkins University Press.

Spitzer, R. L., Endicott, J., & Robins, E. (1978). Research diagnostic criteria: Rationale and reliability. Archives of General Psychiatry, 35, 773-782.

Steinbrueck, S. M., Maxwell, S. E., & Howard, G. S. (1983). A meta-analysis of psychotherapy and drug therapy in the treatment of unipolar depression with adults. Journal of Consulting and Clinical psychology, 51, 856-863.

Strober, M., Green, J., & Carlson, G. (1981). Utility of the Beck Depression Inventory with psychiatrically hospitalized adolescents. Journal of Consulting and Clinical Psychology, 49, 482-483.

Thase, M. (1986). EEG sleep and response to cognitive behavior therapy. Unpublished data, University of Pittsburgh.

Tourigny-Rivard, M. F., Raskind, M., & Rivard, D. (1981). The dexamethsone suppression test in an elderly population. Biological Psychiatry, 16, 1177-1184.

Ulrich, R., Shaw, D. H., & Kupfer, D. J. (1980). The effects of aging on sleep. Sleep, 3, 131-141.

Vitiello, M. V., Bokan, J. A., Kukull, W. A., Muniz, R. L., Smallwood, R. G., & Prinz, P. N. (1984). Biological Psychiatry, 19, 721-734.

Vogel, G. W., McAbee, R., Barker, K., & Thurmond, A. (1977). Endogenous depression improvement and REM pressure. Archives of General Psychiatry, 34, 96-97.

Warburton, D. M. (1984). Brain, behaviour and drugs: Introduction to the neurochemistry of behaviour. London: John Wiley.

Watson, R., Hartmann, E., & Schildkraut, J. J. (1972). Amphetamine withdrawal: Affective state, sleep patterns, and MHPG excretion. American Journal of Psychiatry, 129, 263-269.

Weissman, M. (1983). Psychotherapy in comparison and in combination with pharmacotherapy for the depressed outpatient. In Davis, J. M. & J. W. Maas (Eds.), The affective disorders (pp. 409-418). Washington, DC: American Psychiatric Press.

Whitlock, F. A. (1982). Symptomatic affective disorders. New York: Academic Press.

Whybrow, P. C., Akiskal, H. S., & McKinney, W. T. (1984). Mood disorders:Toward a new pschyobiology. New York: Plenum Press.

Widlocher, D. (1983). Retardation: A basic emotional response? In Davis, J. M. & J. W. Maas (Eds.), The affective disorders (pp. 165-181). Washington, DC: American Psychiatric Press.

Williams, R. L., Karacan, I., & Hirsch, C. J. (1974). Electroencephalography (EEG) of human sleep: Clinical applications. New York: Wiley & Sons.

Yost, E. B., Beutler, L. E., Corbishley, M. A., & Allender, J. R. (1986). Group cognitive therapy: A treatment approach for depressed older adults. New York: Pergamon Press.

Zimmerman, M., Coryell, W., & Pfohl, B. (1986). The validity of the dexamethasone suppression test as a marker for endogenous depression. Archives of General Psychiatry, 43, 347-355.