Available online at www.sciencedirect.com
SCIENCE DIRECT®
Growth Hormone & IGF Research 15 (2005) S21–S25
www.elsevier.com/locate/ghir

# A brief review of the use and utility of growth hormone stimulation testing in the NCGS: Do we need to do provocative GH testing?

Darrell M. Wilson [a,*], James Frane [b]

[a] Pediatric Endocrinology and Diabetes, Stanford University, S-032 Medical Center, Stanford, CA 94305-5208, USA
[b] Biostatistics Consultant to Genentech, Inc., Santa Monica, CA, USA

**Abstract**

True growth hormone deficiency (GHD) in childhood, while rare, has major clinical consequences. GHD is often associated with other pituitary hormone deficiencies, so these children may require multiple hormonal replacement and close clinical follow-up to optimize their outcome. Growth hormone stimulation testing (GHST), as currently conducted, is not a reliable diagnostic tool. Both changes in growth hormone assay methodologies and increases in the diagnostic threshold contribute to the incorrect labeling of a substantial proportion of normal children as having idiopathic GHD. Fortunately, newer imaging technologies and laboratory tests form a more rational basis to diagnose true GHD.

The use of GHST among GH-naive subjects (<20 years of age) enrolled in the National Cooperative Growth Study has declined over the past two decades, from a high of 89% between 1987 and 1989 to only 52% in 2002. Given that GH stimulation testing does not meaningfully aid in distinguishing those few children with true growth hormone deficiency from the much more common short normal child and that alternatives are now available, it is time to discontinue the routine use of GHST in children.


*Keywords:* Growth hormone stimulation testing; National Cooperative Growth Study; Insulin tolerance test; Height standard deviation scores

## 1. Introduction

Historically, growth hormone stimulation testing (GHST) has played a prominent role in diagnosing growth hormone deficiency (GHD). There are growing concerns, however, that GHST, particularly as currently conducted in the USA, is neither precise nor accurate in quantifying a patient's GH secretory status. Recent long-term follow-up data demonstrating increased mortality [1] in patients with hypothalamic–pituitary disorders highlight the importance of accurately diagnosing GHD.

Many investigators have recognized the multiple problems associated with GHST. Gandrud and Wilson [2] recently reviewed many of the aspects of GHST that detract from its utility in aiding the diagnosis of GHD, including a lack of reproducibility. Even when the testing is performed in the same clinical setting, using identical protocols and consistent GH assays, many researchers have reported marked variability in results when GHST is repeated in the same patient [3–5].

When GHST first became a routine component of endocrine clinical practice, GH assays were more consistent, and the cutoff for diagnosing GHD was much lower. Paradoxically, even as some newer GH assays report lower GH concentrations in pooled clinical samples, the diagnostic cutoff to exclude GHD has risen over the years to 10 ng/mL. Marin et al. [6] demonstrated

* Corresponding author. Tel.: +1 650 723 5791; fax: +1 650 725 8375.
*E-mail address:* dwilson@stanford.edu (D.M. Wilson).

1096-6374/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ghir.2005.06.005

that 75% of normal prepubertal children failed to achieve a peak GH of >10 ng/mL; others have published similar findings [7,8]. The current diagnostic criterion of a GH peak greater than 10 ng/mL to exclude GHD is arbitrary and has led to the misdiagnosis of many children and adolescents.

There is no consistent approach to the selection of provocative agents, and many clinicians do not prime prepubertal children with sex steroids. In a recent review, Sizonenko et al. states that "at least 34 provocative tests have been developed and 189 different combination protocols were found. Most investigators have used these tests without normative values for age, pubertal stage or gender" [9]. Clearly, if we had a useful GHST methodology, there would not be 189 different combinations. GHST is expensive, labor-intensive, and, on occasion, risky. Insulin-induced hypoglycemia has been associated with seizures, neurological damage, and, in two patients, death [10].

Many alternative diagnostic strategies to GHST have been developed. As recently reviewed by Wilson and Badaru [11], a combined approach involving auxologic assessments, measurement of GH-responsive factors such as insulin-like growth factor I (IGF-I) and insulin-like growth factor binding protein 3 (IGFBP-3), cranial imaging, and in selected patients, genetic studies, has evolved into a more rational approach to diagnosing GHD.

Against this backdrop, we have reviewed Genentech's National Cooperative Growth Study database to characterize how clinicians' use of GHST has changed over the past two decades.

## 2. Methods

The National Cooperative Growth Study (NCGS) was started in 1985, and its database now contains clinical information from more than 44,000 subjects treated with GH (manufactured by Genentech, Inc.) for more than 165,000 patient-years (the data from 2003 were not complete when the database was queried in mid 2004) [12]. The NCGS historically categorized patients as having either idiopathic short stature (ISS) or isolated GHD, based on a stimulated peak GH concentration of > or <10 ng/mL, respectively. Because the premise of this report is that the GHST is not a useful diagnostic test, subjects classified as ISS and IGHD have been combined for some analyses.

This database was queried for data on subjects naive to GH therapy who were under the age of 20 when GH therapy was started. Specific a priori questions included:

- How many new patients were enrolled each year and what percentage underwent GHST?
- Which stimulating agent(s) was used in GHST, and did this change over time?
- Was the change in height standard deviation score achieved during the first year of GH therapy associated with the peak stimulated GH concentration determined prior to GH therapy?

## 3. Results

As of 2004, the NCGS database contained clinical information for 44,515 GH-naive subjects <20 years old at the time of GH initiation. Of these, 33,073 (74.3%) had undergone GHST. The proportion of patients subjected to GHST was ~89% from 1985 through 1991 (Fig. 1 and Table 1), declining to 52% in 2002. Among subjects in the combined IGHD/ISS group, the frequency of GHST use was higher (Table 1), ranging from 90% to 95% between 1985 and 1992, but it gradually declined to 68% in 2003. Among those subjects with organic GHD, GHST was performed less

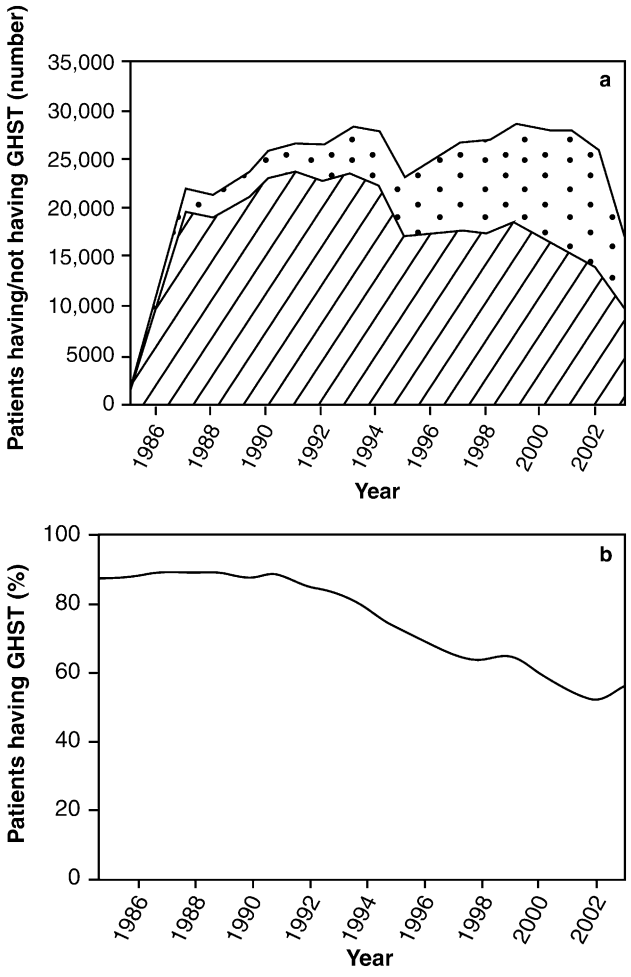


Fig. 1. (a) Total number of growth hormone–naive subjects (dots) under the age of 20 years by year of NCGS enrollment and the number undergoing GHST (diagonal lines). (b) Percentage of growth hormone–naive subjects undergoing GHST.

Table 1
GH-naive patients undergoing GHST

| YEAR | 1985 | 1986 | 1987 | 1988 | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Total naive* | | | | | | | | | | | | | | | | | | | | |
| Number of naive | 132 | 1148 | 2194 | 2136 | 2297 | 2614 | 2661 | 2670 | 2831 | 2802 | 2317 | 2480 | 2691 | 2714 | 2872 | 2835 | 2808 | 2615 | 1698 | 44515 |
| Number GHST | 115 | 1008 | 1959 | 1905 | 2051 | 2297 | 2373 | 2281 | 2375 | 2241 | 1716 | 1743 | 1771 | 1739 | 1875 | 1710 | 1570 | 1372 | 970 | 33071 |
| Fraction GHST | 0.87 | 0.88 | 0.89 | 0.89 | 0.89 | 0.88 | 0.89 | 0.85 | 0.84 | 0.80 | 0.74 | 0.70 | 0.66 | 0.64 | 0.65 | 0.60 | 0.56 | 0.52 | 0.57 | |
| *By etiology* | | | | | | | | | | | | | | | | | | | | |
| *IGHD + ISS* | | | | | | | | | | | | | | | | | | | | |
| Number of naive | 69 | 614 | 1234 | 1239 | 1299 | 1505 | 1493 | 1506 | 1614 | 1655 | 1333 | 1496 | 1635 | 1717 | 1857 | 1798 | 1854 | 1685 | 1183 | 26786 |
| Number GHST | 62 | 567 | 1169 | 1154 | 1222 | 1387 | 1404 | 1389 | 1461 | 1449 | 1145 | 1205 | 1287 | 1267 | 1420 | 1285 | 1238 | 1046 | 802 | 21959 |
| Fraction GHST | 0.90 | 0.92 | 0.95 | 0.93 | 0.94 | 0.92 | 0.94 | 0.92 | 0.91 | 0.88 | 0.86 | 0.81 | 0.79 | 0.74 | 0.76 | 0.71 | 0.67 | 0.62 | 0.68 | |
| *Organic GHD* | | | | | | | | | | | | | | | | | | | | |
| Number of naive | 43 | 292 | 377 | 357 | 410 | 434 | 462 | 412 | 465 | 417 | 346 | 328 | 363 | 351 | 375 | 336 | 304 | 310 | 158 | 6540 |
| Number GHST | 36 | 257 | 341 | 325 | 364 | 387 | 422 | 363 | 390 | 354 | 269 | 243 | 252 | 237 | 262 | 206 | 154 | 156 | 81 | 5099 |
| Fraction GHST | 0.84 | 0.88 | 0.90 | 0.91 | 0.89 | 0.89 | 0.91 | 0.88 | 0.84 | 0.85 | 0.78 | 0.74 | 0.69 | 0.68 | 0.70 | 0.61 | 0.51 | 0.50 | 0.51 | |
| *Chronic renal insufficiency* | | | | | | | | | | | | | | | | | | | | |
| Number of naive | 1 | 6 | 23 | 45 | 50 | 79 | 85 | 87 | 60 | 101 | 90 | 101 | 91 | 83 | 71 | 77 | 95 | 72 | 44 | 1261 |
| Number GHST | 1 | 5 | 20 | 36 | 35 | 65 | 69 | 68 | 49 | 34 | 24 | 25 | 10 | 11 | 4 | 3 | 5 | 3 | 2 | 469 |
| Fraction GHST | 1.00 | 0.83 | 0.87 | 0.80 | 0.70 | 0.82 | 0.81 | 0.78 | 0.82 | 0.34 | 0.27 | 0.25 | 0.11 | 0.13 | 0.06 | 0.04 | 0.05 | 0.04 | 0.05 | |
| *Turner syndrome* | | | | | | | | | | | | | | | | | | | | |
| Number of naive | 2 | 79 | 274 | 225 | 276 | 307 | 300 | 308 | 306 | 279 | 262 | 262 | 311 | 257 | 272 | 249 | 218 | 201 | 118 | 4506 |
| Number GHST | 1 | 52 | 182 | 159 | 201 | 208 | 210 | 171 | 159 | 127 | 78 | 69 | 37 | 28 | 15 | 21 | 16 | 14 | 8 | 1756 |
| Fraction GHST | 0.50 | 0.66 | 0.66 | 0.71 | 0.73 | 0.68 | 0.70 | 0.56 | 0.52 | 0.46 | 0.30 | 0.26 | 0.12 | 0.11 | 0.06 | 0.08 | 0.07 | 0.07 | 0.07 | |

frequently, falling to approximately 50% by 2003. This decrease likely reflects the growing understanding that GHST is generally not helpful in patients with structural abnormalities of the hypothalamic–pituitary region or in patients with multiple pituitary hormone deficiencies.

Changes in GHST frequency in two other conditions, chronic renal insufficiency (CRI) and Turner syndrome, are illuminating (Table 1). The rate of GHST in subjects with CRI decreases markedly in 1994. While overall GHST usage in patients with Turner syndrome fell more gradually, it dropped dramatically between 1996 and 1997. In both disorders, the rate of GHST fell substantially during the year of FDA approval for that particular indication. This implies that clinicians might have been utilizing GHST to justify the use of GH in these disorders prior to formal FDA approval. With regard to the particular pharmacologic agents employed for GHST among NCGS patients, the use of insulin was common in 1985 (61%), but declined to a low of 16% in 2001 (Fig. 2).

Of the subjects who underwent GHST, 21,959 (66.4%) were classified as having either ISS or IGHD. Of these, 14,522 subjects were older than 4 years of age and prepubertal when GH therapy was started. Their median peak stimulated GH was 7.7 ng/mL (interquartile range 4.9–23.0). A first-year treatment growth rate was recorded for 13,134 (90.4%) patients. A weak correlation ($r = -0.16$) was observed between the peak GH concentration and the growth response in the first year of GH therapy (Fig. 3). This association accounts for only 2.7% of the variation in the first-year change in growth. Moreover, only peak GH concentrations <3 ng/mL predicted a small increase in the response to the first year of GH therapy.

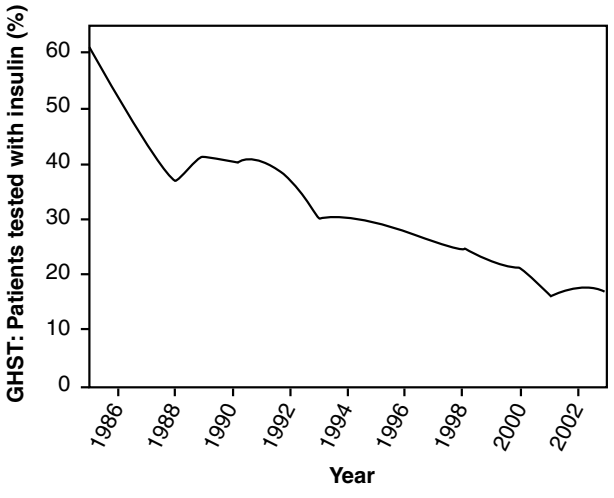


Fig. 2. Percentage of growth hormone–naive subjects who received insulin during GHST.

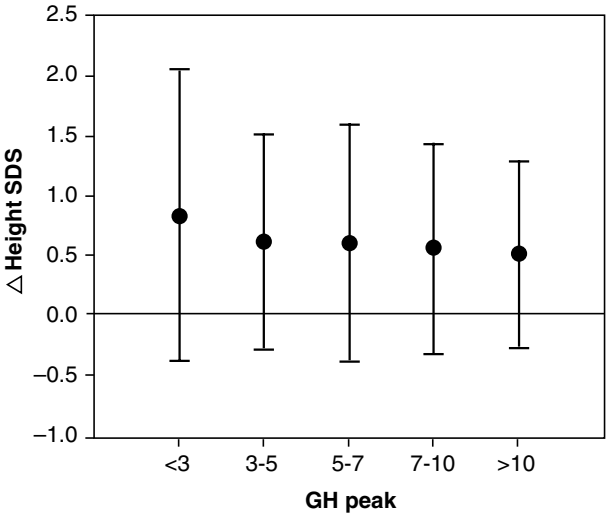


Fig. 3. Mean and range of growth response (indicated by ±2 SD) change in the height standard deviation score over the first year of GH therapy by the peak stimulated GH concentration. Data from 13,134 prepubertal patients diagnosed with ISS or isolated GHD.

## 4. Conclusions

Rates of GHST in young patients being evaluated for GH therapy have decreased over the past 18 years. This decrease has been more dramatic for selected diagnoses, such as CRI and Turner syndrome, which became FDA-approved indications for GH therapy during this time period. It will be interesting to see if the recent FDA approval of GH for the treatment of ISS will also be associated with a similar decrease in GHST frequency over the next few years. The use of insulin, arguably the most potentially dangerous of the GH stimulants, has decreased substantially over the years, reflecting an increased appreciation of the potential hazards, such as insulin-induced hypoglycemia, associated with its use.

Our study is not the first to show a poor correlation between peak GH concentrations during GHST and subsequent growth. Other investigators have also reported that peak GH values obtained during GHST are poor predictors of subsequent growth for all but the lowest peak GH values [13,14]. Interestingly, Wyatt et al. [15] noted in 1995 that while 80% of the pediatric endocrinologists who responded to their questionnaire used two GH stimulation tests, only 32% believe the results of such tests predicted a response to therapy.

Some investigators have suggested that additional standardization of GHST components – namely sex steroid priming, the stimulant combination used, and the GH assay – as well as evidence-based diagnostic criteria would help solve the current problems [16,17]. Even though this might be true, it seems unlikely to occur in the near future. In fact, many of the studies showing poor reliability of GHST have utilized consistent stimulation protocols and GH assays [3–5]. The ''real world''

situation is likely much worse. Moreover, with the advent of other diagnostic tools, particularly high-resolution cranial imaging, it is doubtful that improvements in GHST will meaningfully contribute to the diagnosis of GHD. Ultimately, the combination of GHST's poor performance with the availability of high-quality alternatives has reduced the utility of GHST to the point where it is no longer useful in the diagnosis of GHD in children and adolescents.

## Acknowledgements

This study was made possible by the able assistance of Ken Dana, all of the NCGS staff, and all of the investigators and coordinators who have provided the clinical data over the past two decades.

## References

[1] J.L. Mills, L.B. Schonberger, D.K. Wysowski, et al., Long-term mortality in the United States cohort of pituitary-derived growth hormone recipients, J. Pediatr. 144 (2004) 430–436.
[2] L.M. Gandrud, D.M. Wilson, Is growth hormone stimulation testing in children still appropriate? Growth Horm. IGF Res. 14 (2004) 185–194.
[3] P. Tassoni, E. Cacciari, M. Cau, et al., Variability of growth hormone response to pharmacological and sleep tests performed twice in short children, J. Clin. Endocrinol. Metab. 71 (1990) 230–234.
[4] H.C. Hoeck, P.E. Jakobsen, P. Vestergaard, J. Falhof, P. Laurberg, Differences in reproducibility and peak growth hormone responses to repeated testing with various stimulators in healthy adults, Growth Horm. IGF Res. 9 (1999) 18–24.
[5] A. Zadik, S.A. Chalew, Z. Gilula, A.A. Kowarski, Reproducibility of growth hormone testing procedures: a comparison between 24-h integrated concentration and pharmacological stimulation, J. Clin. Endocrinol. Metab. 71 (1990) 1127–1130.
[6] G. Marin, H.M. Domene, K.M. Barnes, et al., The effects of estrogen priming and puberty on the growth hormone response to standardized treadmill exercise and arginine-insulin in normal girls and boys, J. Clin. Endocrinol. Metab. 79 (1994) 537–541.
[7] E. Ghigo, J. Bellone, G. Aimaretti, et al., Reliability of provocative tests to assess growth hormone secretory status. Study in 472 normally growing children, J. Clin. Endocrinol. Metab. 81 (1996) 3323–3327.
[8] N. Mauras, P. Walton, M. Nicar, S. Welch, A.D. Rogol, Growth hormone stimulation testing in both short and normal statured children: use of an immunofunctional assay, Pediatr. Res. 48 (2000) 614–618.
[9] P.C. Sizonenko, P.E. Clayton, P. Cohen, et al., Diagnosis and management of growth hormone deficiency in childhood and adolescence. Part 1: diagnosis of growth hormone deficiency, Growth Horm. IGF Res. 11 (2001) 137–165.
[10] A. Shah, R. Stanhope, D. Matthew, Hazards of pharmacological tests of growth hormone secretion in childhood, BMJ 304 (1992) 173–174.
[11] A. Badaru, D.M. Wilson, Alternatives to growth hormone stimulation testing in children, Trends Endocrinol. Metab. 15 (2004) 252–258.
[12] D. Wyatt, Lessons from the National Cooperative Growth Study, Eur. J. Endocrinol. 151 (Suppl. 1) (2004) S55–59.
[13] J. Van den Broeck, N. Arends, A. Hokken-Koelega, Growth response to recombinant human growth hormone (GH) in children with idiopathic growth retardation by level of maximum GH peak during GH stimulation tests, Horm. Res. 53 (2000) 267–273.
[14] G.M. Bright, J.R. Julius, J. Lima, S.L. Blethen, Growth hormone stimulation test results as predictors of recombinant human growth hormone treatment outcomes: preliminary analysis of the National Cooperative Growth Study database, Pediatrics 104 (1999) 1028–1031.
[15] D.T. Wyatt, D. Mark, A. Slyper, Survey of growth hormone treatment practices by 251 pediatric endocrinologists, J. Clin. Endocrinol. Metab. 80 (1995) 3292–3297.
[16] R. Pandian, J.M. Nakamoto, Rational use of the laboratory for childhood and adult growth hormone deficiency, Clin. Lab. Med. 24 (2004) 141–174.
[17] C.J. Strasburger, Laboratory assessment of GH, Growth Horm. IGF Res. 8 (Suppl. A) (1998) 41–46.