Growth hormone deficiency (GHD) in adults is a recognised clinical syndrome and its diagnosis is established through GH stimulation testing. The decision to perform GH stimulation testing should be based on clinical findings, medical history and using the appropriate GH stimulation test to obtain biochemical evidence. The insulin tolerance test (ITT) remains the diagnostic test of choice, but this test is labour intensive, contraindicated in the elderly and in adults with seizure disorders and ischaemic heart disease, can be unpleasant for the patient, and is potentially hazardous. The glucagon stimulation test (GST) in recent years has been increasingly used as the alternative test to the ITT in the US and Europe because of its availability, reproducibility, safety, lack of influence by gender and hypothalamic cause of GHD, and relatively few contraindications. In the article, we discuss our recommendations in performing this test, the potential drawbacks in conducting and caveats in interpreting this test, and its future perspectives.

Keywords
Growth hormone, glucagon, insulin tolerance test, growth hormone releasing hormone plus arginine, diagnosis, adult growth hormone deficiency

Disclosure: Kevin CJ Yuen has received research support from Pfizer, Genentech and Novo Nordisk.

Received: 11 September 2010 Accepted: 5 January 2011 Citation: European Endocrinology, 2011;7(2):104–7 DOI:10.17925/EE.2011.07.02.104

Correspondence: Kevin CJ Yuen, Division of Endocrinology, Diabetes and Clinical Nutrition, Oregon Health and Science University, 3181 SW Sam Jackson Park Road, Mailcode: L607, Portland, OR 97239-3098, US. E: yuenk@ohsu.edu

Glucagon Testing in Adult Growth Hormone Deficiency – Current Status and Future Perspectives

Kevin C J Yuen

Abstract
Growth hormone deficiency (GHD) in adults is a recognised clinical syndrome and its diagnosis is established through GH stimulation testing. The maximum or peak GH secretion following GH stimulation testing is used as a surrogate of the capacity of the pituitary to release GH. The insulin tolerance test (ITT) is generally considered the gold standard test for evaluation of GH deficiency and has been endorsed by different consensus guidelines, but this test is labour intensive, may be unpleasant for some patients, has potential risks, and is contraindicated in elderly patients and in patients with seizure disorders and ischaemic heart disease. Thus, there remains a real unmet medical need for an alternative test to the ITT that is safe yet reliable. For this reason, other dynamic tests have been proposed such as arginine (ARG), combined GH releasing hormone plus ARG (GHRH-ARG), levodopa (L-DOPA) in spite of data indicating poor performance of some of these tests for evaluation of adult GHD. A potential alternative to the ITT is the glucagon stimulation test (GST) that has been used extensively in the UK, and is gradually gaining acceptance in the US.

Update on the Glucagon Stimulation Test in Diagnosing Adult Growth Hormone Deficiency
Following the publication of several validation studies and recommendations from current consensus guidelines, the GHRH-ARG test has in recent years emerged as the best and most reliable alternative GH stimulation test to the ITT in diagnosing adult GHD. However, when EMD Serono, Inc decided to discontinue the manufacture of recombinant GHRH (Geref®) in the US in July 2008, this inevitably left a significant gap for an alternative reliable test for evaluation of patients suspected to have GHD in place of the GHRH-ARG test. This is particularly important for endocrinologists in the US who are not comfortable or do not have the necessary logistical or staff support to conduct ITTs in their office or patients who have contraindications to hypoglycaemia in whom the ITT would be inappropriate. With the lack of reliable GH stimulation tests available in the US, we recommended the glucagon stimulation test (GST) as the alternative test to the ITT for diagnosing adult GHD based on its availability, reproducibility, safety, lack of influence by gender and hypothalamic cause of GHD, and relatively few contraindications.
Analysing the data of 13,167 GH-deficient patients enrolled in the KIMS pharmaco-epidemiological database (Pfizer International Metabolic Database) from its inception to the end of 2008, Brabant et al. addressed the question of whether there were regional differences in the use of different biochemical tests to diagnose adult GHD in six large European countries and the US. This analysis revealed striking regional variations in the approach to GH stimulation testing, with the ITT being the most popular test used in 44% of all countries but was less popular (13%) in the US, whereas the GST was more popular in the UK (30%) than in the US. However, the unavailability of the GHRH-ARG test in the US since 2008 has resulted in the GST being more frequently considered as the alternative test to the ITT.

The use of the GST for the assessment of GH reserve was first described in 1969 by Mitchell et al. Since then, the GST has been shown by various investigators to have a GH secretory potency that is similar to or only slightly less than the ITT, suggesting that it is more reliable than other classical agents such as ARG or clonidine for separating GHD patients from normal subjects. The true mechanism by which glucagon induces GH release remains unclear. Some of the hypothesised mechanisms include the glycaemic fluctuations during the test where blood glucose levels increase initially before decreasing later in the test, the generation of a peptidyl fragment associated with the GH- and ACTH-releasing activity and the inactivation of norepinephrine secretion in stimulating GH release via α-receptors. It has also been previously demonstrated that glucagon stimulates GH release more effectively when administered intramuscularly or subcutaneously as opposed to the intravenous route.

The three studies utilising the GST by Gomez et al., Conceicao et al., and Berg et al. evaluated GHD in patients with pituitary disorders. The first two studies were prospective studies that compared the diagnostic characteristics of GST to ITT and included a control group which was matched for age and sex in both studies and for body mass index (BMI) in one study. Using receiver operated curve (ROC) analysis, both studies proposed a peak GH cut-off value of 3 ng/mL as the best cutpoint with the highest combined sensitivity and specificity to differentiate between patients with GHD and healthy controls. In addition, Gomez et al. found no correlation among age, sex and BMI with peak GH levels. In contrast, the study by Conceicao et al. demonstrated that peak GH levels were not affected by age in either the control or patient group, and that there were no gender differences. On the other hand, the study by Berg et al. was a retrospective study that revealed an optimal peak GH cut-off value of 2.5 ng/mL with 95% sensitivity and 79% specificity using ROC analysis. This study also reported lower peak GH levels with GST compared to ITT (5.1 versus 6.7 ng/mL, p=0.01) but a significant positive correlation between peak GH levels during ITT and GST (r=0.88, p=0.0001). Additionally, no correlation between BMI and age to peak GH responses were observed, peak GH responses occurred mainly between 120 and 180 minutes consistent with previous studies, and that, overall, the GST was a well-tolerated test. Nevertheless, these and previous studies did not specifically evaluate patients with glucose intolerance and frank diabetes, and for this reason, the characteristic of the GST and its reliability in testing for GHD in this population remains unclear. This is especially important since performing ITT in patients with diabetes is usually challenging and may not be safe especially if large insulin doses are required to achieve hypoglycaemia in patients with underlying insulin resistance.

### Other Considerations in Performing and Interpreting the Data of the Glucagon Stimulation Test

The diagnosis of adult GHD has proved to be challenging because of the lack of a single biological endpoint such as growth failure, and therefore, the confirmation of adult GHD largely depends on biochemical provocative testing. Clearly, there is no ideal stimulation test and we recommend that the decision to embark on a stimulation test to diagnose adult GHD must factor in the appropriate clinical context of each individual patient together with the number of pituitary hormone deficiencies plus serum IGF-I level, the validity of the chosen test and its appropriate cut-off limits, the sensitivity of the GH assay, and the availability of local resources and expertise.

### Table 1: Recommended Protocol for Performing the Glucagon Stimulation Test in Assessing Growth Hormone Reserve in Adults

<table>
<thead>
<tr>
<th>Contraindications</th>
<th>Precautions</th>
<th>Sampling and Measurements</th>
<th>Normal Response</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malnourished patients or patients who have not eaten for &gt;48 hours</td>
<td>Patients may feel nauseous during and after the test (administration of intravenous anti-emetics can be considered)</td>
<td>Serum GH and capillary blood glucose levels at 0, 30, 60, 90, 120, 150, 180, 210 and 240 minutes</td>
<td>Blood glucose: usually rises to peak around 90 minutes and then gradually declines (not used to interpret the test)</td>
<td>In adults with GH deficiency, peak GH levels fails to rise above 3 ng/mL.</td>
</tr>
<tr>
<td>Weigh patient</td>
<td>Glucagon administered intramuscularly 1 mg (1.5 mg if patient weighs more than 90 kg)</td>
<td>Glucagon administered intravenously 0.8 mg/kg (1.2 mg/kg if patient weighs more than 100 kg)</td>
<td>GH: rises to above 3 ng/mL</td>
<td>GH = growth hormone.</td>
</tr>
<tr>
<td>Ensure patient is fasted from midnight</td>
<td>Late hypoglycaemia may occur (patients should be advised to eat small and frequent meals after completion of the test)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

GH = growth hormone.
contraindicated in patients with malnourishment or who have not eaten for more than 48 hours due to concern of prolonged hypoglycaemia and those with pheochromocytoma in whom a significant exacerbation of blood pressure may be observed.\textsuperscript{18}

The GST was initially described as a four-hour test in older studies,\textsuperscript{11,22} but more recent studies have suggested that the test could be shortened to a three-hour test, and that serum GH levels can be evaluated between three to five time points only (0, 90, 120, 150 and 180 minutes) as the majority of GH peaks occurred between 120 and 180 minutes (85 \%).\textsuperscript{26,29} In a study by Orme et al. comparing standard and simplified GST (0, 150 and 180 minutes), 75 \% of discordant GH results were due to a peak GH level occurring at 210 minutes.\textsuperscript{28} Accordingly, the authors proposed that the diagnostic utility of the simplified GST could be improved further by drawing an additional blood sample at 210 minutes when assessing GH deficiency.\textsuperscript{30} The audit by Leong et al. is the largest study that assessed patients with hypothalamic-pituitary disease whom had undergone the GST, and they reported that the test could be shortened by omitting the 240-minute blood sample.\textsuperscript{31,32} Among 414 patients who underwent GSTs, the majority of peak GH levels occurred between 120 and 180 minutes (85 \%) and five patients had their peak GH levels recorded at 240 minutes.\textsuperscript{33} Hence, it is still not clear whether the ideal timing of the GST is three versus four hours, and continuing the test for four hours may be advisable, at least until there are more definitive data available. This also allows the monitoring for late hypoglycaemia, although truly low blood glucose levels are not common. While the lowest blood glucose level with the GST in the literature was reported at 37 mg/dL,\textsuperscript{34} in our experience, we have rarely observed blood glucose levels falling below 40 mg/dL with this test. The occurrence of hypoglycaemia reported in the literature with blood glucose levels lower than 40 mg/dL during GSTs are also rare events.\textsuperscript{35,36}

The common side-effects in patients with hypothalamic-pituitary disease that underwent testing with the GST included nausea, vomiting and headaches, and have been reported to range from less than 10\% to 34 \%.\textsuperscript{26} In a study of 97 normal subjects, mild nausea in approximately 30 \% of the subjects, and transient vomiting and retching in about 10 \% of the subjects were the only side-effects that were noted,\textsuperscript{26} whereas in our experience of 143 GSTs performed at four large academic centers in the US, the main side-effects reported were nausea (41 \%), fatigue, headaches, weakness and hunger (12 \%).\textsuperscript{37}

Like other GH stimulation tests, there are also limitations associated with the GST. The three- or four-hour GST is still longer than many other GH stimulation tests, and requires an intramuscular injection which may not appeal to some patients. However, as there is a relationship between peak GH response to GHRH-ARG stimulation and ambient glucose levels,\textsuperscript{38} it is unclear whether hyperglycaemia may play a part in influencing the peak GH response to glucagon stimulation. Furthermore, no peak GH responses have been studied using the GST in normal controls over the age of 70 years and none of the previous studies included patients with frank diabetes. Therefore, it is not known whether testing using the GST in subjects with diabetes is valid. Hence, caution should be exercised when interpreting normal GST results in the patients with diabetes. If the suspicion of GHD remains high in these patients, it is reasonable to consider using a second GH stimulation test. Finally, while it is accepted that a peak GH response of 3 ng/mL or less is the best cutpoint to diagnose adult GHD using the GST,\textsuperscript{10,27} there remains a lack of consensus over a peak GH response between 3 ng/mL and 10 ng/mL, and further studies are required to address this.

Other provocative tests that have been proposed include ARG alone and GH secretagogues. ARG alone has been shown to be less reliable than the ITT or GHRH-ARG11 and the mean peak GH response to ARG alone is lower than in the ITT or GST, even in normal lean subjects.\textsuperscript{11} The diagnostic reliability of ARG alone has been previously questioned.\textsuperscript{11,22} Thus, we recommend that ARG alone should only be considered if the ITT and the GST is contraindicated or if glucagon is unavailable. If this test is used, appropriately low peak GH cutoffs should be employed (for 95 \% sensitivity: 1.4 µg/L, for 95 \% specificity: 0.21 µg/L, and to minimise misclassification in either direction: 0.4 µg/L).\textsuperscript{11} In contrast, the reliability of testing with GH secretagogues such as GH-releasing peptide-2 alone,\textsuperscript{44} GH-releasing peptide-6 alone and combined GH-releasing peptide-6 plus GHRH\textsuperscript{45} in comparison with the ITT has also been demonstrated. These agents utilise the same concept as the GHRH-ARG test in stimulating pituitary GH release by mimicking the activity of the natural GH secretagogue receptor ligand (i.e. ghrelin), and appear to demonstrate a good safety profile with relatively few contraindications.\textsuperscript{46} The limitation, however, of these GH secretagogues is that these agents are more likely to explore the pituitary somatotroph releasable pool and might potentially induce misleadingly normal peak GH responses in hypothalamic GHD.\textsuperscript{47} Furthermore, these agents are not readily available in many countries including the US.

**Future Perspectives**

Recent studies have indicated that further refinements to the GST may still be required to improve the sensitivity and specificity of this test. A study by Micmacher et al. demonstrated in a group of healthy men above 50 years old that GH secretion in response to the GST, but not with the ITT, correlated to physiological spontaneous GH secretion.\textsuperscript{48} These data indicate that GH response to the GST reflects the endogenous GH spontaneous secretion and poses the question as to whether the cutpoints of peak GH response to the GST should be age-dependent. More recently, we reported a one-year experience of GSTs conducted from 4 large academic centers in the US and explored the potential of the GST in testing the hypothalamic–pituitary–adrenal axis.\textsuperscript{83} In this study, we found that there was a negative correlation between fasting glucose (r=−0.24, p<0.01) and BMI (r=−0.37, p<0.01) with peak GH levels. When compared with the 250 µg cosyntropin stimulation test with a cutpoint of 18 µg/dL, peak cortisol levels with GSTs were lower (p<0.02), had higher failure rates (44.4 \% versus 33.3 \%) and the 120-minutes peak cortisol of 16.5 µg/dL achieved 83.3 \% sensitivity and 75 \% specificity using ROC analysis. Overall, the GST was well-tolerated and can be performed as an out-patient; however further studies are needed to determine whether GSTs may falsely diagnose GHD in patients with fasting hyperglycaemia, and/or high BMIs. Thus, to improve the diagnostic reliability of the GST especially in patients with glucose intolerance and in those with high BMIs, a priming agent may be required to combine with the GST with appropriate cutpoints to improve its sensitivity and specificity, similar to the GHRH in priming the ARG test. Until such data becomes available, we recommend that a second GH stimulation test should be considered for such patients.

In conclusion, in line with recent published consensus guidelines,\textsuperscript{7,8} the ITT should remain as the test of reference due to its greatest diagnostic accuracy, even in patients with suspected hypothalamic GHRH deficit. We recommend the GST as the alternative test to the ITT.
for diagnosing adult GHD because of its availability, reproducibility, safety, lack of influence by gender and hypothalamic cause of GHD, and relatively few contraindications. Despite some studies demonstrating the comparability of the GST to the ITT in assessing the hypothalamic–pituitary–adrenal axis, further larger, well-controlled studies are still needed to confirm the reliability of the GST in assessing this axis. If the GST can be shown to reliably distinguish adrenocortical insufficiency from insufficiency, then the ability of assessing both the GH and cortisol reserve simultaneously, just as in the ITT, would make this test even more appealing. While previous studies have shown that the GST could be shortened from four to three hours and yet maintain its diagnostic utility, we would still recommend that the GST be conducted over four hours with measurements every 30 minutes for serum GH and capillary blood glucose levels primarily to ensure that delayed peak GH responses and late hypoglycaemia are not missed.

2. de Boer H, Blaj IG, Van der Veen EA, Clinical aspects of growth hormone deficiency in adults, J Clin Endocrinol Metab, 1995;80(1):63–86.