The octreotide test dose is not a reliable predictor of the subsequent response to somatostatin analogue therapy in patients with acromegaly

Ana Pokrajac, Andrew G Claridge, S K Abdul Shakoor and Peter J Trainer
Christie Hospital NHS Trust, Department of Endocrinology, Wilmslow Road, Withington, Manchester M20 4BX, UK

Abstract
In many centres, a test dose (TD) of octreotide is administered before commencing somatostatin analogue therapy (SAT), although the merits of this procedure are uncertain. We have analysed the value of the GH response to a TD in predicting the efficacy of subsequent SAT in 47 patients with acromegaly (25 male, median age 51 years, range 20–82). The primary goal of SAT was a mean GH of < 5 mU/l. Median baseline GH was 19.3 mU/l (2.2–233 mU/l) and with the TD fell by 78% (35–98%) to a nadir of 4.2 mU/l (<0.3–85 mU/l). Optimal predictive power was observed when GH fell to < 5 mU/l after the TD. With this criterion, the TD had a positive predictive value (PPV) of achieving the primary goal on SAT of 82% and a negative predictive value (NPV) of 50%. However, baseline GH was also highly predictive of the likelihood of successful SAT (GH < 5 mU/l). The GH response to the TD had PPV of 83% and NPV of 61% of normalising IGF-I on SAT. In summary, baseline GH and nadir after a TD are highly predictive of a good response to SAT; however, a poor response to a TD does not exclude an optimal response to SAT. Furthermore, failure to achieve biochemical control does not equate to no benefit, as biochemical improvement was seen in every patient; therefore, no patient should be deprived of octreotide therapy because of the result of a TD. In conclusion, our data indicate that the octreotide TD has no place in selecting patients for SAT.

Introduction
The somatostatin analogues (SA) are the mainstay of medical therapy for acromegaly, with most patients receiving SA therapy at some point in their treatment. Octreotide is a synthetic, octapeptide SA with high affinity for the type 2 and 5 somatostatin receptors responsible for the regulation of growth hormone (GH) secretion (1). There is a 20-year experience of the short-acting, subcutaneous preparations used in the treatment of acromegaly, administered at a dose of 100–200 μg three times daily (2). In the last decade, the short-acting formulation has been superseded by a monthly, depot preparation that offers greater convenience for patients (3). Octreotide can control GH and insulin-like growth factor (IGF-I) secretion, on average, in approximately two-thirds of patients with acromegaly and, when used as primary therapy, it can induce significant tumour shrinkage in approximately 30% of patients (4, 5). Lanreotide, an alternative SA of comparable potency, is available as twice-weekly or monthly preparations (4, 6, 7).

SA are expensive and may have side effects. Thus, there has been great interest in identifying the factors which predict a good response to SA therapy (SAT), thereby focusing treatment on those most likely to respond. Studies have suggested that the response to SA is determined by tumour somatostatin receptor expression, but this is not a clinically practicable means of selecting patients for therapy (8). Pretreatment GH levels have been shown to predict the probability of achieving normal GH levels with long-term SAT (5). It has also been argued that the GH response to an acute dose of octreotide is of value in judging the likely response to long-term therapy. The octreotide test dose has become routine practice in many endocrine centres, but there is a dearth of data defining the criteria by which to judge the response and its predictive accuracy. The aim of this study was to analyse our experience of the octreotide test dose (TD) with the
goal of defining optimal criteria for predicting the response to long-term octreotide therapy.

Subjects and methods

Patients
Forty-seven patients (22 female; median age 51 years, range 20–82) with acromegaly had an octreotide test between February 1988 and September 2004. Thirty-seven patients had macroadenomas; 33 patients had undergone hypophysectomy and 37 radiotherapy. There was at least a 1-year interval between radiotherapy and initiation of octreotide. Twenty patients were on a fixed dose of a dopamine agonist at the time of the octreotide test and during subsequent SAT.

Protocols

Octreotide TD Fasted patients received subcutaneously 50 or 100 μg (n = 5) octreotide (Sandostatin, Novartis Pharma Services, Basel, Switzerland). Blood samples were drawn through an indwelling catheter at −15 and 0 min and hourly for up to 9 h. No difference was found in response between the two dose groups (42 vs 5 patients).

Assessment of disease activity The primary biochemical measure of disease activity was by calculation of the mean of five GH samples drawn during a single day (GH day curve (GHDC)) (9). In recent years, IGF-I has been a secondary parameter in the assessment of the activity of acromegaly.

SAT After the octreotide TD, patients proceeded to SAT with either short-acting octreotide (n = 6), long-acting octreotide (Sandostatin LAR, Novartis; n = 35) or lanreotide (Somatuline Autogel, Ipsen Ltd, Slough, Berks, UK; n = 1). The dose of SAT was titrated with the primary goal of achieving a mean GH of < 5 mU/l during a GHDC. A secondary goal was to lower IGF-I into the reference range.

Assays

GH Until 2000, GH was measured by immunoradiometric assay (Netria, London, UK) (sensitivity 0.5 mU/l; inter- and intra-assay coefficients of variation (CV) of < 5%); now one uses a two-site chemiluminescence immunoassay (Nichols Advantage; Nichols Institute Diagnostics, San Clemente, CA, USA) (sensitivity 0.3 mU/l; intra-assay CV 4.2–8.0%, interassay CV 4.1–12.1%).

IGF-I Until 2000, IGF-I was measured by in-house RIA with a single reference range for adults (8–38 nmol/l, precision and reproducibility of < 10%). This was replaced with a two-site chemiluminescence immunoassay (Nichols Advantage), with the age-related reference range provided by the manufacturer. Since 2003, age- and gender-related reference ranges have been used (10). The precision of this assay is 4.4–5.2% with reproducibility of 5.7–7.4% and limit of sensitivity of 6 μg/l.

Statistical analysis
The data had a skewed distribution; therefore, non-parametric tests were used. Correlation analysis was performed by Spearman’s rank correlation test. Sensitivity, specificity, and positive and negative predictive value (PPV and NPV) were calculated as follows: sensitivity = true positives/(true positives + false negatives); specificity = true negatives/(true negatives + false positives); PPV = true positives/(true positives + false positives); NPV = true negatives/(true negatives + false negatives). Numerous criteria were examined to establish the optimal PPV and NPV of the TD, and it was concluded that the criteria providing the best prediction of the subsequent response to SAT were a mean GH of < 5 mU/l and a percentage fall of > 50% and > 75% in GH values during the TD.

Results
No data on SAT are available in five patients (three failed to attend for follow-up, one entered a clinical study and one discontinued therapy due to severe gastrointestinal side effects). In the remaining 42, the median duration of follow-up was 2 years (range 3 months to 13 years), and the median doses were 30 mg per month (10–60 mg), 300 μg per day (150–600 μg) and 40 mg per month for long-acting octreotide (n = 35), octreotide (n = 6) and lanreotide (n = 1) respectively.

Response to SAT
Individual values for pretreatment GH levels, nadir GH after TD and during SAT are illustrated in Fig. 1. The mean pretreatment GH during a GHDC was 19.3 mU/l (range 2.2–233 mU/l). The lowest mean GH achieved on SAT was 3.6 mU/l (range < 0.3–26.6 mU/l), while median percentage fall during SAT was 83% (range −35% to +98%). Sixty-four per cent of patients achieved a mean GH during GHDC of < 5 mU/l.

Response to octreotide TD
There was a positive correlation between pretreatment and nadir GH values during the TD (r = 0.741, P < 0.001). The median GH value fell from a baseline of 19.3 mU/l (range 2.2–233 mU/l) to a nadir of 4.3 mU/l (range < 0.3–85 mU/l), the nadir occurring
at a median interval of 180 min (range 60–540 min). The median fall in GH was 78% (range 35–98%). Fifty-seven per cent of patients achieved a nadir GH of < 5 mU/l (Fig. 1). As described above, octreotide TD was analysed to optimise prediction of the subsequent response to SAT. One absolute criterion and two relative evaluating criteria were established as follows:

- mean GH nadir of < 5 mU/l during a GHDC
- over 50% decline in GH between the TD baseline and nadir
- over 75% decline in GH between the TD baseline and nadir.

**Relationship between the TD and the response to SAT**

There was a positive correlation between the GH nadir after TD and the best mean GH on SAT ($r = 0.54, P < 0.001$) (Fig. 2). The highest PPV (82%) of achieving the desired GH on SAT of < 5 mU/l was observed for TD nadir of < 5 mU/l. Percentage falls of > 50% and > 75% in GH values after TD had PPVs of the subsequent response to SAT of 69% and 72% respectively. Nevertheless, the NPVs of all criteria were unsatisfactory, standing at 50%, 38% and 40% respectively (Tables 1 and 2 and Figs 2–3), indicating that patients who did not have the desired response to a TD might still respond to SAT.

**Relationship between pretreatment GH values and the response to SAT**

There was a positive correlation between pretreatment GH values and those during SAT ($r = 0.38, P = 0.001$) (Fig. 4). The pretreatment GH values were highly predictive of the response to SAT. Eighty-one per cent of patients with a mean pretreatment GH during a GHDC of < 20 mU/l obtained the desired mean GH.
of < 5 mU/l during SAT, the results for patients with mean pretreatment GH of 20–60 mU/l and > 60 mU/l being 62% and 25% respectively (Fig. 4).

### Relationship between the TD and IGF-I response to SAT

At study entry, 22/32 patients in whom a baseline IGF-I was available had an IGF-I value above the upper limit of the reference range (ULRR). The primary reason to treat in the remaining patients was elevated GH levels. In the 22 patients with an initially elevated IGF-I, SAT resulted in a fall of IGF-I from a median of 107% above ULRR (range 29–290%) to 1% above ULRR (287% to 146%), with normalisation being seen in 11 patients (50%). Median pretreatment IGF-I value expressed relative to the ULRR was higher in the patients in whom IGF-I did not normalise on treatment (140% (range 61–290%) vs 29% (range 71% to +242%); P < 0.001). The best predictor of normalisation of IGF-I (PPV 83%, NPV 61%) was a GH nadir of < 5 mU/l after the TD. Although there was a positive correlation (r = 0.48, P = 0.002) between GH and IGF-I during long-term SAT, results were discrepant in 27% of patients, with elevated IGF-I and mean GH of < 5 mU/l in 12%, and normal IGF-I and elevated GH in 15%.

### Discussion

In many centres, the assessment of the GH response to a TD of octreotide has been routine clinical practice since the introduction of octreotide therapy almost 20 years ago and has even been used as an entry criterion for octreotide therapy studies (11–13). The rationale for performing the test and the interpretation of the results have been largely undefined. Although it has been argued that the response to a TD of octreotide predicts the response to subsequent SAT, there is a lack of accepted guidelines for judging the response (14–16). To date, there have been five detailed analyses of the merits of the TD reported which reached markedly differing conclusions (Table 3) (17–21). Colao et al. (17) reported the GH response to 100 μg s.c. octreotide in 68 patients. The response was judged by calculating the mean of GH samples collected at 2–6 h after octreotide, rather than the nadir GH value reported here. Pretreatment GH values correlated with the best mean GH on treatment, although only 41% achieved the desired GH levels of < 5 μg/l. The authors defined a positive response to the TD as a decrease in GH of > 50%, and by this definition, the TD had a PPV of 43% and NPV of 78% of achieving normalisation of GH during long-term octreotide treatment. Colao et al. (17) concluded that the TD was not a suitable means of identifying patients who would respond to long-term octreotide therapy.

Recently, Karavitaki et al. (18) reported the relationship between the TD and SAT in 30 patients. With 100 μg octreotide, 75% of patients achieved the desired GH levels of < 5 mU/l. With receiver operating characteristics (ROC) curves to optimise the diagnostic value of the TD, 100% sensitivity and 80% specificity could be achieved (using the criterion of a GH nadir of < 5.25 mU/l), which equates to a PPV of 94% and NPV of 100%. For lanreotide, the best results were obtained with the criterion of GH nadir of < 6.05 mU/l (sensitivity 92%, specificity 67%, PPV 92%, NPV 67%). However, it must be appreciated that GH assays lack the performance to allow such precise measurement, and thus diagnostic accuracy inevitably deteriorates when more practicable criteria are applied. Karavitaki et al. (18) concluded that the TD is a reliable tool for the selection of patients with active acromegaly that will achieve ‘safe’ GH levels on therapy with octreotide LAR.

### Table 1

<table>
<thead>
<tr>
<th>Baseline GH (mU/l)</th>
<th>n</th>
<th>Achieved GH nadir &lt; 5 mU/l following TD</th>
<th>Achieved GH nadir &lt; 5 mU/l on SAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 20</td>
<td>21</td>
<td>19/21 (90%)</td>
<td>17/21 (81%)</td>
</tr>
<tr>
<td>20–60</td>
<td>13</td>
<td>4/13 (31%)</td>
<td>8/13 (62%)</td>
</tr>
<tr>
<td>&gt; 60</td>
<td>8</td>
<td>1/8 (13%)</td>
<td>2/8 (25%)</td>
</tr>
<tr>
<td>All</td>
<td>42</td>
<td>24/42 (57%)</td>
<td>27/42 (64%)</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Mean GH on SAT &lt; 5 mU/l</th>
<th>GH nadir following TD &lt; 5 mU/l</th>
<th>GH fall &gt; 50% following TD</th>
<th>GH fall &gt; 75% following TD</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPV (%)</td>
<td>82</td>
<td>69</td>
<td>72</td>
</tr>
<tr>
<td>NPV (%)</td>
<td>50</td>
<td>38</td>
<td>40</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>67</td>
<td>81</td>
<td>67</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>69</td>
<td>23</td>
<td>46</td>
</tr>
</tbody>
</table>
Biermasz et al. (19) analysed the relationship between the response to the acute test (50 μg i.v. octreotide) and the outcome of ‘chronic octreotide LAR treatment’ in 18 patients with acromegaly. A nadir GH of < 5 mU/l during the TD predicted with 100% sensitivity and specificity GH levels of < 5 mU/l on treatment. However, GH and IGF-I results were discrepant in 33% of patients, and the GH response to the TD was a poor predictor of normalisation of IGF-I (PPV 73%, NPV 57%). Despite this, the authors supported the use of the acute test as a predictor of the response to octreotide treatment.

Gilbert et al. (20) reviewed the results of a TD (50 μg sc octreotide) in 33 patients. A TD GH nadir of < 5 mU/l had 80% sensitivity and 83% specificity in predicting a mean GH < 5 mU/l (defined goal of therapy) during SAT, from which the authors judged the TD useful. However, a fifth of patients who achieved GH < 5 mU/l on treatment had a GH nadir > 5 mU/l following TD and on that basis would not have qualified for SAT. Furthermore, the majority (10/14) of patients classified by their GH levels as in remission had IGF-I above upper limit of reference range.

De Herder et al. (21) applied 2 relative (> 50 or 75% fall) and 2 absolute criteria (nadir GH < 1.1 or 2 mcg/l), all arbitrary, in their analysis of the TD performance in 24 therapy-naïve patients. In contrast to the other groups, the primary goal of therapy was to normalise IGF-I, with a > 50% fall in GH during the TD having sensitivity and negative predictive value of 100%. No patient in whom GH failed to fall by > 50% during the TD did normalise IGF-I on the treatment; however, the other criteria were of little value and the authors concluded that the acute test was unnecessary prior to Sandostatin LAR therapy.

In contrast to Colao et al. (17) and De Herder et al. (21), our study and the other recent ones (18–20) had as the primary goal of therapy a mean GH

![Figure 3](https://www.eje-online.org)

**Figure 3** No significant correlation \( (r = -0.06, P = 0.72) \) is seen between the percentage fall in GH after the TD and mean GH on SAT. Despite falls in GH values following the TD of > 75%, patients in quadrant B failed to normalise GH during treatment.

![Figure 4](https://www.eje-online.org)

**Figure 4** The pretreatment GH value is predictive \( (r = 0.38, P = 0.001) \) of the lowest GH value during SAT. However, as can be seen in quadrant C, patients with barely elevated GH values can fail to achieve GH values during SAT of < 5 mU/l, the opposite being seen in quadrant D.
<table>
<thead>
<tr>
<th>Publication</th>
<th>Test dose</th>
<th>Patient nos</th>
<th>TD criteria</th>
<th>Treatment goal</th>
<th>Predictive value of TD</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colao et al. 1996 (17)</td>
<td>100 μg s.c.</td>
<td>68</td>
<td>Mean GH 50% fall</td>
<td>GH &lt; 5 μg</td>
<td>PPV 43% NPV 78% sens. 100% spec. 80% PPV 94% NPV 100%</td>
<td>Not valuable</td>
</tr>
<tr>
<td>Karavitaki et al. 2005 (18)</td>
<td>100 μg s.c.</td>
<td>30</td>
<td>Nadir GH &lt; 5.25 mU/l</td>
<td>GH &lt; 5 mU/l octreotide &lt; 6.05 mU/l lanreotide</td>
<td>PPV 92% NPV 67% sens. 92% spec. 67% PPV 92% NPV 100%</td>
<td>Reliable selection tool for octreotide</td>
</tr>
<tr>
<td>Biermasz et al. 2005 (19)</td>
<td>50 μg i.v.</td>
<td>18</td>
<td>Nadir GH &lt; 5 mU/l</td>
<td>GH &lt; 5 mU/l</td>
<td>Norm. IGF-1</td>
<td>PPV 73% NPV 100% sens. 100% spec. 100%</td>
</tr>
<tr>
<td>Gilbert et al. 2005 (20)</td>
<td>50 μg s.c.</td>
<td>33</td>
<td>Nadir GH &lt; 5 mU/l</td>
<td>GH &lt; 5 mU/l</td>
<td>Norm. IGF-1</td>
<td>PPV 73% NPV 77% sens. 80% spec. 83%</td>
</tr>
<tr>
<td>De Herder et al. 2005 (21)</td>
<td>50 μg s.c.</td>
<td>28</td>
<td>Mean GH &gt; 50% fall</td>
<td>Norm. IGF-1</td>
<td>&gt;75% fall</td>
<td>Norm. IGF-1</td>
</tr>
<tr>
<td>Pokrajac et al. 2005 (Present study)</td>
<td>50/100 μg s.c.</td>
<td>47</td>
<td>GH nadir &lt; 5 mU/l</td>
<td>GH &lt; 5 mU/l</td>
<td>&lt;2 μg</td>
<td>Norm. IGF-1</td>
</tr>
</tbody>
</table>
of <5 mU/l (22, 23), and nadir rather than mean values following the octreotide TD were analysed. The difference in therapeutic goal between the studies is irrelevant to the analysis of the TD as in all studies the dose of SAT was titrated to achieve the predetermined therapeutic goal.

In our study, the best predictive value of the TD (PPV 82%, NPV 50%) was achieved using the criterion of a GH nadir <5 mU/l during TD, with patients meeting this criterion consistently achieving the desired goal during long-term treatment of mean GH during GHDC of <5 mU/l. While a positive response to a TD predicts a good response to subsequent therapy, the limitation of the test is that failure to respond to a TD does not preclude achieving GH values of <5 mU/l on long-term treatment. Twenty-seven percent of patients who did not respond to a TD went on to achieve GH levels of <5 mU/l during SAT; therefore, no patient should be denied SAT based on a failure to respond to a TD.

The case for performing octreotide tests is further undermined by our confirmation of the previously reported correlation between pre-treatment GH values and the best mean GH values achieved on subsequent SAT (5). In our study pretreatment GH values were highly predictive of the probability of achieving a GH of <5 mU/l with treatment. Eighty-one percent of patients with a pretreatment GH <20 mU/l achieved the desired GH levels, which contrasts with only 25% of patients with pretreatment GH levels >60 mU/l. Furthermore, failure to achieve the desired GH of <5 mU/l does not mean that patients have not benefited from octreotide therapy; in our cohort of such patients mean GH levels fell by a median of 78% (range 23% to 100%) and IGF-I reduced from 140% ULRR (range 61–290%) to 41% ULRR (range 5–146%).

Our study and those of De Herder et al. (21) and Biermasz et al. (19) are the first to report the value of the octreotide TD in predicting the IGF-I response to SAT. As with GH, we found that the higher the baseline IGF-I, the less the likelihood of normalisation with treatment. The absence of a response to a TD of octreotide does not preclude normalisation of IGF-I during long-term SAT. IGF-I is increasingly being used as the primary goal of therapy in patients with acromegaly in which case the use of the GH response to a TD of octreotide is of even more limited value. In our study, IGF-I remained elevated in 12% of patients obtaining a mean GH of <5 mU/l.

In conclusion, our data indicate that the octreotide TD is able to predict a positive response to SAT therapy. However, the limitation of the test is that failure to respond to a TD does not preclude a good response to long-term therapy. Therefore, it would be inappropriate to deprive a patient of octreotide treatment because of a poor response to a test dose. Furthermore, even if a patient fails to achieve an optimal response to treatment, defined
here as a mean GH < 5 mU/l, there is still symptomatic and long-term benefit from lowering GH and IGF-I levels, in this study by 78% and 44% respectively. The octreotide TD has no place in the assessment of patients prior to octreotide therapy and should be abandoned.

References

3 Stewart PM, Kane RF, Stewart SE, Lancranjan I & Sheppard MC. Depot long-acting somatostatin analogue (Sandostatin-LAR) is an effective treatment for acromegaly. Journal of Clinical Endocrinology and Metabolism 1995 80 3267–3272.