EDITORIAL

The prospects for clinical psychoneuroendocrinology: has the curtain been drawn across the neuroendocrine window?1

'The new psychiatry' has been hailed as the emerging era of our speciality in which objective diagnostic laboratory tests contribute to a more exacting clinical practice (BMJ, 1981). Neuroendocrine tests figure prominently in this new psychiatry. These tests have been proposed as aids in differential diagnosis, in the selection of specific treatments, and in the evaluation of treatment success. Indeed, in a few psychiatric centres in the United States neuroendocrine tests already have become a routine part of the clinical evaluation. However, rather than their being established laboratory aids, the utility of these neuroendocrine tests in psychiatry remains controversial, and their promise remains to be fulfilled. Much of the recent adverse publicity about neuroendocrine tests in psychiatry stems from conflicting research reports. Many of these reports, both the original favourable ones and the more recent conflicting ones, have methodological differences which contribute to their disparate conclusions. A review of some aspects of the methodology of these studies on the 'neuroendocrine window' to the brain should facilitate a sober reappraisal of the place of neuroendocrine testing, both in the elucidation of altered brain-endocrine physiology and in the clinical assessment of psychiatric illness.

Hypothalamo-pituitary-adrenocortical (HPA) hyperactivity in depression, and its measurement by the Dexamethasone Suppression Test (DST), can serve as a convenient paradigm for this appraisal. Some 15 years of research preceded the major report by Carroll et al. (1981) which, in its title, portrayed the DST as 'a specific laboratory test for the diagnosis of melancholia'. This report has been followed by a plethora of studies on the DST in depression and other psychiatric illnesses—Index Medicus catalogued some 40 in 1982, 100 in 1983, and 50 in the first six months of 1984. A few of these studies have focused on the physiological mechanism of an abnormal DST, i.e. at which level in the HPA axis the dysregulation occurs, but most have considered how the DST might be clinically relevant. Both of these aspects will be considered from a methodological perspective (Rubin & Poland, 1984a).

While one group of investigators has implicated the adrenal cortex itself in an abnormal DST, based on discordant post-dexamethasone adrenocorticotropic hormone (ACTH) and cortisol patterns (Fang et al. 1981), the consensus is that ACTH is indeed elevated in cortisol escapers compared with suppressors (Reus et al. 1982; Kalin et al. 1982; Yerevanian et al. 1983; Holsboer et al. 1984a). Of particular importance in these studies are the frequency of blood sampling for ACTH, which is secreted in short bursts, and the validity of the ACTH radioimmunoassays, since all immunoreactive ACTH may not be biologically active (Schöneshofer et al. 1981). Furthermore, studies of the ACTH response to synthetic corticotropin releasing factor (CRF) in depressives indicate that the corticotrophs of the pituitary gland are either normal or somewhat suppressed by an assumed increase in hypothalamic production of endogenous CRF (Holsboer et al. 1984b; Gold et al. 1984).

The occurrence of adrenocortical hyperactivity in depression, as determined by increased plasma and CSF cortisol concentrations and urinary cortisol excretion, has been known for years (Rubin & Mandell, 1966; Carroll, 1976). It also appears that these measures of adrenal cortical hyperactivity and an abnormal DST might not share a common physiological mechanism, since they sometimes are not highly correlated in depressed patients. However, this imperfect correlation also can be explained on methodological grounds – in particular, sources of variance in the DST itself (see

1 Address for correspondence: Professor Robert T. Rubin, Division of Biological Psychiatry, Harbor–UCLA Medical Center, Torrance, CA 90509, USA.
In the absence of compelling data to suggest a special mechanism for adrenocortical hyperactivity in depressive illness as reflected by an abnormal DST, i.e. a functional defect at the pituitary and/or adrenal level, increased limbic–hypothalamic stimulation of CRF production, putatively on the basis of altered neurotransmitter function, must remain the most tenable explanation (Carroll, 1976; Asnis et al. 1981).

An interesting clinical question concerns the resultant metabolic effects, if any, of HPA axis hyperactivity in depression. Cushing’s disease has been proposed as a ‘psychosomatic’ illness (Gifford & Gunderson, 1970), and a continuum between depression with HPA hyperactivity and frank Cushing’s disease has been suggested (Reus, 1984). However, while Cushing’s patients often show depressive symptoms (Starkman et al. 1981), even severely depressed patients with elevated circulating cortisol rarely manifest overt stigmata of Cushing’s syndrome. Here, too, methodological issues are relevant. Most hypercortisolism in depression has been documented by the measurement of total (both free and protein-bound) serum or plasma cortisol concentrations. Saliva cortisol, being a diffusate of serum cortisol, accurately reflects the free, metabolically active fraction of serum cortisol (Riad-Fahmy et al. 1982; Landon et al. 1984) and can distinguish between true hypercortisolemia (elevated free fraction) and apparent hypercortisolemia (elevated bound fraction) (Evans et al. 1984). Saliva cortisol concentrations in depression, in the few studies conducted to date, do not appear to be increased into the Cushing’s range (Poland & Rubin, 1982; Copolov et al. 1985; Hanada et al. 1985).

What of the clinical utility of the DST? A host of methodological and other factors has been cited as contributing to between-study variability in DST results (Checkley & Rush, 1983). Some of the methodological factors are of major importance, such as dexamethasone dosage, frequency of cortisol measures, concurrent physical illness and/or medication usage and, especially, the clinical characteristics of the patients studied.

The most widely used dexamethasone dose has been 1·0 mg, given orally at 2300 h or 2400 h, based on the work of Carroll et al. (1981). Several-fold differences in serum concentrations of dexamethasone occur in patients after this fixed dose; low concentrations can yield ambiguous test results (Meikle, 1982). Refinements in the form (e.g. oral elixir), route (e.g. parenteral), and dose (e.g. based on body weight) of dexamethasone administration, as well as measurement of both dexamethasone and cortisol in serum, should considerably reduce this source of variance.

Collecting blood samples, particularly from out-patients, is often logistically difficult. Several groups of investigators have shown that, of the 8-, 16- and 24-hour post-dexamethasone serum cortisols, the 16-hour measure detects the most non-suppressors, with the 24-hour sample contributing a few more. Carroll et al. (1981) thus advocated a single 16-hour post-dexamethasone (late afternoon) blood sample for cortisol measurement and suggested a criterion value for non-suppression of 50 ng/ml, based on their laboratory technique (competitive protein binding assay). Radioimmunoassay for cortisol, both by commercially available kits and by individual laboratory protocols, is becoming the preferred measurement technique. This has necessitated a focus on laboratory aspects of the DST, such as quality control of assays and differing criteria for cortisol non-suppression (Rubin et al. 1980; Wilens et al. 1983; Meltzer & Fang, 1983; Rubin & Poland, 1984b). Perhaps the most unequivocal indicator of an abnormal DST is the pattern of post-dexamethasone cortisol values – initial suppression at 8 hours, indicating adequate bioavailability of dexamethasone, and an escape pattern thereafter. This, however, requires multiple blood sampling (inconvenient) and cortisol determinations (more costly), but it might prove to be essential, particularly since cortisol secretion resumes in an episodic manner (Sherman et al. 1984).

The issues of concurrent illness and medication usage by patients have been covered adequately in other reports (Carroll et al. 1981; Carroll, 1982; Holsboer, 1983). The lists of both grow inexorably longer. So does the list of patients with non-major depressive psychiatric illnesses who can exhibit positive DSTs – for example, manics, alcoholics, dementants, and schizophrenics. And, so does the list of intervening variables affecting the DST – for example, recent weight loss (Berger et al. 1983; Edelstein et al. 1983), and the stress of hospitalization (Coccaro et al. 1984).

A major methodological issue concerns the clinical characteristics of the patient samples studied to date. Determinations of the sensitivity and specificity of the DST have been conducted, inter alia,
on Research Diagnostic Criteria (RDC) (Spitzer et al. 1978) major depressives, RDC endogenous major depressives, and DSM-III (APA, 1980) melancholic major depressives versus other depressives, psychiatric patients, and normal controls. The necessarily conflicting results from these various patient samples have generated much debate about the predictive value of the DST. Comparative diagnostic studies have shown that these categories of depression are all different – for example, among 70 major depressives at the Royal Edinburgh Hospital, 35 were rated as RDC definitely endogenous, and among these 35 only 20 were rated as DSM-III melancholic (Copolov et al. 1985). Similarly, in an American sample of 50 depressives (46 with major depression) diagnosed according to several systems, 50% were RDC definitely endogenous, 40% were Newcastle scale endogenous, and only 30% were DSM-III melancholic (Davidson et al. 1984a). Thus, although criterion-based diagnostic systems are more reliable, different ones can produce quite different frequencies of the type of depression for which the DST is supposed to be a useful ancillary diagnostic aid.

Since no diagnostic system in psychiatry has yet been established as a ‘gold standard’, how do we untie this Gordian knot? Some investigators have used the DST as the benchmark and determined which clinical symptoms correlate with it (Brown & Shuey, 1980; Reus, 1982; Nasr et al. 1983; Copolov et al. 1985). This is a good idea, but requires sample sizes of several hundred patients for stable multivariate analysis, in addition to consistent and reliable raters for all the patients. Other investigators have taken a more pragmatic approach – in the absence of a valid benchmark they consider the use of multiple diagnostic systems for comparison with biological measures (Kendell, 1982; Kasper & Beckmann, 1983). Ultimately, of course, the DST and all other psychoneuroendocrine and biological tests, both separately and in combination, must be validated not only against diagnosis, even by multiple systems (Davidson et al. 1984b), but also against genetic and familial factors, treatment response, and long-term prognosis.

Concluding with our paradigm, is the DST, as an ‘objective diagnostic laboratory test’, and perhaps also other psychoneuroendocrine tests, destined for the same fate in our specialty as befall psychoanalytical treatment, namely having more restrictions on its use than indications? Only methodologically precise and carefully controlled studies will provide an answer to this question. And, in a more general sense, is the foregoing implied lament about the unfulfilled promise of ‘the new psychiatry’ an expression of unwarranted pessimism or a realistic concern? Only time, and many more research studies, will tell. For the moment, the curtain may have been drawn across the neuroendocrine window to the brain, but it is translucent, and we remain fascinated by the mysterious silhouette it still reveals.

ROBERT T. RUBIN

This study was supported by NIMH Research Scientist Award MH 47363 and NIMH Research Grant MH 28380.

REFERENCES


