depression
in the 70's

Published by Excerpta Medica
under a grant from Lakeside Laboratories, Inc.

Ronald R. Fieve, M.D.
DEPRESSION IN THE 1970's

MODERN THEORY AND RESEARCH

PROCEEDINGS OF THE SYMPOSIUM
NEW YORK, N.Y., OCTOBER 1970

Program Director and Chairman | Editor
RONALD R. FIEVE
Columbia University College of Physicians and Surgeons,
New York, N.Y.

EXCEPERTA MEDICA
Current biometric approaches to depression*

JOSEPH ZUBIN AND JOSEPH FLEISS

Biometrics Research, New York State Department of Mental Hygiene, New York, N.Y., U.S.A.

Almost 20 years ago in the Presidential Address of the American Psychopathological Association volume entitled Depression (Zubin, 1954), we laid down the blueprint for a biometric approach to the mental disorders, including depression:

‘In the biometric evaluation of the depression syndrome, the following studies have been undertaken: (1) studies for rigorously defining the concept; (2) techniques for detecting incipient depression; (3) quantification of the degree of depression present; (4) determination of personalities susceptible to depression; and (5) provision of methods for studying depression in the pure state. Since Adolph Meyer’s dictum – that we do not know how to define depression – still holds true today, we must resort to some other way of determining its essence than by logical definition. One solution is to find the common elements that characterize the depressed. T. V. Moore has used the method of factor analysis for classifying patients in accordance with their traits and confirmed Kraepelin’s nosological categories of ‘agitated depression’ (patients who were sad and anxious) and ‘retarded depression’ (physically tired and worn out). Intellectual deficits were not involved, but feelings and emotional behavior were affected. This statistical confirmation lends a more rigorous framework for nosology than previously available.’

Today in 1970, despite the facts that investigations of depression have increased quantitatively, if not qualitatively, that antidepressant drug therapy has bloomed, and that epidemiological investigations have flourished, the five basic problems of 1952 remain unsolved and continue to defy our most energetic attempts at resolution. The reason for this slow progress is not a lack of methods and techniques but rather that the basic feature of depression is an inner feeling of sadness and dejection; the rest is epiphenomena. As we have been unable to

* This work was supported in part by Public Health Service grant MH-09191, Diagnosis of Mental Disorder in the United States and United Kingdom, from the National Institute of Mental Health.
determine the cause of this inner sadness, if we could induce it experimentally we might eventually be able to determine its cause, elimination, and prevention. For the present, however, we are left with a subjective, intuitive feeling which manifests itself directly, if at all, by verbal report and indirectly by such different types of behavior as fearfulness, diminished appetite, psychomotor retardation, and a host of others.

Added to this basic problem, we have such additional problems as deciding whether depression is a transient mood state, a symptom, a syndrome, a disease, an illness, or a combination of all or some of these. Before we distinguish between these possibilities, we must first determine the purpose of the distinction— is it for (1) classification, (2) etiology, (3) clinical treatment, (4) prognosis, or (5) prevention?—for each of these purposes must be dealt with in different ways. The primary approaches for dealing with the above five purposes may be categorized as (1) clinical, (2) phenomenological, and (3) biometric. Elsewhere we have described these three contrasting approaches more fully (Zubin, 1968). Here we need merely point to the clinical approach as primarily focused on treatment and relief of illness; the phenomenological, on description and individual empathic understanding of the phenomena experienced by the patient; and the biometric, on measuring those aspects of the phenomena that are amenable to measurement or objective assessment. Further, the biometric approach is primarily focused on classification for diagnosis and prognosis and on etiology, and serves the purpose of treatment and prevention only as an evaluative method.

From the point of view of classification, it is important to have a description of the phenomena of depression that is as complete as possible. That depression is a many-faceted phenomenon is attested to by the variety of depressive states that have been recognized: normal mood swings, reactive versus endogenous depression, manic-depressive psychosis versus unipolar depression, involutional melancholia, depression in old age and schizophrenia (schizoaffective), as well as depression with a presumably organic basis (cerebrovascular disease, tumors, etc.). The multifaceted nature of depression had already been noted by Robert Burton in the early 17th century in his Anatomy of Melancholy (1927).

'To discern all which symptoms are better, Phasis the Arabian makes three degrees of them. The first is false conceits and idle thoughts: to misconstrue and amplify, aggravating everything they conceive or fear; the second is to talk to themselves, or to use inarticulate voices, speeches, obscure gestures, and plainly to utter their minds and conceits of their hearts by their words and actions, as to laugh, weep, to be silent, not to sleep, eat their meat, etc.: the third is to put in practice that which they think or speak ... Who can sufficiently speak of these symptoms, or prescribe rules to comprehend them? The four and twenty letters make no more variety of words in diverse languages, than melancholy conceits produce diversity of symptoms in several persons. They are irregular, obscure, various, so infinite, Proteus himself is not so diverse: ...' (346–347).

It is beyond the scope of this paper to deal with the total biometric approach to depression, and we shall limit ourselves to the following current issues: (1) comparison of the prevalence of hospitalized depression in the United Kingdom,
and United States; (2) the measurement of depression; and (3) the controversy regarding the continuity of neurotic and psychotic depression.

COMPARISON OF THE PREVALENCE OF HOSPITALIZED DEPRESSION IN THE UNITED KINGDOM AND UNITED STATES

Because of the tremendous difference reported in the national statistics of the United Kingdom and United States regarding the prevalence of affective disorders—especially depression—and schizophrenia (Kramer, 1969), a cross-national diagnostic study was organized in 1963 by a group representing the Institute of Psychiatry at the Maudsley Hospital in London and our own Biometric Research Unit.

Many instruments have been proposed for measuring psychopathology, including self-reporting paper and pencil questionnaires, interviewing methods, and physiological indicators. In this study we limited ourselves to the use of the systematic structured interviews developed by our own and collaborating staffs. One of the advantages of these interviews is that they are not geared to depression alone, nor to schizophrenia alone, nor to any other disorder alone, and can therefore provide a wide spectrum of measures of psychopathology for any disorder.

The systematic structured interview used in this study consisted of some 400 questions related to current mental state. Each question in the interview was accompanied by several possible response items, indicating the presence or absence of psychopathology, which could be recorded by the interviewer as he proceeded with the interview. There were a total of nearly 700 such items, 481 from the Present State Examination (PSE) of Wing et al. (1967) and 197 from the Psychiatric Status Schedule (PSS) of Spitzer et al. (1970). The mental state interview was administered within 72, and usually within 48, hours of admission. The mental state ratings were made with a high degree of reliability (Kendell et al., 1968); they were concurrently valid as they were strongly associated with the hospital’s independent diagnoses in London (although not in New York) (Gurland et al., 1970); and these ratings were predictively valid in that they were strongly associated with short-term outcome of hospitalization (Gurland et al., 1971).

Two hundred and fifty consecutive admissions to a New York State Mental Hospital and 250 consecutive admissions to a London Area Mental Hospital were studied by the project. All patients were aged between 20 and 59 years, but were otherwise unselected from all successive admissions. The major result of the study (Cooper et al., 1969; Gurland et al., 1969) has been to deflate the alleged discrepancies between the prevalence of hospitalized schizophrenic and affective disorders in the two countries (see Fig. 1).

According to the diagnostic labels provided by the indigenous psychiatrists in the two countries, schizophrenia was over 3.4 times more prevalent than affective disorders in the United States, but was only three-quarters as prevalent in the United Kingdom. The hospital diagnoses therefore indicated that the relative prevalence of schizophrenia to affective disorders was 5 times greater in the United States than in the United Kingdom.

On the other hand, according to the labels provided by the project psychiatrists,
the contrast between the two countries was substantially reduced. The project found that schizophrenia was 80% as prevalent as affective disorders in the United States and 40% as prevalent in the United Kingdom. The project diagnoses, therefore, indicated that the relative prevalence of schizophrenia to affective disorders was only 2 times greater in the United States than in the United Kingdom, i.e. 40% of the ratio based on indigenous diagnoses.

Another striking result of this study (Gurland et al., 1971) was the fact that a group of patients who suffered essentially from morbid depression and from little else were found to be labeled as depressives by their indigenous hospitals in 60% of the cases in the United Kingdom but in only 20% of the cases in the United States. Despite this apparent mislabeling, 84% of these patients received antidepressant treatment (drugs or ECT) in the United Kingdom, but in the United States only 50% received such treatment. This paradox of labeling wrongly and yet tending to treat properly remains an enigma, but note that 16% of the morbid depressives in the United Kingdom and fully 50% of those in the United States failed to obtain treatment aimed at their depression.

THE MEASUREMENT OF DEPRESSION – FACTOR ANALYSIS RESULTS

One of the most frequently used methods for classifying responses from patients into dimensional structures is factor analysis. In his review of the major factorial
analyses of psychopathology, Costello (1970) points out how inconsistent the results of one factor analysis typically are when compared to another. One of the few consistent results, however, has been the combination of depression and anxiety into a single factor. We thus have the anxious depressed factor of Wittenborn and Holzberg (1951), the anxious intropunitiveness factor of Lorr et al. (1963), the depression-anxiety factor of Cohen et al. (1966), and the depression-anxiety factor of Spitzer et al. (1967).

In only one of the factor analytic studies cited by Costello did anxiety and depression emerge as separate factors (Wittenborn, 1963). The second successful separation of depression and anxiety into two factors has been provided by the Cross-National Study of Diagnosis of the Mental Disorders just described.

The 700 item-responses to the mental state interview were subjected to a Varimax factor analysis and the results are shown in Table I. In the first column, the factors could be regarded as mood factors, with the exception of the last — obsessions; in the second column, most could be regarded as disorganization factors. These two types of factors could be regarded as reflecting affective and schizophrenic disorders, respectively.

TABLE I

25 factors of psychopathology found by U.S.-U.K. diagnostic project

<table>
<thead>
<tr>
<th>Depression</th>
<th>Sloppy appearance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety</td>
<td>Disorientation</td>
</tr>
<tr>
<td>Reported restlessness</td>
<td>Lack of insight</td>
</tr>
<tr>
<td>Observed restlessness</td>
<td>Depersonalization-derealization</td>
</tr>
<tr>
<td>Retarded speech</td>
<td>Paranoid delusions</td>
</tr>
<tr>
<td>Retarded movement</td>
<td>Grandiose delusions</td>
</tr>
<tr>
<td>Hypomania</td>
<td>Control delusions</td>
</tr>
<tr>
<td>Somatic concerns</td>
<td>Visual hallucinations</td>
</tr>
<tr>
<td>Non-delusional suspiciousness</td>
<td>Auditory hallucinations</td>
</tr>
<tr>
<td>Reported belligerence</td>
<td>Bizarre behavior</td>
</tr>
<tr>
<td>Observed belligerence</td>
<td>Non-social speech</td>
</tr>
<tr>
<td>Obsessions</td>
<td>Flat affect</td>
</tr>
<tr>
<td></td>
<td>Incomprehensibility</td>
</tr>
</tbody>
</table>

The most striking and novel outcome of this analysis was the separation of the factors of anxiety and depression. The correlation between these two factors was still appreciable (0.48), but the correlations among the items of the anxiety factor and among the items of the depression factor were considerably higher than the cross-correlations between the two sets of items.

The depression factor

Forty-five items were assigned to the depression factor: 16 from the PSS and 29 from the PSE. Some of the items correlating highest with the total factor score were:

'Has less interest than usual in things,'
‘Has difficulty in concentrating,’
‘Admits he is often sad or depressed,’
‘Has felt life wasn’t worth living,’ and
‘Feels overwhelmed with life.’

Other items contributing to the factor were:

‘Has too little energy,’
‘Keeps losing his train of thought,’
‘Has thoughts about killing himself,’ and
‘Diminished appetite.’

The anxiety factor

Twenty-one items were assigned to the anxiety factor: 4 from the PSS and 17 from the PSE. Some of the items correlating highest with the total factor score were:

‘Presence of situation giving rise to uneasiness or anxiety in the past month,’
‘Tried to avoid that situation in the past month,’
‘Trembling, hand shaky, weak at the knees,’
‘Sweating, clammy hands,’ and
‘Hot and cold feelings, blushing, pallor.’

It is striking that neither of two PSS items describing general feelings of anxiety, ‘Admits that he is often anxious’ and ‘Admits he feels anxious most of the time,’ contributed to this factor. In fact, these items correlated slightly higher with the items in the depression factor than with those in the anxiety factor.

Our anxiety factor is measured, not by general and nonspecific feelings of unease or fear, but rather by the presence of specific situations giving rise to anxiety and by the physiological concomitants of anxiety. The specific situations described among the twenty-one items were:

‘Staying home alone,’
‘Going out alone,’
‘Being in an enclosed space,’ and
‘Being in crowds.’

In addition to those already cited, the physiological signs of anxiety which contributed to the factor were:

‘Butterflies or sinking feeling in stomach,’
‘Heart pounds or flutters,’
‘Dry mouth, mouth coated,’
‘Dizziness, faintness, giddiness,’ and
‘Difficulty in getting breath, choking, tightness in chest.’

The success in defining a separate anxiety factor cannot be attributed to the method of analysis, because Varimax rotation has been employed in many other factorial analyses of psychopathology. Rather, the success is due to the inclusion in the interview schedule of the many kinds of situations which give rise to anxiety, and of the many physical ways in which anxiety may be manifested.
Reliability data

Three studies of the reliability of the mental state ratings were conducted. In one, 37 of the London patients were interviewed by one of the project’s psychiatrists with a second project member sitting in as an observer, making ratings but not asking questions himself. In a second study, 25 of the London patients were reinterviewed within a week of the first interview by a second psychiatrist who was ignorant of the first psychiatrist’s ratings. In the third study, the interviews of 24 of the New York patients were recorded on audiotape and rated by a different psychiatrist who was ignorant of the ratings made by the interviewer.

The intraclass correlation coefficients of interrater reliability for the depression and anxiety factors are presented in Table II. They are at least as high as those reported by others for factors combining depression and anxiety. The slight reduction in reliability upon repeated assessment is quite to be expected.

| TABLE II |
|-----------------|-----------------|-----------------|-----------------|
| **Intraclass correlation coefficients of interrater reliability** | | | |
| **Factor** | **Interviewer-Observer** | **Repeat interview** | **Audiotape** |
| | \(N=37\) | \(N=25\) | \(N=24\) |
| Depression | 0.97 | 0.88 | 0.93 |
| Anxiety | 0.92 | 0.86 | 0.91 |

To assess the added discrimination provided by the separation of depression and anxiety in distinguishing related diagnostic categories, patients variously diagnosed by the project team were compared on these two factors. Scores were standardized to have a mean of 50 and a standard deviation of 10 across the entire sample of 500 patients.

**Depressive versus anxiety neurosis**

Means for depressive and anxiety neurotics on the factors of depression and anxiety are presented in Figure 2. There was no difference of any practical magnitude on the depression factor \(t = 0.09, \text{n.s.}\), but a strongly significant difference on the anxiety factor \(t = 3.17, p < 0.001\). Anxiety neurotics seem to be distinguished by having greater levels of anxiety than depressive neurotics and not by having lower levels of depression.

**Schizoaffective psychosis versus other schizophrenias**

Mean standard scores for schizoaffectives and for all other schizophrenias are presented in Figure 3. The mean difference on the depression factor was statistically significant \(t = 2.65, p < 0.01\), but the difference on the anxiety factor was not \(t = 0.75, \text{n.s.}\). When the other schizophrenics were subdivided by type of schizophrenia, it was found that the schizoaffectives still had the highest mean on depression, the differences being significant at the 0.05 level or beyond between the schizoaffectives and all but the hebephrenic schizophrenics. None of the
Fig. 2. Mean standard scores for depressive and anxiety neurotics.

differences on the anxiety factor between the schizoaffectives and the other subtypes of schizophrenia even approached statistical significance.

Fig. 3. Mean standard scores for schizoaffectives and other schizophrenics.

Neurotic versus psychotic depression

Figure 4 presents means for neurotic and psychotic depressives. The mean standard score on the depression factor was greater for the psychotic than for the neurotic depressives. The t-ratio for this factor and that for anxiety, however, were both approximately unity, indicating no significant difference on either
depression or anxiety. These two groups of depressives were distinguished, however, on other factors. These other differences will be discussed below.

![Graph showing standard scores for psychotic and neurotic depressives.](image)

*Fig. 4. Mean standard scores for psychotic and neurotic depressives.*

To summarize these findings briefly, both the depression and anxiety factors can be rated with an excellent degree of reliability. The depression, but not the anxiety, factor succeeded in distinguishing schizoaffectives from other schizophrenics; and the anxiety, but not the depression, factor succeeded in distinguishing anxiety from depressive neurotics. These diagnostic discriminations would have been diluted if depression and anxiety had been combined into a single factor.

The successful separation of depression and anxiety factors also has implications for the evaluation of treatments aimed at specified target symptoms. Antidepressants may now be evaluated on a factor of depression uncontaminated by anxiety variables, and tranquilizers may now be evaluated on a factor of anxiety uncontaminated by depression variables. The success of these evaluations will only be determined by controlled trials.

**THE CONTINUITY OF NEUROTIC AND PSYCHOTIC DEPRESSION**

There is a long-standing controversy regarding the possibility that neurotic depressions are qualitatively different from psychotic depressions. We already know that response to ECT treatment differentiates neurotic from psychotic depressives, but whether this relationship is continuous has not yet been determined. Suppose, for example, that there is but a single homogeneous distribution for depression, and suppose that the probability of a positive response to treatment is correlated with position on this continuum. Then, if psychotic and neurotic depressives are distinguished only by their being above or below an arbitrary cut-off point, the two types will indeed be found to be distinguished by their proportions of positive response to treatment.
In our own studies we postulated, as did Kendell (1968), that if the two disorders are actually qualitatively different, they should produce different patterns among the factors of psychopathology which emerged from our factor analysis. One way of demonstrating such a difference is the method of the discriminant function, which provides the best set of weights for maximizing the mean difference on a total weighted score between two groups. If the two groups are qualitatively different, there should be little or no overlap between the two groups on the distribution of the total discriminant scores, and a bimodal distribution of these discriminant scores should be produced.

Such a discriminant analysis was applied to 104 depressive patients from the London sample included in the original 500 patients referred to previously. We sought a linear combination of the factors which would best distinguish the 3 patients diagnosed as neurotic depressives from the 65 diagnosed as psychotic depressives. Included in the latter category were involutional melancholia, the depressed type of manic-depressive psychosis, and psychotic depressive reaction.

Before applying the multiple factor comparison afforded by the discriminant function, separate analyses for each factor were carried out. The neurotic depressives scored significantly higher than the psychotic depressives on only one factor, reported belligerence (including such reported behaviors as having fits of anger and having hit or attacked people), but they scored significantly lower than the psychotic depressives on three factors — observed restlessness (including such observed behaviors as marked agitation, fidgeting, and pacing), retarded movement (including such observed behaviors as a slowness of movement and shuffling gait), and somatic concerns (including such behaviors as preoccupation with physical complaints and the mistaken belief that an organ is diseased).

Each of the 104 depressives was given a score on the discriminant function by applying the derived set of optimal weights to his set of factor scores. The resultant distribution of discriminant function scores shown in Figure 5 is unimodal. The neurotic depressives tended to have low, and the psychotic depressives to have

---

**Fig. 5.** Observed and fitted distributions on discriminant function.
high discriminant function scores. Their mean difference was highly significant ($t = 6.89, p < 0.001$). The superimposed normal distribution indicates, however, that the variation from neurotic to psychotic depression may be continuous (Fleiss, 1970).

It should be remembered that we do not get the full spectrum of depression in the clinic, and that the distribution of scores just presented is for in-patients only. What would happen if we included the entire population of a given locality? If we regard the depression dimension as analogous to the dimension of intelligence, we might note that intelligence is normally distributed up to about 3.3 standard deviations below the mean. At that point a hump is often found in the curve (Roberts, 1950). The hump may be regarded as consisting primarily of lesional or pathological cases, while the milder cases may represent the lower end of the distribution of normal intelligence.

Since we have only the cases at one end of the distribution in depression, we can only guess at what the total distribution would look like. It may very well be that in the case of depression, too, we shall have the majority of people distributing themselves on a normal curve, while our in-patients will represent the hump on the curve consisting for the most part of lesional cases with a few ‘sad sacks’ who constitute the extreme end of the normal curve of depression. Only community surveys can reveal these cryptodepressions, i.e. depressed individuals who never come to treatment and remain outside of known psychopathology. By studying them, we may be able to understand the basic nature of depressive illness since we can contrast those who exhibit the syndrome, but do not succumb, with those who succumb and need help. Furthermore, methods for assessing the assets as well as the liabilities are needed to find out what enables some to overcome the depression or at least contain it.

IMPLICATIONS FOR FUTURE RESEARCH

It is quite clear from the analyses of the United States-United Kingdom Diagnostic Project (Gurland et al., 1971) of the present state of diagnosis that any comparison of studies of depressives based on clinical diagnosis is tenuous because of the non-comparability of diagnoses among clinicians. For research purposes, where the comparability of patients is essential, the clinical label of depression has so far been of little value, since no homogeneous group is described by the label. For this reason much of the literature on depression must be viewed skeptically.

To make studies more comparable, systematic structured interviews revealing objectively the dimensions of psychopathology are essential. By subjecting groups of patients possessing similar profiles of psychopathology to certain field and laboratory investigations, we can use an iterative method for relating sequentially interview profiles to test results. As a consequence both the interview and the field and laboratory tests will be improved as we move from the one to the other and, in addition, highly purified groups for research purposes will be created. Such culture-free techniques as evoked potentials may prove useful here.

In order to contain the entire spectrum of depression, we cannot depend on
treatment cases alone but must investigate an entire population to find those individuals who may exhibit amounts of depression equal to those exhibited by individuals who come for treatment, but who themselves never cross the threshold of a clinic or a therapist. A knowledge of why they fail to do so may unearth some of the hidden factors separating the ‘ill’ from the ‘well’ as well as reveal individuals who need treatment but for some reason do not get it.

SUMMARY

The current biometric approaches to depression have produced a systematic structured interview, yielding objective and reliable assessments of psychopathology. Further, its use has enabled us (1) to demonstrate that the reported differences between the United States and United Kingdom in the prevalence of depression is due more to differences in diagnostic practice than to differences in patient prevalence, (2) to separate factorially depression from anxiety, which permits better differentiation between various categories of patients, and (3) to provide further evidence for the essential continuity of the distribution of depression in patient populations. The descriptive part of our work sets the stage for the next step, etiology, a task in which we have made some progress biometrically but which will be discussed here by other panelists. We hope that our provision of objective description and measurement of psychopathology will facilitate future etiological research.

ACKNOWLEDGEMENT

The authors are grateful to Karen M. Olson for her editorial assistance in the preparation of this paper.

REFERENCES


