

## DIAGNOSIS OF ENDOGENOUS DEPRESSION

### Comparison of Clinical, Research and Neuroendocrine Criteria

BERNARD J. CARROLL, MICHAEL FEINBERG, JOHN F. GREDEN,  
ROGER F. HASKETT, NORMAN McI. JAMES, MEIR STEINER and JANET TARIKA

*Clinical Studies Unit and Mental Health Research Institute, Department of Psychiatry,  
University of Michigan, Ann Arbor, MI 48109 (U.S.A.)*

(Received 17 March, 1980)

(Accepted 17 March, 1980)

---

#### SUMMARY

Eighty-nine depressed outpatients were studied by clinical criteria, Research Diagnostic Criteria (RDC), and the dexamethasone suppression test (DST) of neuroendocrine regulation. A simple outpatient version of the DST, requiring only one blood sample, correctly identified 40% of patients diagnosed clinically as endogenous depression (ED), with a specificity of 98% and a diagnostic confidence of 95%. Differences in age, sex, or severity of symptoms between endogenous and non-endogenous depressives did not account for these results. By comparison, the diagnostic performance of the DST was weaker for the RDC categories Major Depressive Disorder (MDD) and primary MDD. These were less selective and more heterogeneous than the clinical category ED. The clinical diagnoses of ED were supported in 98% of cases by the RDC, but 22% of RDC endogenous MDD diagnoses were not supported by the clinical diagnoses. Abnormal DST results were found only in patients with both the clinical diagnosis of ED and the RDC diagnosis of endogenous MDD. Patients with definite endogenous MDD had a significantly higher frequency of abnormal DST results (42%) than those with probable endogenous MDD (14%), or those with other RDC diagnoses (3%). A significant association was found between positive DST results and a positive family history of depression. These results support other evidence for use of a positive DST result as an external validating criterion for ED. The category MDD contained all cases diagnosed clinically as ED, but was diluted by cases diagnosed clinically as non-endogenous depression who had no neuroendocrine disturbance. The results also confirmed that the endogenous/non-endogenous and primary/secondary classifications of depression are not identical.

We conclude: (1) that the DST can be used in the differential diagnosis of depressed outpatients as well as inpatients; (2) that the RDC category primary MDD and the

---

This work was supported by USPHS Grant MH 28294, by research funds from the State of Michigan Department of Mental Health, and by the Mental Health Research Institute.

Washington University category primary depression are more heterogeneous and probably less valid than the clinical category ED; (3) that the RDC for endogenous MDD have only moderate validity; (4) that RDC diagnoses cannot substitute for careful clinical diagnoses in research studies, (5) that the best use of the RDC is to support clinical diagnoses, but not to generate diagnoses independently as a free-standing system; (6) that the concept of endogenous or endogenomorphic depression has validity and should be retained in research studies of depression.

---

## INTRODUCTION

The most persistent issue in the long history of debate about the nosology of depression has been to discriminate illnesses with a presumed biological basis from those which are related primarily to psychological or social factors. The former group are frequently termed endogenous, endogenomorphic, melancholic, autonomous or vital depressions in various classificatory systems, whereas the latter group tend to be described as reactive, neurotic or characterological depressions. The differential diagnosis of patients with depressed mood has until recently been based on clinical information alone, without the assistance of laboratory tests commonly used in other areas of medicine, and it is obvious from the current state of the field that no clinical system of classification has gained a general consensus. Furthermore, some current nosologic proposals in the United States, such as the Washington University criteria (Feighner et al. 1972), have adopted an 'agnostic', atheoretical approach which avoids the issue of endogenous depression altogether (Woodruff et al. 1974).

In clinical practice disagreements about whether patients have endogenous or non-endogenous depression do not arise with typical cases but with patients who may have only mild endogenous depression (Watts 1957), or a mixture of both endogenous and 'neurotic' features. As Kendell (1976) emphasized recently, 'the differences — in symptomatology, premorbid personality, treatment response, and lifetime course — between the two extremes are too extensive to be regarded merely as differences in severity and chronicity'. The real problem is to define an agreed boundary between them, and the various clinical checklist approaches to defining this boundary, such as the Newcastle Index (Carney et al. 1965) or the Research Diagnostic Criteria (RDC) (Spitzer et al. 1977) are not generally accepted because they lack operational validity in the hands of several investigators (McConaghy et al. 1967; Kendell 1968; Nelson et al. 1978; Feinberg et al. 1979). The Newcastle Index relies on a discriminant function which contrasts endogenous with neurotic features, and to a certain extent it implies that patients with endogenous depression will lack neurotic or characterologic symptoms. The RDC simply require the presence of at least 4 of 11 classical endogenous symptoms for a 'probable' diagnosis and 6 symptoms for a 'definite' diagnosis of endogenous Major Depressive Disorder. Both approaches suffer from the problem recognised long ago by Gillespie (1929), namely, that clinical features alone are not 'going to serve as a touchstone

for diagnosis, or prognosis, or as a therapeutic indication'. Even more importantly, both schemes pay no attention to certain clinical features that are usually weighted highly by clinicians, such as family history, bipolar course, and (in the case of the RDC) previous episodes.

A laboratory procedure to assist in identifying the type of depression would clearly have wide clinical and research application, and several promising advances in this area have appeared recently (Carroll 1980). Four years ago Carroll et al. (1976a,b) reported that the dexamethasone suppression test (DST) of neuroendocrine regulation correctly identified almost 50% of inpatients with endogenous depression at a high level of confidence. They also demonstrated that, in principle, a simplified version of the inpatient DST could be used for the diagnosis of depressed outpatients (Carroll and Curtis 1976). This report describes a confirmation of that prediction and compares several current nosologic proposals with the results obtained by the DST.

#### PATIENTS AND METHODS

We evaluated eighty-nine outpatients presenting consecutively to the Clinical Studies Unit (CSU) who received a clinical diagnosis of either endogenous depression (47 patients) or non-endogenous depression (42 patients). Each patient was seen by a psychiatrist for a standard clinical assessment which usually included an interview with a close relative. Each completed a structured psychiatric interview, the Schedule for Affective Disorders and Schizophrenia (SADS) (Spitzer and Endicott 1978), which was administered by a trained social worker, research assistant or psychiatrist. Other sources of information included in the diagnostic evaluation were the patient's previous psychiatric history, family history and past hospital records. The psychiatrist and the SADS interviewer then met for a diagnostic conference where items of information were cross-checked before consensus diagnoses were reached. Diagnoses were recorded as (1) clinical diagnosis and (2) RDC diagnosis. Most patients were rated for the severity of depressive symptoms by the Hamilton rating scale (Hamilton 1960) and by the Carroll self-rating scale (Feinberg et al. 1979). The latter is a new self-rating instrument designed to match closely the items contained in the clinician-rated Hamilton scale (Carroll et al. 1973).

The clinical features employed in the CSU for the diagnosis of endogenous depression are listed in Table 1. Rather than adopt a fixed number of features as a diagnostic criterion we require that the illness should be perceived by the psychiatrist as consistent with the syndrome ordinarily recognized as endogenous depression or melancholia. This requirement of 'pattern recognition' by the psychiatrist is a well recognized element in the performance of skilled diagnosticians (Dudley 1968). At the same time, of course, variations in this skill are also a major source of diagnostic disagreements among clinicians. As will be shown, however, our clinical diagnoses of

TABLE 1

Diagnostic features of endogenous depression

*Major*

- History of mania, hypomania or endogenous depression
- Definite family history
- Severe agitation or retardation
- Depressive psychosis
- Pervasive anhedonia
- Definite pathological guilt

*Minor*

- Depressed or dysphoric mood
- Middle or delayed insomnia
- Loss of appetite or weight
- Retardation or agitation
- Unrealistic self-reproach
- Anhedonia, not pervasive (includes decreased sex drive)
- Lack of reactivity
- Distinct quality of mood
- Diurnal variation
- Age at onset over 40 (unless bipolar)
- Duration more than 1 month and less than 1 year
- Response to previous treatments

endogenous depression were supported by the RDC in 46 of 47 cases (98%). In making the diagnoses of endogenous depression we did not necessarily give negative weight to features such as lack of adequate personality, apparent psychogenesis, or anxiety, which are weighted negatively by the Newcastle Index. Rather, we regard such 'neurotic' features as belonging to an independent clinical dimension, in keeping with Kiloh et al. (1972), Klein (1974) and Akiskal et al. (1977, 1979, 1980). The clinical features in Table 1 are quite similar to those listed by Shopsin et al. (1976). They also include all the RDC items for the endogenous subtype of Major Depressive Disorder (MDD).

The diagnoses of non-endogenous depression were based not only on the absence of endogenous features but also on the clinical assessment of the illness as a whole. The 42 non-endogenous cases all presented with a major complaint of dysphoric mood and other associated symptoms. In 13 of the 42 cases some RDC endogenous features were present, sufficient to warrant an RDC diagnosis of probable or definite endogenous MDD, if the RDC were used simply as a checklist. Despite the suggestion of an endogenous functional shift, however, the psychiatrists did not perceive these illnesses as consistent with endogenous depression. The most frequent reasons for the disagreement of our clinical and RDC diagnoses of these patients were that, although they satisfied the literal operational RDC for certain endogenous items (such as anhedonia, self-reproach, psychomotor retardation or distinct

quality of mood), their presentations of these features were felt to be unreliable and/or unconvincing. At the same time it should be noted that the diagnoses of non-endogenous depression were *not* based on personality type, manipulative, dependent or histrionic features, understandable psychological precipitants, or inferred unconscious conflicts.

Patients in whom a pre-existing psychiatric disorder was diagnosed (before the onset of the current depression) were classified as endogenous or non-endogenous by the psychiatrists on the basis of their current clinical profiles, without reference to their previous character disorders, neuroses, or alcohol and drug abuse. This decision is consistent with the position of the Washington University group that the endogenous/non-endogenous and primary/secondary classifications are independent of each other (Woodruff et al. 1974). Although diagnoses according to the Washington University criteria are not presented as such in the Results, it should be noted that all patients who met the RDC for MDD also met the Washington University criteria for depression.

Patients were excluded from this study if they had a previous or current diagnosis of schizophrenia, drug-related psychosis or organic brain disease. Patients taking certain drugs or suffering from medical conditions known to affect the dexamethasone suppression test also were excluded (see Carroll 1980). Over 80% of the patients were completely drug-free for at least 7 days before the DST.

A simplified version of the dexamethasone suppression test was given before treatment began. The patients took an oral dose of 1 mg dexamethasone tablets at 11:30 p.m. and came to the hospital for a single blood sample for plasma cortisol at 4 p.m. the next day. Except for the visit to the hospital they followed their usual schedule of work or activity on the day of the test. They also collected a 24-h urine sample from the time of dexamethasone administration. The results of urinary free cortisol measurements added very little to the information given by the plasma cortisol results, and have been reported in detail elsewhere (Carroll et al. 1980a). The criterion of an abnormal DST response was a plasma cortisol level greater than 6  $\mu\text{g}/\text{dl}$  at 4 p.m. post-dexamethasone. The test performance was evaluated with the calculations of sensitivity, specificity, and predictive value (diagnostic confidence) as defined by Vecchio (1966) and Galen and Gambino (1975). Plasma cortisol was measured by the competitive protein binding method of Murphy (1967) with minor modifications as described by Carroll et al. (1976b). The laboratory samples were assayed without knowledge of the clinical diagnosis, and the results were not released to the clinicians until the diagnostic evaluations were completed.

## RESULTS

### *Clinical diagnoses, age, sex, severity*

Endogenous depression was diagnosed in 47 patients, of whom 29 were

unipolar and 18 bipolar. Twenty of the unipolar patients were experiencing at least their second episode of endogenous depression, while 9 were seen in their first episode. Eight bipolar patients had a past history of mania, while 10 had a history of hypomania. Non-endogenous depression was diagnosed in 42 patients. The majority (32) received a clinical diagnosis of neurotic depressives ( $t = 2.65, P < 0.01$ ). The two groups did not differ with respect to a character disorder (5), obsessive compulsive neurosis (4) or anxiety neurosis (1).

Seventy percent of the endogenous depressives were female, as were 60% of the non-endogenous depressives. The mean age of the 47 endogenous depressives was 40.9 yr, compared with 32.7 yr for the 42 non-endogenous depressives ( $t = 2.65, P < 0.01$ ). The two groups did not differ with respect to the rated severity of depressive symptoms. The mean Hamilton rating scores were 18.2 for the endogenous cases and 16.2 for the non-endogenous cases. The mean Carroll self-rating scores were 23.4 and 23.5, respectively. Hamilton ratings were available for 74% of all patients, and Carroll ratings for 81% of all patients. These subgroups with available ratings did not differ from the total group with respect to age, sex ratios, or frequency of abnormal DST results.

#### *RDC diagnoses*

The RDC diagnoses of the endogenous and non-endogenous cases are given in Tables 2 and 3, respectively. An RDC diagnosis of probable or definite endogenous Major Depressive Disorder was recorded in 46 of the 47 endogenous cases (98%). In 43 endogenous cases the depression was primary according to the RDC, while in 4 cases it was secondary to obsessive compulsive disorder (2), drug use disorder (1) or a significant medical disorder (advanced loss of vision due to diabetic retinopathy) (1).

Twenty-five of the 42 non-endogenous cases (60%) met the RDC for Major Depressive Disorder and 13 (31%) also qualified for the endogenous RDC subtype (definite or probable). Most of these 25 patients diagnosed as MDD by the RDC were primary cases (19 of 25, or 76%). Six cases were secondary to obsessive compulsive disorder (1), phobic disorder (1), alcohol-

TABLE 2

RDC diagnoses of 47 cases diagnosed clinically as endogenous depression

Major depressive disorder (MDD)	47
endogenous (definite)	32
endogenous (probable)	14
non-endogenous	1
Primary MDD	43
Secondary MDD	4

TABLE 3

RDC diagnoses of 42 cases diagnosed clinically as non-endogenous depression

<i>Major depressive disorder</i>	25	<i>Minor depressive disorder</i>	13
endogenous (definite)	6		
endogenous (probable)	7		
non-endogenous	12	<i>Other</i> <sup>a</sup>	4
		obsessive compulsive disorder	1
		generalized anxiety disorder	1
Primary MDD	19	labile personality disorder	1
		other psychiatric disorder	1
Secondary MDD	6		

<sup>a</sup> All with dysphoric mood.

ism (1), drug use disorder (1) or a combination of alcoholism and drug use disorder (2). Thirteen non-endogenous cases received the RDC diagnosis of minor depressive disorder, while miscellaneous other RDC diagnoses were recorded in 4 cases.

### *Comparison of clinical and RDC diagnoses*

#### *(1) Major depressive disorder*

Among the total group of 89 patients the clinical diagnosis of endogenous depression was made only 47 times, while the RDC diagnosis of MDD was found appropriate much more often (72 cases). These results suggest that patients diagnosed as having Major Depressive Disorder by the RDC (or depression by the Washington University criteria) are likely to be a distinctly heterogeneous group when viewed from other clinical perspectives, as was emphasized also by Nelson et al. (1978) in their study of inpatients. Even the category primary MDD was diagnosed in 62 cases by the RDC and, as predicted by Woodruff et al. (1974), a large proportion of these cases (31%) were regarded as non-endogenous by the clinicians. These findings thus confirm that the endogenous/non-endogenous and primary/secondary classifications are not identical. Of the 47 cases diagnosed clinically as endogenous, 43 (91%) were called primary MDD by the RDC, but so were 19 of the 42 non-endogenous cases (45%).

#### *(2) Endogenous depression*

By contrast to the rather broad criteria for diagnosing MDD (or depression in the Washington University system), the RDC subtype of endogenous MDD requires a more selective profile of symptoms. The comparison of the clinical and RDC diagnoses of endogenous depression is given in Table 4, which indicates that the two diagnostic systems agreed in 84% of the cases. Both probable and definite cases of endogenous MDD have been included in this comparison. This decision is in keeping with the fact that the clinical diagnoses also covered a range of diagnostic confidence from probable to

TABLE 4

Comparison of clinical and RDC diagnoses of endogenous depression

RDC diagnosis	Clinical diagnosis	
	Endogenous depression	Other
Endogenous major depressive disorder <sup>a</sup>	46	13
Other	1	29

Kappa 0.68.

Diagnoses agree in 84% of cases.

22% of RDC endogenous cases are not endogenous by clinical diagnosis.

<sup>a</sup> Includes probable and definite cases.

definite. As stated above, the clinical diagnoses of endogenous depression were confirmed by the RDC in all but one case. On the other hand, the RDC diagnosed endogenous MDD in 59 cases, compared to only 47 clinical diagnoses of endogenous depression. Thus, in 13 of 59 cases (22%) the RDC diagnosis of endogenous MDD was not supported by the clinical diagnosis. These results suggest that even the more restrictive RDC endogenous criteria are likely to yield a heterogeneous group of subjects for research studies.

#### *Dexamethasone suppression test*

Using the plasma cortisol criterion of 6  $\mu\text{g}/\text{dl}$ , we found that 19 of the 47 cases (40%) diagnosed clinically as endogenous depression had an abnormal DST result. By contrast, 41 of the 42 patients (98%) with non-endogenous depression had normal test results (Chi square 18.4,  $P < 0.0001$ ). Based on these findings, the DST showed a sensitivity of 40%, with a specificity of 98%. Thus the predictive value (diagnostic confidence) of an abnormal or positive DST result was 95%. That is, 19 of the 20 positive test results were true-positive for endogenous depression. On the other hand, the predictive value of a negative or normal DST result was only 59% for non-endogenous depression. The frequency of abnormal DST results in the various subgroups of patients with endogenous depression was: unipolar, first episode, 33%; recurrent unipolar 40%; bipolar with previous mania 63%; bipolar with previous hypomania 30%; all unipolar 38%; all bipolar 44%. There was no effect of sex on the DST results. Positive results were found in 36% of the males and 42% of the females with the clinical diagnosis endogenous depression. Age also had no effect on the DST results. Among the 47 endogenous depressives the correlation of age with post-dexamethasone plasma cortisol was not significant ( $r = -0.14$ ). By age cohorts the frequency of positive DST results was: 15–24 yr ( $n = 9$ ) 78%; 25–34 yr ( $n = 13$ ) 23%; 35–44 yr



( $n = 5$ ) 40%; 45–54 yr ( $n = 10$ ) 40%; 55–64 yr ( $n = 7$ ) 29%; 65–84 yr ( $n = 3$ ) 33%. Thus the difference in frequency of abnormal DST results between the endogenous and non-endogenous depressives was not explained by the significantly greater mean age of the endogenous group.

In 41 of the 47 cases with endogenous depression we had adequate information about the family psychiatric history. Six cases were rejected because the patients were adoptees or because the sources of information were considered incomplete, based on blind chart review. The frequency of abnormal DST results was 43% in the remaining 41 cases, compared with 40% in the total group of 47 endogenous depressives. Of the 41 informative cases, 30 (73%) had a definite or probable family history of depression. A positive DST result was found in 16 of these 30 cases (53%) with a positive family history, compared with 2 of 11 cases (18%) with a negative family history (Fisher exact test,  $P 0.047$ ). This association of positive family history with positive DST result was seen both in unipolar cases ( $n = 25$ ) and bipolar cases ( $n = 16$ ).

The positive DST results by RDC categories are shown in Table 5, which includes for comparison the clinical diagnoses as well. It can be seen that most of the RDC categories had a lower frequency of abnormal DST results than did the clinical endogenous depression group. All 20 patients with positive DST results were found in the RDC category Major Depressive Disorder but the frequency of neuroendocrine disturbance was only 28% in this

TABLE 5

Frequency of abnormal DST results in diagnostic subgroups

Diagnosis	N	Positive DST	% Positive
<i>(1) Clinical</i>			
Endogenous depression	47	19	40
Non-endogenous depression	42	1	2
<i>(2) RDC</i>			
Major depressive disorder (MDD)	72	20	28
Primary MDD	62	19	31
Secondary MDD	10	1	10
Endogenous MDD (definite and probable)	59	19	32
Endogenous MDD (definite)	38	16	42
Endogenous MDD (probable)	21	3	14
Primary endogenous MDD (definite and probable)	51	18	35
Minor depressive disorder	13	0	0
Other	4	0	0

group. All 17 patients with RDC diagnoses other than MDD had negative DST results, as did 9 of the 10 cases with secondary MDD. The one positive DST result with secondary MDD occurred in a patient diagnosed clinically as endogenous depression, who also met endogenous RDC criteria, and who had a serious progressive medical condition (advanced visual failure due to diabetic retinopathy). No positive DST results were found in 9 patients with RDC diagnoses of MDD secondary to other psychiatric disorders. The RDC category definite endogenous MDD had a frequency of abnormal DST results slightly above the 40% figure found in the patients diagnosed clinically as endogenous depression. However, only 16 of the 20 positive tests were found in patients with definite endogenous MDD, compared with 19 in the group diagnosed clinically as endogenous depression.

The frequency of abnormal DST results was distinctly lower in the patients with only a probable endogenous MDD diagnosis (14%) than in those with a definite diagnosis (42%) (Fisher exact test,  $P$  0.026) and it was only 3% in those with neither definite nor probable endogenous MDD. Thus, if a high frequency of abnormal DST results is used as an external validating criterion, the RDC category definite endogenous MDD is significantly more valid than the category probable endogenous MDD, which in turn is more valid than the category not endogenous MDD. However, the category definite endogenous MDD is not more valid than the clinical category endogenous depression. When the analysis above was restricted to the cases confirmed as endogenous by the clinicians, then the frequency of abnormal DST results was: definite endogenous MDD 50%; probable endogenous MDD 21%; not endogenous MDD 2%.

The diagnostic performance of the DST for the major categories of interest is shown in Table 6. The overall performance was no better for any of the RDC categories than it was for the clinical category endogenous depression. The RDC categories showed a lower sensitivity or predictive value for the DST. The lowering of sensitivity, as in the case of Major Depressive Disorder, suggests that a more heterogeneous group of patients was selected by these criteria, although all the cases with neuroendocrine disturbance were identified. The lowering of specificity and predictive value, as in the case of definite endogenous MDD, suggests that this RDC diagnosis is more restrictive than the clinical diagnosis of endogenous depression, by excluding more cases with neuroendocrine disturbance, even though the sensitivity was slightly higher at 42%. The well known balance or 'trade-off' between sensitivity and predictive value is illustrated clearly by these results (see Galen and Gambino 1975).

When the disagreements between the clinical and RDC diagnoses of endogenous depression (Table 4) were examined with respect to the DST results, the value of the DST as an independent validating criterion was reinforced. Table 7 shows the same data as Table 4, with the additional information of positive DST results in each of the four cells. The results in Table 7 indicate that 95% of the abnormal DST results occurred in patients whose clinical diag-

TABLE 6  
 Diagnostic performance of DST according to clinical and RDC categories

Category	N	Prevalence %	Positive DST	Sensitivity %	Specificity %	Predictive value %
<i>Endogenous depression (clinical)</i>	47	53	19	40	98	95
<i>RDC</i>						
Major depressive disorder (MDD)	72	81	20	28	100	100
Primary MDD	62	70	19	31	96	95
Endogenous MDD definite	38	43	16	42	92	80
probable and definite	59	66	19	32	97	95
Primary endogenous MDD (probable and definite)	51	57	18	35	95	90

TABLE 7

Clinical and RDC diagnoses of endogenous depression in relation to DST results  
Number of abnormal DST results indicated in parentheses.

RDC diagnosis	Clinical diagnosis	
	Endogenous depression	Other
RDC diagnosis		
Endogenous major depressive disorder	46(19)	13(0)
Other	1(0)	29(1)

nosis of endogenous depression was also confirmed by the RDC. Thus, the neuroendocrine disturbance was strongly associated with the concordant diagnoses of endogenous depression. By contrast, there were no abnormal DST results among the 13 cases diagnosed as endogenous by the RDC, but about which the clinicians disagreed. Among the 59 patients diagnosed as endogenous MDD the distribution of abnormal DST results was significantly related to the clinical diagnosis (Fisher exact test,  $P < 0.003$ ). For the 13 cases about whom the clinicians disagreed with the RDC diagnosis, definite endogenous MDD was the RDC diagnosis in 6 patients, while probable endogenous MDD was the RDC diagnosis in 7 patients. Thus, the low rate of abnormal DST results in these 13 cases cannot be explained by their being only probable RDC endogenous cases. Only one abnormal DST result was found among the 29 patients diagnosed as non-endogenous by both the clinicians and the RDC. This low frequency (3%) is the same as we have found in non-patients (Carroll et al., unpublished observations) and appears to represent the background rate of nonspecific false-positive DST results in the general population.

## DISCUSSION

The results of this outpatient study are consistent with our previous studies of inpatients (Carroll et al. 1976b). The DST correctly identified 40% of the endogenous depressed patients at a confidence level of 95%. Our findings are supported by two recent studies of other investigators. Brown et al. (1979) found a sensitivity of 40% and a predictive value of 100% for the DST in a group of 20 inpatients with 'primary endogenomorphic' depression, compared with 29 inpatients with other diagnoses. Similarly, Yiannis et al. (1980) found a DST sensitivity of 43% and a predictive value of 100% in 16 inpatients with 'melancholic' depression, compared with 22 other patients. By extending the work to outpatients in this study we have expanded the utility of this diagnostic test, since the majority of depressed patients are evaluated and treated in the outpatient setting. The high specificity of the

test could not be attributed to a lower severity of depressive symptoms in the non-endogenous cases, nor to age and sex differences. Rather, it was a function of the categorical distinction between endogenous and non-endogenous depressions.

The test procedure we used is simple enough to be implemented by any treatment facility, and most clinical laboratories can provide the plasma cortisol assays. The elimination of a urine collection (Carroll et al. 1980a) is an advantage both for the patients and for the laboratory, since we had found that great care was needed in the accurate measurement of low urinary free cortisol values. By using only a single post-dexamethasone blood sample some loss of sensitivity will occur. Based on our studies of inpatients (from whom post-dexamethasone blood samples were obtained at 8 a.m., 4 p.m. and 11:30 p.m.) we can state that 80% of the positive DST results will be detected by the 4 p.m. sample (Carroll et al., unpublished observations). We have opted for this 20% loss of sensitivity in the interest of developing a practical DST procedure for outpatient use. This decision is consistent with the fact that a positive DST result has a much higher predictive value or diagnostic significance (95%) than does a negative DST result (59%). Thus, a positive DST result is strongly associated with the diagnosis endogenous depression, whereas a negative DST result will not necessarily rule out this diagnosis.

Our results indicate that the clinical category endogenous depression is not entirely congruent with the RDC category Major Depressive Disorder (or the Washington University category of depression). Even the RDC category primary MDD was not congruent with the clinical category endogenous depression: 31% of the primary MDD cases were regarded as non-endogenous by clinical diagnosis. This finding is of course consistent with the explicit prediction of the Washington University group (Woodruff et al. 1974), but it has not been demonstrated previously by the actual study of patients.

The extent of agreement between clinical diagnoses, RDC diagnoses and Washington University diagnoses will not necessarily be consistent in different treatment settings. For example, Kupfer and associates found that virtually all inpatients with an RDC diagnosis of primary MDD also satisfied RDC endogenous criteria, whereas the majority of their secondary depressives did not (D. Kupfer, personal communication). For this study, however, the inclusion criteria specified that all patients should remain seriously depressed beyond a 2-week inpatient evaluation period. Non-endogenous primary depressives did not meet this criterion and thus were excluded from the sample. In an outpatient treatment setting, however, non-endogenous primary depressives are quite common, as we have shown. Similarly, Schlessler et al. (1979) reported DST results very similar to our own in 86 inpatients with primary unipolar depression (by Washington University criteria) and, like ourselves, a very low rate of positive DST results among secondary depressives (G. Winokur, personal communication). It seems

likely, then, that there would be a high degree of congruence between 'primary depression' and endogenous depression among inpatients at Iowa City. The logic of the two classificatory systems does not require a high degree of congruence, however, and the extent of agreement may be quite variable from one center to another. We would expect the greatest discordance between diagnoses of primary and endogenous depression in out-patient studies, in community surveys, and in studies of symptomatic volunteers who respond to newspaper advertisements for depression research studies.

In attempting to compare the validity of various clinical and research classifications we have used the DST as an independent validating criterion. The rationale for this approach comes from several lines of evidence. An abnormal DST result has *face validity* in view of the high specificity and predictive value, confirmed now by several independent groups (Carroll et al. 1976b, 1980a; Brown et al. 1979; Schlessner et al. 1979; Yiannis et al. 1980). The significant increments in sensitivity between the 'absent', 'probable' and 'definite' endogenous MDD categories (Table 5) also are *prima facie* evidence for the use of the DST as an independent criterion of endogenous depression. *Predictive validity* is derived from observations that a conversion of abnormal DST responses to normal is associated with good prognosis after treatment, while failure of the DST response to convert to normal is associated with early relapse (Carroll 1972; Goldberg 1980; Greden et al. 1980a). Predictive validity is derived also from the findings that an abnormal DST response is associated with a significantly increased rate of response to antidepressant treatments (Brown et al. 1979; Greden et al., unpublished observations). Finally, *construct validity* for this approach is derived from several studies. DST responses return to normal with complete recovery, or with a switch into mania (Carroll 1972; Carroll et al. 1976a). A gradual improvement in DST responses occurs when the test is performed serially during a course of antidepressant drug treatment or electrotherapy (Carroll 1972; Carroll et al. 1976a; Dysken et al. 1979; Alcala and Greden 1980; M. Fink, personal communication). The association with a positive family history reported above is further evidence for construct validity of abnormal DST responses, and this association was found also by Schlessner et al. (1979). Construct validity also derives from studies which suggest that the neuroendocrine disturbance revealed by the DST indicates dysfunction in the limbic system, the presumed site of pathology in the biological or endogenous depressions (Rubin and Mandell 1966; Carroll et al. 1968; Carroll 1972; Sachar et al. 1973).

The obvious limitation of the DST as an independent validating criterion is that its sensitivity is only moderately high (about 45%). While this figure is comparable to that of some other diagnostic tests used in internal medicine (Carroll et al. 1980a), it is still clear that a consistently accurate laboratory diagnosis of endogenous depression will not be achieved by the DST alone. The heterogeneity of DST results among endogenous depressives

is so far unexplained. It may be related to genetic factors (alcoholism and sociopathy) (Schlesser et al. 1979) or to certain clinical features (Carroll et al. 1980b). Despite this limitation there is good evidence to support the validity of an abnormal DST response as an independent criterion for nosology. The prediction would be that the validity and homogeneity of a clinical diagnostic category are related to the frequency of abnormal DST results detected in that category.

When our results were examined in light of this prediction they indicated that none of the RDC categories was superior to the clinical category endogenous depression. In particular, the RDC categories MDD and primary MDD had a lower frequency of positive DST results. The conclusion from these data is that the MDD and primary MDD categories are more heterogeneous than the clinical category endogenous depression. Only the RDC category definite endogenous MDD was as powerful as the clinical endogenous depression category with respect to the frequency of abnormal DST results. In this case, however, there was a loss of specificity and predictive value.

From our results we would suggest that the RDC for endogenous MDD can be misleading if they are used merely as a checklist to generate a diagnosis among unselected depressed outpatients. Indeed, 22% of cases identified as endogenous MDD were considered non-endogenous by clinical diagnosis. Furthermore, the DST results strongly validated the clinical diagnoses rather than the RDC diagnoses of these cases. As described earlier, diagnostic disagreements arose with these patients because discriminating diagnosticians considered several of their apparently endogenous features to be unreliable, or unconvincing, or qualitatively different from those associated with definite endogenous depression. We stress, however, that these patients did satisfy the operational SADS-RDC criteria for endogenous MDD. We conclude that the RDC, as presently formulated, do not identify endogenous depression with the same validity as experienced clinicians do.

On the other hand, 98% of the clinical diagnoses of endogenous depression were confirmed by the RDC (Table 4). The best use of diagnostic checklists like the RDC, therefore, may be to support clinical diagnoses rather than to generate diagnoses independently. Use of the RDC as a system of research documentation of clinical diagnoses can facilitate communication among clinical research centers, but there is no evidence available to support the notion that RDC diagnoses could substitute for discriminating clinical diagnoses in research studies. Our results indicate that there is an impressive concordance of neuroendocrine disturbance only with concordant clinical *and* RDC diagnoses of endogenous depression (Table 7).

This need for clinical judgement in making research diagnoses was emphasized recently by another group of investigators (Carpenter et al. 1980): They stressed that their research diagnostic system for schizophrenia cannot be used as 'a free-standing system for identifying . . . patients or for differential diagnosis (but) rather . . . for ascertaining the confidence level of a diagnosis . . . and for defining a . . . cohort with reliable criteria under-

standable to other clinicians and investigators. It *must* be embedded in a broad clinical diagnostic evaluation'. When using the SADS and RDC instruments we have insisted (in the interest of reliability) on strict accordance with the operational definitions. Indeed, the very purpose of such instruments would only be vitiated if clinicians were free to impose the filter of clinical judgement on individual patients when making SADS-RDC evaluations. To the extent that this does occur (but is not acknowledged) in clinical research studies, statements that operationally strict 'research' diagnoses have been used are unrealistic. While disagreements between the diagnoses endogenous depression and MDD, or primary MDD, in the present study were largely definitional in origin, those between endogenous depression and endogenous MDD were related precisely to the point made by Carpenter et al. (1980). The RDC as a 'free-standing' system diluted the clinical category endogenous depression with cases that were not validated by the DST.

A final conclusion from our results is that the concept of endogenous depression deserves to be retained in research studies. The 'agnostic' categories of primary depression in the Washington University system and Major Depressive Disorder in the RDC are distinctly more heterogeneous not only from clinical perspectives but also with respect to neuroendocrine function. While the primary/secondary proposal of the Washington University group may possibly have some utility, it was unfortunately joined with a loose set of diagnostic criteria for depression. These criteria in turn were the direct precursors of those for the RDC category MDD. By avoiding the historically difficult problem of diagnosing endogenous depression it was hoped that these new categories would be more reliable. Our results indicate, however, that they are more heterogeneous and probably less valid than the clinical category endogenous depression. Their unquestioned use as research criteria might thus be viewed with serious reservations. Carey and Gottesman (1978) recently presented an extensive discussion of this problem of reliability -v- validity, and our results bear out several of their *caveats* about 'reliable criteria'. Our work with the DST is one approach towards the development of independent laboratory criteria for subtypes of depression defined by pathophysiology in addition to clinical features. As these biological approaches to the nosology of depression are developed further (Carroll 1980) the 'old fashioned' concept of endogenous or endogenomorphic depression is likely to be reinstated rather than discarded.

#### ACKNOWLEDGEMENTS

We thank Anne Commorato and Anne Wissler, Psychiatric Social Workers for their careful help with the clinical research evaluations, and our laboratory chiefs, K. John Jarrett and James C. Ritchie, for their reliable collaboration.



## REFERENCES

- Akiskal, H.S., Djenderedjian, A.H., Rosenthal, R.H. and Khani, M.K., Cyclothymic disorder — Validating criteria for inclusion in the bipolar affective group, *Amer. J. Psychiat.*, 134 (1977) 1227—1233.
- Akiskal, H.S., Rosenthal, T.L., Haykal, R.F., Lemmi, H., Rosenthal, R.H. and Scott-Strauss, A., Characterological depressions — Clinical and sleep electroencephalographic findings separating 'subaffective dysthymias' from 'character-spectrum disorders', *Arch. Gen. Psychiat.*, 137 (1980) 383.
- Akiskal, H.S., Rosenthal, R.H., Rosenthal, T.L., Kashgarian, M., Khani, M.K. and Puzantian, V.R., Differentiation of primary affective illness from situational, symptomatic, and secondary depressions, *Arch. Gen. Psychiat.*, 36 (1979) 635—643.
- Albala, A.A. and Greden, J.F., Serial dexamethasone suppression tests (DST) in affective disorders, *Amer. J. Psychiat.*, 137 (1980) 383.
- Brown, W.A., Johnston, R. and Mayfield, D., The 24-hour dexamethasone suppression test in a clinical setting — Relationship to diagnosis, symptoms and response to treatment, *Amer. J. Psychiat.*, 136 (1979) 543—547.
- Carey, G. and Gottesman, I.I., Reliability and validity in binary ratings, *Arch. Gen. Psychiat.*, 35 (1978) 1454—1459.
- Carney, M.W.P., Roth, M. and Garside, R.F., The diagnosis of depressive syndromes and the prediction of ECT response, *Brit. J. Psychiat.*, 111 (1965) 659—674.
- Carpenter, W.T., Strauss, J.S and Bartko, J.J., Diagnostic systems and prognostic validity, *Arch. Gen. Psychiat.*, 37 (1980) 228—229.
- Carroll, B.J., The hypothalamic-pituitary-adrenal axis in depression. In: B. Davies, B.J. Carroll and R.M. Mowbray (Eds.), *Depressive Illness — Some Research Studies*, Ch.C. Thomas, Springfield, IL, 1972, pp. 23—201.
- Carroll, B.J., Implications of biological research for the diagnosis of depression. In: J. Mendlewicz (Ed.), *New Advances in the Diagnosis and Treatment of Depressive Illness*, Elsevier, Amsterdam, In press.
- Carroll, B.J. and Curtis, G.C., Neuroendocrine identification of depressed patients, *Aust. New Zealand J. Psychiat.*, 10 (1976) 13—20.
- Carroll, B.J., Curtis, G.C. and Mendels, J., Neuroendocrine regulation in depression, Part 1 (Limbic system-adrenocortical dysfunction), *Arch. Gen. Psychiat.*, 33 (1976a) 1039—1044.
- Carroll, B.J., Curtis, G.C. and Mendels, J., Neuroendocrine regulation in depression, Part 2 (Discrimination of depressed from non-depressed patients), *Arch. Gen. Psychiat.*, 33 (1976b) 1051—1057.
- Carroll, B.J., Martin, F.I.R. and Davies, B., Resistance to suppression by dexamethasone of plasma 11-OHCS levels in severe depressive illnesses, *Brit. Med. J.*, 3 (1968) 285—287.
- Carroll, B.J., Fielding, J.M. and Blashki, T.G., Depression rating scales — A critical review, *Arch. Gen. Psychiat.*, 28 (1973) 361—366.
- Carroll, B.J., Feinberg, M., Steiner, M., Haskett, R.F., James, N.McI. and Tarika, J., Diagnostic application of the dexamethasone suppression test in depressed outpatients. In: J. Mendlewicz (Ed.), *Advances in Biological Psychiatry*, Karger, Basel, 1980a, In press.
- Carroll, B.J., Greden, J.F., Feinberg, M., James, N.McI., Haskett, R.F., Steiner, M. and Tarika, J., Neuroendocrine dysfunction in genetic subtypes of primary unipolar depression, *Psychiat. Res.*, (1980) In press.
- Dudley, H.A.F., Pay-off, heuristics, and pattern recognition in the diagnostic process, *Lancet*, ii (1968) 723—726.
- Dysken, M.W., Pandey, G.N., Chang, S.S., Hicks, R. and Davis, J.M., Serial postdexamethasone cortisol levels in a patient undergoing ECT, *Amer. J. Psychiat.*, 136 (1979) 1328—1329.
- Feighner, J.P., Robins, E., Guze, S.B., Woodruff, R.A., Winokur, G. and Munoz, R.,

- Diagnostic criteria for use in psychiatric research, *Arch. Gen. Psychiat.*, 26 (1972) 57-63.
- Feinberg, M., Carroll, B.J., Steiner, M. and Commorato, A., Misdiagnosis of endogenous depression with the Research Diagnostic Criteria, *Lancet*, i (1979) 267.
- Feinberg, M., Carroll, B.J., Smouse, P., Rawson, S., Haskett, R.F., Steiner, M., Albala, A. and Zelnick, T., Comparison of physician and self-ratings of depression. Abstract No. 80, Annual Mtg. Society of Biological Psychiatry, Chicago, 1979.
- Galen, R.S. and Gambino, S.R., *Beyond Normality — The Predictive Value and Efficiency of Medical Diagnoses*, Wiley, New York, NY, 1975.
- Gillespie, R.D., The clinical differentiation of types of depression, *Guy's Hospital Rep.*, 79 (1929) 306-344.
- Goldberg, I.K., Dexamethasone suppression test as indicator of safe withdrawal of anti-depressant therapy, *Lancet*, i (1980) 376.
- Greden, J.F., Albala, A.A., Haskett, R.F., James, N.McI., Goodman, L., Steiner, M. and Carroll, B.J., Normalization of dexamethasone suppression test — A laboratory index of recovery from endogenous depression, *Biol. Psychiat.*, 15 (1980) 449-458.
- Hamilton, M., A rating scale for depression, *J. Neurol. Neurosurg. Psychiat.*, 23 (1960) 56-62.
- Kendell, R.E., *The Classification of Depressive Illness*, Oxford University Press, London, 1968.
- Kendell, R.E., The classification of depressions — A review of contemporary confusion, *Brit. J. Psychiat.*, 129 (1976) 15-28.
- Kiloh, L.G., Andrews, G., Neilson, M. and Bianchi, G.N., The relationship of the syndromes called endogenous and neurotic depression, *Brit. J. Psychiat.*, 121 (1972) 183-196.
- Klein, D.F., Endogenomorphic depression — A conceptual and terminological revision, *Arch. Gen. Psychiat.*, 31 (1974) 447-454.
- McConaghy, N., Joffe, A.D. and Murphy, B., The independence of neurotic and endogenous depression, *Brit. J. Psychiat.*, 113 (1967) 479-484.
- Murphy, B.E.P., Some studies of the protein-binding of steroids and their application to the routine micro and ultramicro measurement of various steroids in body fluids by competitive protein-binding radioassays, *J. Clin. Endocrinol. Metab.*, 27 (1967) 973-990.
- Nelson, J.C., Charney, D.S. and Vingiano, A.W., False-positive diagnosis with primary-affective-disorder criteria, *Lancet*, ii (1978) 1252-1253.
- Rubin, R. and Mandell, A., Adrenal cortical activity in pathological emotional states — A review, *Amer. J. Psychiat.*, 123 (1966) 387-400.
- Sachar, E.J., Hellman, L., Roffwarg, H., Halpern, F., Fuskushima, D. and Gallagher, T., Disrupted 24-hour patterns of cortisol secretion in psychotic depression, *Arch. Gen. Psychiat.*, 28 (1973) 19-24.
- Schlesser, M.A., Winokur, G. and Sherman, B.M., Genetic subtypes of unipolar primary depressive illness distinguished by hypothalamic-pituitary-adrenal axis activity, *Lancet*, i (1979) 739-741.
- Shopsin, B., Freidman, E. and Gershon, S., Parachlorophenylalanine reversal of tranylcypromine effects in depressed patients, *Arch. Gen. Psychiat.*, 33 (1976) 811-819.
- Spitzer, R.L. and Endicott, J., *Schedule for Affective Disorders and Schizophrenia*, New York State Psychiatric Institute, New York, NY, 1978.
- Spitzer, R.L., Endicott, J. and Robins, E., *Research Diagnostic Criteria (RDC) for a Selected Group of Functional Disorders*, 3rd edition, New York State Psychiatric Institute, New York, NY, 1977.
- Vecchio, T.J., Predictive value of a single diagnostic test in unselected populations, *New Engl. J. Med.*, 274 (1966) 1171-1173.
- Watts, C.A.H., The mild endogenous depression, *Brit. Med. J.*, 1 (1957) 4-8.
- Woodruff, R.A., Goodwin, G.W. and Guze, S.B., *Psychiatric Diagnosis*, Oxford Press, London, 1974.
- Yiannis, P., Lee, J., Johnson, L. and Fink, M., Neuroendocrine effects of ECT, *Society of Biological Psychiatry, Abstracts*, 1980.