CLINICAL STUDY

Depth and duration of hypoglycaemia achieved during the insulin tolerance test

O Ajala, H Lockett, G Twine and D E Flanagan
Department of Endocrinology, Derriford Hospital, Level 9, Plymouth PL6 8DH, UK and 1Royal Devon and Exeter Hospital, Exeter, UK

(Correspondence should be addressed to O Ajala; Email: olubukola.ajala@nhs.net)

Abstract

Context: The insulin tolerance test (ITT) is the gold standard for assessment of the pituitary adrenal axis but its use is limited because of concerns relating to the risk of hypoglycaemia.

Objective: This study examined the depth and duration of hypoglycaemia achieved during the test in a large cohort of patients.

Design: Two hundred and twenty ITTs were performed from 2005 to 2010.

Setting: A 1200-bed University Teaching Hospital.

Patients: Two hundred and twenty ITTs were carried out in patients with suspected or known pituitary disorders.

Interventions: Intravenous insulin was administered to achieve nadir plasma glucose (NPG) of 2.2 mmol/l (39.6 mg/dl). Blood chemistry to show the cortisol and GH response to hypoglycaemic stress was measured.

Main outcome measures: Predictors of depth and duration of hypoglycaemia, adverse events and within-subject variability of nadir glucose, peak cortisol and peak GH were studied.

Results: Thirty percent of the cohort achieved a nadir glucose of $<2.0$ mmol/l (36 mg/dl) that lasted for 60 min or more. The NPG correlated positively with fasting plasma glucose (FPG; $r=0.56; P<0.0005$), insulin dose ($r=0.27; P<0.0005$) and weight ($b=0.21; P<0.004$). The within-subject variability of nadir glucose was 15.2%, peak cortisol was 11.7% and peak GH was 6.4%. The factors determining nadir blood glucose were FPG ($b=0.56, P<0.0005, 20\%$ contribution) and weight ($b=0.14, P<0.05, 2\%$ contribution). The five patients with adverse events had NPG and insulin dose comparable with the rest of the population.

Conclusions: The hypoglycaemia achieved during the ITT is much lower than the target required. However, adverse events are few and do not relate to the depth of hypoglycaemia.

Introduction

Pituitary adenomas together with a variety of other sella-based abnormalities account for $\sim 10–15\%$ of intracranial tumours (1). With improvement in pituitary imaging, the recognition and diagnosis of pituitary abnormalities is increasing (2). It is part of the routine work of the endocrine clinic to assess pituitary hormone function in each of these individuals. An accurate test of the hypothalamic–pituitary–adrenal (HPA) axis is an essential part of this assessment (1, 3, 4, 5). In some cases, for example, after pituitary radiotherapy, annual repeat testing may be required (6, 7, 8). Undiagnosed hypoadrenalism is associated with significant morbidity and mortality (5, 9). Conversely inappropriate replacement with glucocorticoid may be associated with an increased risk of osteoporosis, weight gain, hypertension and diabetes mellitus together with increased cardiovascular risk (10). An assessment of the GH axis is also important as a proportion of adult patients with hypopituitarism will also benefit from GH replacement (11, 12, 13).

The gold standard test for the assessment of these two axes is the insulin tolerance test (ITT) (3, 12, 13, 14, 15, 16, 17). It was developed in the 1960s using an insulin bolus to produce hypoglycaemia. It was found that if blood glucose was lowered to 2.2 mmol/l, the cortisol response was similar to that seen following major surgery (18). The modern ITT protocol used in most endocrine units is essentially unchanged from the original description (1, 3, 11, 19, 20). An i.v. insulin bolus based on body weight is used to produce a rapid onset of hypoglycaemia. Blood glucose is monitored at the bedside using a capillary glucose meter. The depth of hypoglycaemia is later confirmed in the laboratory on the basis of whole blood or plasma glucose measures. Protocols vary as to the depth of hypoglycaemia necessary for the test to be considered adequate. Most units quote either 2.2 mmol/l (39.6 mg/dl), 2.5 mmol/l (45 mg/dl) or 2.8 mmol/l (50.4 mg/dl), although it is
not always clear whether whole blood or plasma is being used for this definition (3, 14, 15, 17, 19, 20). The advantage of the ITT is that hypoglycaemia is a reproducible stressor that tests the integrity of the whole HPA axis (3, 16, 17, 18). The response to this hypoglycaemic stress is vital when considering the need for corticosteroid replacement.

The disadvantages of the test include the unpleasantness of profound hypoglycaemia for the individual, the need for close monitoring by experienced clinicians and the perceived dangers of the test (14, 15, 21). The ITT is contraindicated in individuals with a history of epilepsy or at risk of seizures and those with a history of ischaemic heart disease. Many centres extrapolate this to exclude elderly patients and some centres do not perform the test in children due to previous case reports of fatalities (21).

The decision on how best to assess the HPA axis involves balancing the risk of the test against the risks of a less than accurate diagnosis of HPA dysfunction.

The purpose of this study was to define the duration, depth and risks associated with insulin-induced hypoglycaemia and also the factors that predict nadir blood glucose.

**Materials and methods**

Two hundred and twenty ITTs were performed on the programmed investigations unit at Derriford Hospital, a 1200-bed referral centre, between January 2005 and June 2010. Exclusion criteria for the test were evidence of ischaemic heart disease, history of epilepsy or unexplained blackouts and patients who declined the test.

Weight was measured to calculate insulin dose and venous blood samples were taken for glucose, GH, cortisol, TSH, free thyroxine (FT4), LH, FSH, prolactin, insulin-like growth factor 1 (IGF1), testosterone (in males) and oestradiol (in females). Venous plasma glucose was measured in a clinical laboratory on a Siemens Immulite 2000 (Siemens medical Solutions, B Bagsvaerd, Denmark) was administered. The insulin dose was determined by clinical judgement based on the presence of severe obesity, known cortisol deficiency, acromegaly, Cushing’s syndrome or insulin resistance.

A dose of 0.15 IU/kg was used in the majority of patients (185/218). The dose was 0.3 U/kg if the patient was considered to be insulin resistant (26/218) and 0.1 U/kg if thought to be insulin sensitive (7/218).

The tests were terminated when clinical features of hypoglycaemia were present and by confirmation from the biochemistry laboratory that nadir plasma glucose (NPG) was below 2.2 mmol/l. The patient was then fed a supervised mixed meal or administered i.v. glucose if hypoglycaemia lasted >30 min. Verbal and written consent were obtained from all patients.

**Statistical analysis**

Of the test values examined (weight, insulin dose per kilogram of body weight, fasting plasma glucose (FPG), GH and cortisol) all except baseline and peak GH followed a normal distribution.

Pearson and Spearman analyses were used for correlations and considered significant if P was ≤0.05. An unpaired t-test was performed when comparing groups of normally distributed values and Mann–Whitney U test for those values not normally distributed.

The variability between individual responses in repeat tests (within-subject variability) was calculated according to the method described by Bland & Altman (27).

**Results**

The summary characteristics and outcome of the ITTs are shown in Table 1.

The major indication in our cohort was non-functioning pituitary macroadenomas, half of whom had undergone pituitary surgery. The ‘other’ group comprised patients with sellar/parasellar lesions following radiotherapy for childhood leukaemia and congenital disorders causing GH deficiency. The age range of the population was 17–88 years with 22 (10%) patients over the age of 70 and four (2%) patients 18 years or younger.
Most of the tests involved the use of the 'standard' dose of insulin (0.15 U/kg; Table 2). About one-third of the tests done in patients with previous Cushing's disease involved using the sensitive dose (0.1 U/kg); all of these patients were post-operative and considered to be steroid deficient.

Seventy percent of the patients with acromegaly had the standard dose of insulin. All of these were either after pituitary surgery or radiotherapy, so considered to have normal insulin sensitivity.

Figure 1 shows the bedside and laboratory venous plasma glucose measured at 30-min intervals over the course of the tests.

Thirty percent of the cohort had glucose of <2.0 mmol/l (36 mg/dl) for 60 min or more. The patients in this group were characterised by a lower body weight (86 vs 93 kg, \(P<0.05\)), a lower total insulin dose (14.4 vs 23.5 U, \(P<0.001\)), lower baseline GH (0.4 vs 0.6 mmol/l, \(P<0.001\)), lower GH response (peak GH 3.8 vs 6.4 mmol/l, \(P<0.0001\)) and lower fasting glucose (4.8 vs 5.2 mmol/l, \(P<0.001\)) compared with the rest of the cohort. Peak cortisol was, however, about the same in both groups (570 vs 460 nmol/l).

The patients with a nadir glucose of <1 mmol/l (18 mg/dl) had a lower body weight, lower total insulin dose and lower fasting glucose (79 vs 89.5 kg, \(P<0.05\); 13 vs 17 U, \(P<0.01\); and 4.6 vs 5.1 mmol/l, \(P<0.006\) respectively) compared with the rest of the cohort.

Determinants of nadir glucose
Using multiple regression analysis, the independent factors determining nadir blood glucose were fasting venous glucose \(b=0.56, P=0.0005, 20\%\) contribution) and weight \(b=0.14, P<0.05, 2\%\) contribution).

### Table 1: Patient characteristics and outcome (s.d.; median).

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>NFPA</th>
<th>Acromegaly</th>
<th>Prolactinomas</th>
<th>Cushing's</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>220</td>
<td>83</td>
<td>41</td>
<td>15</td>
<td>9</td>
<td>72</td>
</tr>
<tr>
<td>Age (years)</td>
<td>50.5 ± 17</td>
<td>58.3 ± 15</td>
<td>54.5 ± 13</td>
<td>49.5 ± 14</td>
<td>42.6 ± 20</td>
<td>42.5 ± 16</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>87 ± 21</td>
<td>90 ± 22</td>
<td>87 ± 19</td>
<td>85 ± 29</td>
<td>94 ± 21</td>
<td>84 ± 19</td>
</tr>
<tr>
<td>Post-pituitary surgery</td>
<td>83/220</td>
<td>40/83</td>
<td>25/41</td>
<td>1/15</td>
<td>9/9</td>
<td>8/72</td>
</tr>
<tr>
<td>Post-radiotherapy</td>
<td>48/220</td>
<td>24/83</td>
<td>6/41</td>
<td>3/15</td>
<td>1/9</td>
<td>14/72</td>
</tr>
<tr>
<td>Insulin dose (U/kg)</td>
<td>0.17 (0.05)</td>
<td>0.17 (0.05)</td>
<td>0.21 (0.07)</td>
<td>0.15 (0.0)</td>
<td>0.13 (0.03)</td>
<td>0.15 (0.03)</td>
</tr>
<tr>
<td>Baseline plasma glucose</td>
<td>5.0 (0.9)</td>
<td>5.0 (1.1)</td>
<td>5.4 (1.0)</td>
<td>4.9 (0.5)</td>
<td>4.6 (1.0)</td>
<td>4.9 (0.7)</td>
</tr>
<tr>
<td>NPG</td>
<td>0.1–4.6 (1.5)</td>
<td>0.1–4.3 (2.2)</td>
<td>0.4–3.8 (1.8)</td>
<td>0.9–3.4 (1.6)</td>
<td>1.1–3.9 (2.4)</td>
<td>0.4–3.6 (1.4)</td>
</tr>
</tbody>
</table>

| Baseline cortisol   | 474       | 341       | 381        | 367           | –         | 360    |
| Post-pituitary surgery | 318    | 279       | 281        | 315           | 363       | 295    |
| Baseline GH         | 2.0 (5.3) | 0.36 (0.45)| 8.3 (10.6) | 3.0 (3.1)     | –         | 4.2 (4.9)  |
| Post-pituitary GH   | 1.1 (3.7) | 0.70 (1.10)| 2.7 (7.4)  | 3.3 (5.4)     | 0.2 (0.2)  | 0.9 (1.6)  |
| Peak GH             | 6.1 (8.9) | 3.9 (4.5) | 16.4 (15.7) | 3 (3.1)       | –         | 4.2 (4.9)  |
| Target NPG achieved | 89        | 90        | 83         | 100           | 87.5      | 89     |
| Requiring ‘2nd dose’ of insulin (%) | 12  | 9         | 17        | 0             | 12.5      | 12     |
| Achieved peak cortisol | 65    | 64        | 75        | 100           | 44        | 59     |
| >550 mmol/l (%)     | 103/137   | 27/43     | 16/16      | 7/14          | –         | 30/64   |
| Proportion of participants with normal GH response | 21/83 | 31/40     | 13/25      | 1/1           | 5/9       | 5/8     |

| Proportion of participants with normal GH response | 103/137 | 27/43 | 16/16 | 7/14 | – | 30/64 |
| Proportion of participants with normal GH response | 21/83 | 31/40 | 13/25 | 1/1 | 5/9 | 5/8 |

NFPA, non-functioning pituitary adenoma; second dose of insulin administered in those who failed to achieve adequate hypoglycaemia following first dose of insulin; NPG, nadir plasma glucose; glucose (mmol/l) to convert to mg/dl, multiply by 18.016; cortisol (nmol/l); GH (mg/l). Normal GH response defined as peak GH response of ≥6.67 μg/l.

### Table 2: Insulin dose.

<table>
<thead>
<tr>
<th>Insulin dose</th>
<th>All</th>
<th>Acromegaly</th>
<th>Cushing's</th>
<th>Prolactinomas</th>
<th>NFPA</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>187</td>
<td>26</td>
<td>7</td>
<td>27</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Percentage (%)</td>
<td>85</td>
<td>12</td>
<td>3</td>
<td>70</td>
<td>62.5</td>
<td>37.5</td>
</tr>
</tbody>
</table>

St, standard dose (0.15 U/kg); R, resistant dose (0.3 U/kg); and Se, sensitive dose (0.1 U/kg).

Most of the tests involved the use of the 'standard' dose of insulin (0.15 U/kg; Table 2). About one-third of the tests done in patients with previous Cushing's disease involved using the sensitive dose (0.1 U/kg); all of these patients were post-operative and considered to be steroid deficient.

Seventy percent of the patients with acromegaly had the standard dose of insulin. All of these were either after pituitary surgery or radiotherapy, so considered to have normal insulin sensitivity.

Figure 1 shows the bedside and laboratory venous plasma glucose measured at 30-min intervals over the course of the tests.

Thirty percent of the cohort had glucose of <2.0 mmol/l (36 mg/dl) for 60 min or more. The patients in this group were characterised by a lower body weight (86 vs 93 kg, \(P<0.05\)), a lower total insulin dose (14.4 vs 23.5 U, \(P<0.001\)), lower baseline GH (0.4 vs 0.6 mmol/l, \(P<0.001\)), lower GH response (peak GH 3.8 vs 6.4 mmol/l, \(P<0.0001\)) and lower fasting glucose (4.8 vs 5.2 mmol/l, \(P<0.001\)) compared with the rest of the cohort. Peak cortisol was, however, about the same in both groups (570 vs 460 nmol/l).

The patients with a nadir glucose of <1 mmol/l (18 mg/dl) had a lower body weight, lower total insulin dose and lower fasting glucose (79 vs 89.5 kg, \(P<0.05\); 13 vs 17 U, \(P<0.01\); and 4.6 vs 5.1 mmol/l, \(P<0.006\) respectively) compared with the rest of the cohort.

The NPG (Figs 2 and 3) ranged from 0.1 to 4.6 mmol/l (1.8–82.8 mg/dl) and correlated positively with FPG \(r=0.56; P=0.0005\), insulin dose \(r=0.27; P=0.0005\) and weight \(r=0.21; P<0.004\).

Determinants of nadir glucose
Using multiple regression analysis, the independent factors determining nadir blood glucose were fasting venous glucose \(b=0.56, P=0.0005, 20\%\) contribution) and weight \(b=0.14, P<0.05, 2\%\) contribution).

www.eje-online.org
Insulin dose

Twenty-four patients received an insulin dose exceeding 0.15 U/kg body weight. These patients were characterised by higher weight (mean 93 vs 86 kg, \( P < 0.001 \)) and FPG (mean 5.7 vs 4.9 mmol/l, \( P < 0.01 \)) compared with the rest of the population.

Outcome

Sixty-five percent of all ITTs met criteria for normal cortisol response to hypoglycaemic stress (Table 1), the highest proportion in those with prolactinomas.

Reproducibility

Fifty patients had ITTs done more than once. After excluding those whose tests were done after either surgery or radiotherapy (28 patients), the within-subject variability of NPG was 15.2%, peak cortisol was 11.7% and peak GH was 6.4%.

Adverse effects

Five patients had adverse events during the test (2%). One patient (aged 67 years) developed chest pain during the test and a slight troponin rise and subsequently had coronary artery bypass surgery following a finding of three-vessel coronary artery disease.

Four patients had blackouts: two occurred before insulin was administered and may represent vasovagal events (aged 43 and 75 years), and two occurred during confirmed hypoglycaemia (aged 56 and 68 years). The insulin dose and NPG and in these patients were comparable with those of the total population (patient with chest pain: 0.15 vs 0.17 U/kg and 1.9 vs 1.7 mmol/l (34.2 vs 30.6 mg/dl); patients with blackouts: 0.16 vs 0.17 U/kg and 1.68 vs 1.7 mmol (30.2 vs 30.6 mg/dl) respectively). The age range of the patients who experienced adverse events was 43–75 years, while the rest of the cohort was 17–88 years.

Fifty-eight patients had nadir glucose of \( \leq 1 \) mmol/l (18 mg/dl). These patients had to stay on average 30 min longer compared with the rest of the study population to ensure normal glucose at discharge. However, only two of these patients had an adverse event (blackouts).

Discussion

In this study we have assessed the duration and depth of hypoglycaemia achieved by patients undergoing the ITT. A significant proportion of patients achieved nadir glucose far below the target glucose for a prolonged period.

The ITT has been the gold standard for the assessment of the pituitary–adrenal axis and GH reserve for about 40 years. It was developed based on the response of normal and corticosteroid-treated subjects to the stress of surgery (18). The major advantage of this test is that hypoglycaemia as a stressor will activate the whole of the HPA and GH axes and as such will give a direct assessment of the integrity of a critical endocrine function (3, 4, 21, 23).

Although surrogate tests are available, they are often less sensitive and specific (3, 14, 15, 16). The glucagon stimulation test is thought to be inadequate as a stimulus for GH response and has relatively poor specificity (24). The administration of GH-releasing hormone (GHRH) with L-arginine (GHRH/arginine test) is an increasingly popular alternative (25, 26, 28, 29) as the GH response is reproducible with less variability (30). The short synacthen test is a widely used
alternative in the assessment of the integrity of the HPA axis in patients with pituitary disease (31, 32, 33, 34, 35); however, although it has been shown to be as good as the ITT in excluding ACTH deficiency, it does not assess the GH axis and has a lower sensitivity particularly soon after pituitary surgery (31).

The major disadvantage of the test is that hypoglycaemia is unpleasant and potentially dangerous (14, 16, 21). The literature would suggest that the risk of the ITT predominantly relates to the depth of hypoglycaemia rather than duration (36, 37, 38), although some of the risk may relate to the use of i.v. dextrose to rapidly correct blood glucose (14, 21). The magnitude of the cortisol and GH rise also relates to the depth rather than duration of hypoglycaemia (4). Thus an ideal test would be one where the target glucose is achieved for a short period but then goes no lower.

Although previous groups have audited their experience (19, 20), with one small study describing the depth of hypoglycaemia in a cohort of 16 GH-deficient patients during (39) the ITT, our data is novel as it shows the depth and duration of hypoglycaemia in a large number of patients undergoing this test.

The target glucose values used by other units varies from 2.2 to 2.8 mmol/l (3, 4, 16, 17, 19, 23, 40, 41), with most using 2.2 mmol/l, and we therefore currently use a target glucose of 2.2 mmol/l which was achieved by 89% of patients. However, over a third of this group experienced blood glucose of <2 mmol/l for more than an hour, with the lowest venous plasma glucose measured being 0.1 mmol/l, which must pose an unacceptable risk. In addition, the type of pituitary disorder does not predict predisposition to severe or prolonged hypoglycaemia.

Severe insulin-induced hypoglycaemia can have deleterious effects on brain activity (22, 42); however, as the counter-regulatory response occurs as the glucose falls below 3.5 mmol/l, the magnitude of this response depends on the depth of hypoglycaemia achieved (36, 43). The target nadir glucose during the ITT is the level at which normal subjects will show cortisol and GH responses (18) and it is not dependent on the rate of fall of glucose or the duration of hypoglycaemia. There is therefore no clinical need for the glucose to fall below this level.

The dose of insulin chosen is based on clinical assessment of the individual taking into account the presence of endocrine conditions that are associated with insulin resistance or increased insulin sensitivity and clinical markers of insulin resistance. Our data suggests that these criteria are poor at predicting the glucose fall in response to an insulin bolus. FPG appears to be the best predictor of subsequent sensitivity to insulin. As in previous studies, the reproducibility of the test is poor (21); our data showed an inter-subject variability in nadir glucose of 15.2%.

Our sample size is large and, although retrospective, includes all consecutive subjects that underwent the ITT. A weakness of this design is that we cannot provide data about subjects that were considered unfit for the ITT or declined the procedure. These patients were offered a synacthen test and basal pituitary profile as an alternative. A significant number of subjects may have undergone the ITT but needed repeated testing and declined following the experience. This information although interesting was unfortunately not available. We recorded adverse events in 2% of patients, with one serious adverse event (0.5%); Biller et al. (28), when comparing six different tests of the HPA axis, found no serious adverse events in the 68 patients who underwent the ITT but found that it was the least preferred test. We do not have data on patient experience but the majority of our patients who required repeat testing for monitoring purposes, e.g. following cranial/pituitary irradiation, were willing to have it done.

We have no measures of hypoglycaemia symptom scores during the tests. Venous blood glucose was also measured at 30-min intervals as a minimum, with the endocrine nurse present at the bedside throughout the test. This allowed for immediate venous glucose testing with a glucometer if signs or symptoms of hypoglycaemia were detected at any point during the test. The test might be improved by measuring venous glucose on a more frequent basis (e.g. every 10–15 min) to avoid missing the exact time that nadir glucose is achieved and reduce the risk of the severe hypoglycaemia experienced by a significant proportion of our cohort.

A further weakness of the study is that we do not routinely perform this investigation in children hence our findings cannot be extrapolated to the paediatric age group.

The ITT should be used with caution in the elderly (11, 12) because of the risk of underlying ischaemic heart disease, particularly in those with hypopituitarism. Similar to other centres (19, 20), our data shows...
that despite the inclusion of older patients and the depth of hypoglycaemia, adverse events are infrequent; this would suggest that our patient selection is adequate. We would however suggest caution in the older population, with careful assessment of their cardiovascular risk assessment before the test is performed.

To summarise, in order to produce adequate hypoglycaemia during the ITT, patients were routinely exposed to severe hypoglycaemia. The reasons for this are threefold: i) the depth and duration of hypoglycaemia that will be achieved following an i.v. insulin bolus is difficult to predict and is not reproducible; ii) the use of bedside handheld glucose meters systematically underestimates the depth of hypoglycaemia in this dynamic situation; and iii) laboratory glucose values are not available during the test and therefore hypoglycaemia is often prolonged to ensure an adequate test.

Potential areas for improvement to this test would include bedside testing of venous plasma glucose using more accurate means that would closely approximate that obtained from a laboratory, a means whereby the fall in blood glucose is controlled perhaps by the use of an insulin infusion combined with dextrose infusion at variable rates. One unit has successfully used a form of hypoglycaemic clamp to produce controlled hypoglycaemia albeit time and labour intensive (44). Our unit is currently undertaking a study using this method of achieving controlled hypoglycaemia in a small group of patients and following analysis of the results of this study might use this in all future ITTs.

In conclusion, the current paradigm for the investigation of HPA insufficiency includes the choice between uncontrolled hypoglycaemia following an i.v. insulin bolus or surrogate tests such as the synacthen test. The ITT is an old test. We perhaps need to update the test using modern technology to produce more controlled hypoglycaemia.

Declaration of interest
The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding
This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.

Acknowledgements
We thank Jinny Jeffery, Senior Clinical biochemist, for her assistance.

References


