How reliable is the short synacthen test for the investigation of the hypothalamic-pituitary-adrenal axis?

V Bangar and R N Clayton

Department of Diabetes and Endocrinology, City General Hospital, Stoke on Trent, Staffordshire ST4 6QG, UK and \(^1\)Department of Medicine, Keele University, Thornburrow Drive, Hartshill, Stoke on Trent, Staffordshire ST4 7QB, UK

Correspondence should be addressed to R N Clayton

Abstract

The best test for the assessment of the hypothalamic–pituitary–adrenal (HPA) axis remains a matter of controversy. We compared the performance of the short synacthen test (SST, 250 \(\mu\)g) with the insulin stress test (IST) to assess the reliability of the former as a first line test. Patients with pituitary disease underwent both the SST and the IST. The results in patients who had both tests within 4 weeks of each other, and where these were not separated by a therapeutic intervention, were compared. Basal, 30 and 60 min cortisol levels were obtained from the SST. Basal and maximal cortisol level after adequate hypoglycaemia (glucose < 2.2 mmol/l) were recorded for the IST.

Sixty-nine paired test results were available for analysis. With a 30 min ‘pass’ plasma cortisol value of 500 nmol/l on the SST, 7/69 (10%) patients who passed the SST failed the IST set at a ‘pass’ maximum value of 500 nmol/l. At a ‘pass’ cortisol value of 600 nmol/l on the SST, 3/69 (4%) who passed the SST failed the IST. Assuming the IST as the gold standard, the sensitivity of an SST ‘pass’ of 600 nmol/l is 85% with a specificity of 96%.

During the conventional dose SST (250 \(\mu\)g) a 30 min plasma cortisol value of 600 nmol/l is more reliable than a value of 500 nmol/l, and using the former criterion the SST can safely be used as a first line test for the evaluation of the HPA axis in patients with pituitary disease. However, if the result is borderline or there is clinical suspicion of mild hypocorticotrophism an IST or other test of the HPA axis may be warranted.

Introduction

In the investigation of function of the hypothalamic–pituitary–adrenal (HPA) axis, the insulin stress test (IST) is accepted as the gold standard investigation. Cortisol responses to hypoglycaemia were defined by Plumpton and Besser (1) and found to correlate well with the peri-surgical cortisol response. Accordingly, an adequate maximal hypoglycaemic cortisol response of 580 nmol/l, using the fluorometric assay of Mattingley, was adopted as an indicator of an intact pituitary adrenocortical axis. The test has been shown to be reliable, although it is unpleasant for patients and has resulted in loss of consciousness, arrhythmias, angina, myocardial infarction and death. It is contra-indicated in patients with cardiac disease, in extremes of age, and epilepsy. Close medical and/or nursing supervision is required and the test is relatively costly.

The short synacthen test (SST) has been advocated as an alternative to the IST (2–6), using 250 \(\mu\)g synthetic adrenocorticotropic hormone (ACTH) i.m. or i.v. This dose maximally stimulates the adrenal gland. The test relies on atrophy of the adrenal cortex due to removal of the trophic stimulus (ACTH) from the pituitary, as a consequence of hypothalamic pituitary disease or chronic suppression of the HPA axis from exogenous glucocorticoids. The atrophic adrenal cannot then respond to an acute ACTH challenge. In acute pituitary insufficiency (such as post hypophysectomy) the test is, however, unreliable (7, 8) since about 2 weeks is required for involution of the zona fasciculata (7). A good correlation has been demonstrated repeatedly between the IST and the SST such that the latter has enjoyed increasing popularity as a substitute for the IST. Twenty-four percent of UK endocrinologists adopted the SST as a first line test in 1988 (4) and 50 percent had adopted it by 1995 (9). Whilst the majority of studies show close correlation between the responses to the two tests (3, 4, 10), discrepancies have been demonstrated (11–15), leading some to rely on the IST as the first line test despite its unpleasantness. In our unit at present we use the SST as a first line investigation of the HPA axis. Previously, however, it was policy to perform both tests. Accordingly, we have the opportunity to analyse the results of those paired tests to determine the extent of any discrepancies.
Patients and methods

A retrospective analysis of results from patients who had assessment of the HPA axis by both the SST and IST was performed. Ninety-one pairs of tests were identified but 22 pairs were excluded for the following reasons: inadequate hypoglycaemia (12 subjects with nadir plasma glucose >2.2 mmol/l); more than 4 weeks between tests (6 subjects); administration of hydrocortisone on the day of the test (4 subjects). This left 69 pairs of tests for analysis in patients with a diagnosis of pituitary adenoma pre- or post-treatment (54); craniopharyngioma (8); idiopathic hypopituitarism or empty sella (6); optic nerve germinoma (1). In none of the patients were the two tests separated by a therapeutic intervention.

The test pairs were analysed for concordance using a ‘pass’ on the SST as a 30 min plasma cortisol value of more than 500 nmol/l or more than 600 nmol/l. For the IST it was necessary that the maximum cortisol level was more than 500 nmol/l for it to be classified as a ‘pass’. A plasma cortisol value of 500 nmol/l is equivalent to 580 nmol/l (1) since the old fluorimetric assay has a 20–30% positive bias compared with radioimmunoassay of cortisol (16) over the 300–900 nmol/l range of plasma cortisol (15). Plasma cortisol was measured by the assay currently in use at the time which was either a radioimmunoassay (Amerlex, Aylesbury, Bucks, UK) or after 1990 a non-isotopic immunoassay (Abbott TDX, Diagnostics Division, IL, USA) with intra-assay and interassay coefficients of variation of 3–4% and 4–7% respectively.

Sensitivity is defined as the number of true positive results×100/number of all cases with deficiency, and specificity is defined as the number of true negative results×100/number of all cases without deficiency.

Results

The results of the 69 pairs of tests are shown in Table 1.

With the ‘pass’ set at 500 nmol/l for the SST, 7 patients passed the SST but failed the IST. The IST maximum cortisol levels in these patients were 496, 496, 473, 396, 362, 313 and 140 nmol/l. If the SST ‘pass’ value is increased to 600 nmol/l, then the number of patients passing the SST but failing the IST decreases to three (3/69 = 4%) and the IST peak cortisol levels for these were 496, 496 and 313 nmol/l. Assuming that the IST is the gold standard test, the sensitivity of the SST at a ‘pass’ value of 600 nmol/l was 17/20 (85%) and the specificity was 47/49 (96%). However, there is no guarantee that the IST is the most appropriate gold standard even though it tests the integrity of the entire HPA axis. Indeed, in the report by Tsatsoulis et al. (17) in children with multiple pituitary hormone deficiencies, failure to achieve a cortisol value of 500 nmol/l in the IST was reported in two patients who on subsequent IST testing easily achieved this value. Although the intra-individual reproducibility of the cortisol response to the IST is reasonable (coefficient of variation 10–15%) in normal adults, this has not been assessed in patients with pituitary disease (18). Thus, for example, if an SST ‘pass’ of >600 nmol/l were accepted as the gold standard, and this is reasonable based on responses of acute medical emergencies (19), then the sensitivity of the IST in our study is 17/19 (89%) and the specificity is 47/50 (94%), values not different from the converse situation. Therefore, the IST and the SST have similar sensitivities and specificities for the detection of true deficiency of the HPA axis. Where the tests were concordant, i.e. pass/pass, or fail/fail the SST could substitute for the IST, that is in 64/69 (93%) of tests. A fail on the SST and a ‘pass’ on the IST (2/69 = 3%) does not expose the patient to any risks.

Table 1 Outcomes of 250 μg SST compared with IST using two definitions of SST ‘pass’ (of 500 or 600 nmol/l). In all patients nadir plasma glucose during IST <2.2 mmol/l.

<table>
<thead>
<tr>
<th>Number of Patients</th>
<th>SST IST*</th>
<th>SST pass &gt;500 nmol/l</th>
<th>SST pass &gt;600 nmol/l</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>Fail</td>
<td>13</td>
<td>17</td>
</tr>
<tr>
<td>Fail</td>
<td>Pass</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Pass</td>
<td>Fail</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Pass</td>
<td>Pass</td>
<td>47</td>
<td>47</td>
</tr>
</tbody>
</table>

* IST pass at peak cortisol of 500 nmol/l.

Table 2 Published comparisons of 250 μg SST with IST cortisol responses in patients with suspected pituitary-adrenal insufficiency.

<table>
<thead>
<tr>
<th>Study</th>
<th>Reference number</th>
<th>SST pass (&gt;600 nmol/l) and IST fail (&lt;500 nmol/l)/ total no. of subjects</th>
<th>False reassurance rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lindholm et al. (1987)</td>
<td>(3)</td>
<td>4/200</td>
<td>2</td>
</tr>
<tr>
<td>Stewart et al. (1988)</td>
<td>(4)</td>
<td>1/70</td>
<td>1.5</td>
</tr>
<tr>
<td>Hurel et al. (1996)</td>
<td>(10)</td>
<td>6/166</td>
<td>3.6</td>
</tr>
<tr>
<td>Ammari et al. (1996)</td>
<td>(12)</td>
<td>4/30</td>
<td>13.2</td>
</tr>
<tr>
<td>Mukerjee et al. (1997)</td>
<td>(8)</td>
<td>0/21</td>
<td>0</td>
</tr>
<tr>
<td>Present study</td>
<td></td>
<td>3/69</td>
<td>4</td>
</tr>
</tbody>
</table>
risk since it triggers the IST or other test of the HPA axis. We have examined recently published comparative reports identically, taking an IST ‘pass’ as a peak cortisol value of >500 nmol/l, and an SST ‘pass’ as a cortisol value of >600 nmol/l (Table 2). All but one study showed a false reassurance rate of <5% by using the SST.

**Discussion**

For decision making regarding hydrocortisone substitution in individual patients a test with 100% sensitivity is desirable, and the IST has been accepted as such. However, even this is not totally reliable since it failed to reveal partial ACTH deficiency in six patients who had symptoms of cortisol deficiency after pituitary radiotherapy (17). The IST is therefore an acceptable alternative as shown in this and other studies. The clinically important discordance is when the SST implies adequate adrenal reserve but the IST shows a suboptimal cortisol response, in either case false reassurance is given by the SST and the patient does not receive substitution therapy.

The false reassurance rate for the SST varies. At best, in the series of Stewart et al. (4), the one patient with discordant results passed the SST at 600 nmol/l but had an IST maximum of 491 nmol/l which is arguably within the variance limit of the assay. In the majority of studies the false reassurance rate with the SST is acceptably low at <4%. However, caution is required since Ammari et al. (12), in a small series of 30 cases, reported 4 patients who passed the SST at 600 nmol/l but failed the IST. The maximum IST values reported in these patients were 484, 467, 435 and 372 nmol/l. Their study shows a discordance rate of 13.2%.

The IST ‘pass’ value was originally defined as the minimum cortisol response required to safely undergo a stressful event (elective surgery) i.e. 580 nmol/l (1). Since then, the ‘pass’ value has been modified to account for the increased specificity of modern cortisol assays, a value of 500 nmol/l being adopted as being equivalent to 580 nmol/l (4, 10, 15, 16). This is the minimal cortisol response that is deemed safe for individuals to undergo maximal stress such as surgery. If the IST is adopted as a first line substitute for the IST, most patients would only have an SST. Under those circumstances, with our figures 3 people out of 69 would be in danger of a suboptimal cortisol response during severe stress.

The SST is likely to be unreliable in the first two weeks after pituitary surgery (8). These authors compared the SST with IST both pre-operatively and within two weeks of pituitary surgery. Using their criteria of a ‘pass’ in both tests of >580 nmol/l, 4/21 (19%) of patients pre-operatively and 4/16 (25%) of patients post-operatively had discordant results. However, taking the criteria we have used of an IST ‘pass’ of >500 nmol/l and an SST ‘pass’ of >600 nmol/l none of the patients who achieved 600 nmol/l on the SST post-operatively failed the IST. At a ‘pass’ value of 500 nmol/l 30 min post SST, 4/21 (20%) of subjects failed the IST and would have been falsely reassured. However, 10/21 (50%) of the post-operative SSTs produced 30 min cortisol values of <600 nmol/l and in these subjects an IST or other test of the HPA axis would be required. In another study of 12 patients in whom SST and IST were compared 4–18 days post hypophysectomy all 5 who achieved a 30 min cortisol value >600 nmol/l with the SST achieved >500 nmol/l with the IST (20). Thus, even early after pituitary surgery the SST may provide an adequate and safe assessment of the HPA axis if a conservative value of >600 nmol/l is adopted as a ‘pass’ value.

Therefore the results of the present study and review of the published series indicate that an SST ‘pass’ defined by a 30 min cortisol value >600 nmol/l is safer for clinical decision making than lower values (i.e. 500–550 nmol/l). Moreover, all 50 elderly patients with unselected medical emergencies who had an SST within 24 h of admission achieved 30 min cortisol values >600 nmol/l (range 602–2265 nmol/l) validating this criterion as safe during acute illness (19).

A recent report has concluded that both low dose (1 μg) and standard dose (250 μg) SSTs were equally sensitive for the diagnosis of moderate to severe secondary adrenal insufficiency when compared with IST, metyrapone, or corticotrophin-releasing hormone tests (21). However, in subjects defined as having ‘mild’ cortisol deficiency, with a response to IST of between 500–550 nmol/l or to metyrapone of 179–200 nmol/l, both doses of synacthen produced ‘normal’ or ‘pass’ responses in 8/9 subjects, implying false reassurance would be given. However, there are at least two problems with this interpretation. First, it is known that the intra-individual variation of the IST is at least 10% in normal individuals (18) so if retested on a subsequent occasion they may well achieve a peak cortisol of >550 nmol/l and therefore qualify as a ‘pass’. The intra-individual variability of the metyrapone test has not been reported. Secondly, if the authors take 500 nmol/l as the ‘pass’ value on IST, as we have argued is logical with current cortisol assays, then there were only one or two discrepancies and these were on the side of caution, i.e. they are failures of the proposed first-line test (SST) rather than the gold standard IST and would therefore trigger an IST or alternative test.

When considering a test a considered approach needs to be taken balancing side effects, contra-indications, resource requirement and cost as well as efficacy. Accepting that no test is perfect in all these respects, we feel that the SST can continue to be used as a first line test. However, its fallibility needs to be remembered and any clinically doubtful cases investigated further by the IST or another stimulation test e.g. the glucagon or metyrapone test.

Accordingly, at a cut off value of 600 nmol/l at 30 min, the 250 μg SST can be used as the first line test for investigation of the HPA axis in patients with pituitary disease.
References


