PROFESSOR LUCA PERSANI (Orcid ID : 0000-0003-2068-9581)

Article type : Original Article - Europe, excluding UK

Clinical benefits of sex steroids given as a priming prior to GH provocative test or as a growth promoting therapy in peripubertal growth delays: results of a retrospective study among ENDO-ERN centers

Short title: Role of priming in peripubertal growth delay

Elena Galazzi[1*], Nicola Improda[2*], Manuela Cerbone[3,4], Davide Soranna[5], Mirella Moro[1], Letizia Maria Fatti[1], Antonella Zambon[6], Marco Bonomi [1,7], Mariacarolina Salerno[2], Mehul Dattani[3,4], Luca Persani[1,7].

[1]Department of Endocrine and Metabolic Diseases & Lab of Endocrine and Metabolic Research, Istituto Auxologico Italiano, Milan, Italy; *[2]Department of Translational Medical Sciences, Pediatric Endocrinology Unit, Federico II University of Naples* ; *[3]London Centre for Paediatric Endocrinology and Diabetes at Great Ormond Street Children's Hospital and University College London Hospitals, London, UK* ;*[4] Genetics and Genomic Medicine Programme, University College London Great Ormond Street Hospital Institute of Child Health, London, UK* ; *[5]Istituto Auxologico Italiano, IRCCS, Milan, Italy*; *[6]Dept of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy;[7]Department of Biotechnology and Translational Medicine, University of Milan, Italy*;

*These authors contributed equally to the work

Acknowledgments: The Authors want to thank Elena Monti and Harriet Gunn for their help in accessing patients' data from the London cohort. Endo-ERN is a European Reference Network co-funded by the Third European Union's Health Programme (CHAFEA Framework Partnership Agreement number: 739527 – Endo-ERN).

Declaration of Interest: Authors have nothing to disclose on the content of this research.

Fundings: This work was partially supported by Ricerca Corrente funds of Istituto Auxologico Italiano (codes: 05F401_2014 and 02C502_2005)

Data availability statement: No specific data to be shared on public repository. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Correspondence and Reprint Requests:

Luca Persani, MD PhD, University of Milan & IRCCS Istituto Auxologico Italiano; email: luca.persani@unimi.it

Summary

Objectives: Sex steroids, administered as a priming before GH stimulation tests (GHST) to differentiate between growth hormone deficiency (GHD) and constitutional delay of growth and puberty (CDGP), or as growth-promoting therapy using Low Dose Sex Steroids (LDSS) in CDGP, is much debated. We aimed to compare auxological outcomes of CDGP or GHD children undergoing primed or unprimed GHST and to evaluate LDSS treatment in CDGP.

***Design*:** Retrospective study among three pediatric University Hospitals in Italy and UK.

Methods: 184 children (72 females) aged 12.4 ± 2.08 years underwent primed (/P^+) or unprimed (/P^-) GHST and were followed-up until Final Height (FH). CDGP patients were untreated ($CDGP^-$) or received LDSS ($CDGP^+$). The cohort included 34 $CDGP^-/P^+$, 12 $CDGP^+/P^+$, 51 GHD/P^+, 29 $CDGP^-/P^-$, 2 $CDGP^+/P^-$ and 56 GHD/P^-. FH standard deviation score (SDS), Δ SDS FH-target height (TH) and degree of success ($-1 \leq \Delta$ SDS FH-SDS TH $\leq +1$) were outcomes of interest.

Results: GHD/P^+ had better FH SDS (-0.87 vs -1.49; p = 0.023) and ΔSDS FH-TH (-0.35 vs -0.77; p = 0.002) than $CDGP^-/P^+$. Overall, GHD/P^+ showed the highest degree of success (90%, p= 0.006). Regardless of priming, both rhGH and LDSS improved degree of success compared to no treatment (89% and 86% vs 63%, p= 0.0009). GHD/P^+ showed a trend towards a higher proportion of permanent GHD compared to GHD/P^- (30.43% vs 15.09%; p = 0.067).

Conclusion: In peripubertal children, priming before GHST improves diagnostic accuracy of GHST for idiopathic GHD. LDSS treatment improves auxological outcomes in CDGP.

Keywords: Puberty; idiopathic growth hormone deficiency; short stature; sex steroid priming; constitutional delay of growth and puberty; final height

MAIN TEXT

Introduction

Children in the peripubertal phase or with delayed puberty may present with growth deceleration or short stature, prompting endocrine evaluation. In this scenario, differentiating idiopathic GHD (GHD) from constitutionally delayed non-GH deficient short children (CDGP) remains a challenge[1].

GH secretion is regulated by sex steroids, as demonstrated by the threefold increase in GH secretion occurring along with an increase in gonadal steroid concentrations during puberty [2]. Blunted GH peaks after physiological or pharmacological stimuli may be found in normal pre-pubertal children [3], and thus, although GH stimulation tests (GHST) are required to diagnose GHD, their results need to be evaluated carefully and integrated with other clinical, biochemical and neuro-imaging data.

Early retesting (1-6 months after diagnosis) of children labelled as idiopathic GHD upon unprimed GHST showed normal GH responses in up to 85% of subjects, indicating poor specificity of unprimed GHST in the diagnosis of idiopathic GHD [4]. On the other hand, the administration of sex steroids before GHST (also known as priming), aimed at sensitizing the pituitary gland of pre/ peri-pubertal children to sex steroids and at increasing the secretion of GH under stimulation performed soon afterwards, has been shown to reduce false positive GHD results from 39% to 5%[3]. Current International Pediatric Guidelines give a conditional recommendation to use priming with sex steroids prior to GHST for prepubertal boys older than 11 or prepubertal girls older than 10 years with adult height prognosis above - 2 SD of mean for the reference population, aiming to prevent unnecessary GH treatment of children in whom a diagnosis of CDGP may be likely [5].

However, evidence supporting this approach is limited, as the studies available [3 6 7 8 9] report conflicting results, possibly due to the small cohorts reported, and the lack of long-term follow up.

Widespread adoption of priming in clinical practice has also been hampered by the fact that some clinicians consider it as an artificial stimulus leading to a transient increase in GH peaks, thus masking peri-pubertal GHD [10].

Other than being used to improve specificity of GHST, low dose sex steroids (LDSS) can also be given as a quasi-physiological growth-promoting therapy in those patients diagnosed as CDGP after GH deficiency has been excluded, in order to improve growth and sexual maturation and mitigate their psycho-social discomfort.[11] Nevertheless, this treatment is not universally accepted and in CDGP patients a wait-and-see approach is frequently adopted [12].

To date, long-term data on large cohorts evaluating the clinical benefits of the use of sex steroids as a pharmacological tool before GHST (priming) or as a growth-promoting therapy in CDGP patients after exclusion of GHD are scanty.

Therefore, the primary aim of our multicenter retrospective study was to evaluate if priming with sex steroids improves the diagnostic accuracy of GHST in the diagnosis of GHD, by comparing the auxological outcomes of a large cohort of GHD/CDGP children undergoing a primed GHST compared to a control group not diagnosed with the use of priming. Besides, we compared the reversal rates of idiopathic GHD at retesting between patients undergoing primed and unprimed GHST. Secondary aim of our study was to assess whether LDSS administered as a growth-promoting therapy in CDGP subjects had a significantly impact in their long-term auxological outcomes.

Material and methods

Population

We retrospectively collected auxological data of 184 children (74 females) who underwent primed or unprimed (depending on local or individual preferences) GHST in the peripubertal phase at 3 European tertiary centers from 2002 to 2015 (Naples and Milan) and from 2008 to 2015 (London). Data were retrieved from an internal audit conducted within the three hospitals after approval of the Institutional Ethic Committeee (02C502_2005). All parents gave a signed informed consent to the use of anonymous clinical data for research purposes.

All patients sought medical attention for short stature (IH, initial height) per se and/or in relation to their target height (TH), as defined by IH-TH < - 1.5 SDS and/or for growth deceleration (defined

as a height velocity - 1.0 SDS below the mean for age and sex). All patients had attained FH (height measured at growth velocity < 2 cm/year) at the time of data collection. Height was measured with a wall-mounted stadiometer to the nearest 0.1 cm.

Testicular volume was measured using a Prader orchidometer [14] and pubertal stage was determined according to the criteria of Marshall and Tanner. [15] All boys diagnosed as having G1 stage of puberty had mean testes volume between 2 and 3 mL, whereas those classified as G2 stage had mean testes volume between 4 and 6 mL.

Inclusion criteria were: age 11-14 years for males or 10-13 years for females; Tanner stage < -1 SDS according to puberty nomograms,[13] bone age no greater than chronological age. Exclusion criteria were: born small for gestational age; familial short stature; obesity or severe malnutrition; chronic diseases or steroid use; combined pituitary hormone deficiencies; syndromic patients; history of intracranial or systemic tumors; midline brain defects; signs of skeletal dysplasia.

Other systemic causes of short stature (malabsorption, coeliac disease, liver or kidney diseases, systemic autoimmune diseases) were ruled out.

The key differential diagnosis between CDGP and GHD was based on the response entity of GH peaks under GHST for evaluation of GH reserve and was indeed supported by other clinical and/or laboratory criteria. Patients were diagnosed as isolated growth hormone deficiency (GHD) if peak GH upon GHST was < 8 µg/L, possibly together with IGF-1 concentrations below two Standard Deviation Score (SDS). Among subjects with peak GH ≥ 8 µg/L, diagnosis of CDGP was established if puberty had not started yet or had a slow or stuttering progression (defined as Tanner stage < -1 SDS according to puberty nomograms) associated with at least two of these criteria; i) bone age delay > 1 year compared to chronological age; ii) being short for target height (TH); and iii) family history of pubertal delay [13].

CDGP patients were either untreated or received low-dose sex steroids (LDSS) on an individual basis until a Tanner stage 3 or 4 was achieved, in order to mitigate their psychosocial distress and improve their growth.

All patients diagnosed as GHD underwent a pituitary MRI and evaluation of pituitary gland size was based on its cranio-caudal height, measured on a sagittal t1-weighted scan.

GHD children were treated with recombinant growth hormone (rhGH) at replacement doses (20-35 µg/kg/day).

Patients were divided into 6 groups: untreated CDGP diagnosed with or without priming ($CDGP^-/P^+$, n = 34; $CDGP^-/P^-$, n = 29), CDGP receiving LDSS diagnosed with or without priming ($CDGP^+/P^+$ n =12; $CDGP^+/P^-$ n =2), GHD diagnosed with or without priming receiving rhGH therapy (GHD/P^+ n = 51; GHD/P^- n = 56) [**Figure 1 Supportive Information**].

Evaluation of GH reserve and priming

Growth hormone reserve was investigated locally by using either primed or unprimed hypothalamic stimuli (arginine, clonidine, insulin tolerance or glucagon tests, depending on local preferences). In males, primed GH stimulation was carried out following either low-dose (50 mg) or high-dose (100 mg) intramuscular Testosterone (Propionate or Enanthate) injections for 4 to 7 days, depending on local preferences. In females, the two priming regimens used before GHST were Ethinylestradiol 100 ug/day for 3 days and Stilbestrol 1 mg twice daily for 2 days.

Retesting of GH reserve was performed by using either ITT or GHRH plus arginine tests at least one month after rhGH therapy discontinuation.

Outcome assessments

Anthropometric measurements [initial height (IH), weight and body mass index (BMI)] were standardized for age and gender and expressed as SDS, according to reference standards [16]. TH was defined by Tanner's formula: (mother's height + father's heigh)/2 + 6.5 for males and – 6.5 for females. Bone age was evaluated by using either Greulich and Pyle or Tanner-Whitehouse methods [17].

Height gain from presentation was defined as the difference between standard deviation score of FH and IH (Δ SDS FH-IH). The difference between FH and TH SDS was calculated as the difference between standard deviation of final height and target height (ΔSDS FH-TH). Degree of success, defined as the proportion of patients reaching a final height in the range TH ± 1 SDS, was considered as a dichotomic variable.

Assays

Serum GH and IGF1 concentrations were measured using immunometric assays (LIASON DiaSorin, Saluggia Italy or Immulite, Siemens). GH immunoassays had a similar functional sensitivity (<0.1 μg/L) and intra- and inter-assay coefficients of variation (CV) were <3.5% and <10%, respectively. Plasma IGF-I was measured by immunometric assay after ethanol extraction. The functional sensitivity was 0.8 μg/L. The intra- and inter-assay CVs were <3.5% and 8.2%, respectively.

Statistical analysis

Continuous covariates were reported as mean and standard deviation (or, if data were not normally distributed, as median and interquartile range), whereas categorical covariates were reported as number of patients and percentage. Normal distribution of the data was verified by means of Shapiro-Wilk test. Comparisons among groups were performed by means of Analysis of variance (or Wilcoxon test) for continuous covariates and Chi-square test (or Fisher test) for categorical ones.

For each continuous outcome (SDS FH, Δ SDS FH-IH, Δ SDS FH-TH) ANOVA model was applied to verify that the outcomes, as well as the covariates, resulted statistically significantly different among groups. A head to head comparisons between groups were performed considering Tukey adjustments approach to control the inflation error I type rate. Analogously, for dichotomous outcomes (degree of success), a log-binomial model was considered in order to prevent overestimation of odds ratio due to an uncommon outcome. In this model we controlled the inflation error I type rate by means of the False Discovery Rate approach. Finally for both models we evaluated the following comparisons between groups: "$CDGP^-$ / P^+ and $CDGP^-$ / P^- vs GHD/P^+"; "$CDGP^-$ / P^+ and $CDGP^-$ / P^- vs GHD/P^-".

All tests were two tailed and a p-value < 0.05 was considered as statistically significant. All the analyses were performed by means of SAS software.

Results

Description of the population

Overall, 184 children were included: 104 from Milan, 54 from Naples and 26 from London. The main features of all groups at diagnosis and at final height are presented in **Table 1**; data

comparisons shown were confirmed in multivariate analyses and group CDGP^{+}/P^{-} was not considered due to its small size (n = 2).

Within the whole cohort, GHD/P^{-} patients were younger ($p < 0.0001$) and had bone ages more delayed compared to chronological age. Mean bone age at presentation was 11.90 years (10.10 to 12.50) for CGDP^{-}/P^{+}, 12.20 years (10.50 to 13.00) for CDGP^{+}/P^{+}, 12.00 years (10.70 to 13.00) for GHD/P^{+}, 11.00 years (9.90 to 12.50) for CDGP^{-}/P^{-}, 11.20 years (9.5 to 12.90) for CDGP^{+}/P^{-} and 9.00 years (8 to 11.30) for the GHD/P^{-} cohorts. For 3 of the groups, male predominance was greater than 70%. BMI and Tanner stage at presentation were similar between groups ($p = 0.122$ and 0.202 respectively). Among all groups, CDGP^{+}/P^{+}, GHD/P^{+} and GHD/P^{-} had the most severely compromised median SDS IH ($p= 0.017$) and ΔSDS IH-TH ($p = 0.045$) at presentation.

SDS IGF-1 at diagnosis was not different between groups ($p= 0.368$).

Priming with sex steroids pre-GHST was overall performed in 97/184 children (30 females), of which 46 (17 females) were diagnosed with CDGP and 51 (13 females) with GHD. LDSS was given as a growth promoting therapy on an individual basis in 14/77 CDGP patients, by using increasing dosages (up to 100 mg monthly) of either Propionate or Enanthate Testosterone or Oxandrolone 0.03 mg/kg/day for a median duration of 13 months (interquartile range 5-18 months). No concerns or side effects were reported in patients during or after LDSS treatment. Moreover, positive feedbacks were obtained throughout the use and after withdrawal of LDSS in terms of psycho-social wellness.

Mean GH peaks upon GHST were similar between CDGP/P^{+} and CDGP/P^{-} and between GHD/P^{+} and GHD/P^{-} cohorts (data not shown).

Pituitary hypoplasia was found in 12 out of 51 GHD/P+ children, and in 8 out of 56 GHD/P- patients, while the rest had a normal MRI. The prevalence of having a small anterior pituitary gland was not statistically different between the two groups.

Among CDGP patients, only one female patient belonging to the CDGP^{+}/P^{+} group and presenting with a SDS IH -2.11, a Δ SDS IH-TH of -2.53, a bone age delay of 2 years and a B1 Tanner stage of puberty, turned out to have permanent hypogonadotropic hypogonadism and reached a stature well within her target height after sequential estradiol and estrogen-progestins patches.

Multivariate analysis

Mean SDS FH, ΔSDS FH-TH and ΔSDS FH-IH were statistically different between CDGP/P^+, CDGP$^+$/P^+, GHD/P^+, CDGP$^-$/P^-, GHD/P^- groups (p-value =0.017, p-value= 0.007 and p-value=0.002, respectively). [**Figure 1, 2** and **3**]

More specifically, the head to head comparisons for SDS FH were significantly lower for CDGP$^-$/P^+ group compared to both GHD groups: CDGP$^-$/P^+ vs GHD/ P^+ LS means (SE): -1.49 (0.16) vs -0.87 (0.14), p-value =0.023; CDGP$^-$/P^+ vs GHD/ P^-: -1.49 (0.16) vs -0.87 (0.14), p-value = 0.022. [**Figure 1**].

Concordantly, ΔSDS FH-TH proved to be significantly higher only between CDGP$^-$/P^+ and GHD/P^+ [-0.77 (0.13) vs -0.35 (0.11), p-value = 0.002]. [**Figure 2**]

Hence, the differences in SDS FH and ΔSDS FH-TH between CDGP$^-$/P^- and both GHD groups were similar. [**Figures 1** and **2**]

Better long term auxological outcomes in GHD patients diagnosed upon a primed GHST (GHD/P^+) compared to GHD/P^- were also confirmed in the comparison between overall untreated CDGP patients (groups CDGP$^-$/P^+ and CDGP$^-$/P^- taken together). In fact, in GHD/P^+ compared to untreated CDGP patients overall, SDS FH nearly achieved statistical significance with a p-value approaching the significant threshold [SDS FH mean difference -0.328 (SE 0.177); p-value =0.065), whilst ΔSDS FH-TH did achieve a significant statistical difference [ΔSDS FH-TH mean difference -0.454 (SE 0.145); p-value = 0.004]. In contrast, these outcomes were not both concordantly statistically significant among GHD/P^- patients compared to untreated CDGP patients overall [SDS FH: mean difference -0.326 (SE 0.174), p-value= 0.064; ΔSDS FH-TH: mean difference -0.238 (SE 0.143), p-value= 0.196].

Similarly, catch-up growth from presentation (ΔSDS FH-IH) was lower in CDGP$^-$/P^+ compared to both GHD groups [CDGP$^-$/P^+ vs GHD/ P^+: 0.81 (0.13) vs 1.50 (0.11), p-value < 0.001; and CDGP$^-$/P^+ vs GHD/P^-: 0.81 (0.13) vs 1.33 (0.11), p-value = 0.018], whereas it was found to be similar between CDGP$^-$/P^- and both GHD groups. [**Figure 3**]. A lower catch-up growth from presentation was indeed confirmed considering untreated CDGP patients overall vs GHD/P^+ or vs GHD/P^- [ΔSDS FH-IH: vs GHD/P^+ mean difference -0.515 (SE 0.145), p-value =0.001; vs GHD/P^- mean difference -0.342, (SE 0.143), p-value =0.018].

Among all groups, the highest degree of success was gained by GHD/P^+ (p = 0.006) [**Table 1**]. However, regardless of priming, treatment with rhGH was associated with better long term auxological outcomes compared to untreated CDGP. In fact, the head to head comparison between groups showed that GHD/P^+ had a higher degree of success compared to untreated CDGP groups (GHD/P^+ vs CDGP$^-$/P^+: 90% vs 65%, p-value = 0.012; GHD/P^+ vs CDGP$^-$/P^-: 90% vs 62% p-value = 0.012) and analogously that GHD/P^- group had a higher degree of success compared to untreated CDGP (GHD/P^- vs CDGP$^-$/P^+: 86% vs 65%, p-value = 0.015; GHD/P- vs CDGP$^-$/P^-: 86% vs 62%, p-value = 0.015)**.** [**Figure 4**] These results were also confirmed in the comparisons between both GHD groups and untreated CDGP patients overall: proportions of success for CDGP patients overall (63%) vs GHD/P^+ (90%), p-value =0.001; vs GHD/P^- (86%), p-value = 0.003.

Finally, considering the entire cohort and grouping patients for diagnosis and treatment, regardless of the use of priming before GHST, both treatments with either rhGH or LDSS were associated with a higher degrees of success compared to the untreated CDGP cohorts (GHD 89% vs CDGP$^+$ 86% vs CDGP$^-$ 63% p= 0.0009) [**Figure. 5**].

Retesting

After attainment of final height, 46 out of 51 (90.19%) patients in group GHD/P^+ (mean age 17.76 ±1.43 years) and 53 out of 56 (92.52%) patients in group GHD/P^- (mean age 16.69±1.05 years) were retested for GH reserve and a trend towards a higher proportion of permanent GHD was documented in patients diagnosed upon a primed GHST (GHD/P^+ vs GHD/P^-: 30.43% vs 15.09%; p = 0.067) [**Figure 2 Supporting Information**].

No statistically significant differences were found in the reversal rates of GHD at retesting between subjects with normal (n=87) vs abnormal (n=20) aspects of anterior pituitary gland on MRI (81% in those with normal MRI vs 70% in those with small anterior pituitary, N.S.).

No significant post therapy IGF-1 differences were observed between normalized vs persistent GHD at retesting both in GHD/P^+ (normalized: 0.23 ± 1.21 vs persistent GHD: 0.97 ± 1.21, N.S.) and in GHD/P^- groups (normalized: -0.46±0.97 vs persistent GHD: -1.04 ± 1.64, N.S.).

Discussion

To our knowledge this is the first study investigating long term auxological outcomes in a large cohort of pre/peri-pubertal children diagnosed with CDGP or GHD, undergoing a primed GHST, in comparison to a group of children who did not receive priming.

The results of our study indicate that priming with sex steroids prior to GHST may improve the diagnostic accuracy of the test in the diagnosis of GHD.

In fact, we documented that GHD patients diagnosed upon primed GHST reach a FH that is greater and closer to the TH, in comparison with untreated CDGP. Conversely, in patients in whom the diagnosis was made upon an unprimed GHST, we found no differences in auxological outcomes between treated GHD and untreated CDGP patients, suggesting that priming plays a key role in selecting those children who may benefit the most from rhGH treatment.

Important strengths of our study are both the involvement of a homogeneous cohort (with respect to age, gender, pubertal and auxological status at diagnosis) as well as the inclusion of a control group of unprimed CDGP/GHD. In addition, we ruled out secondary causes of GHD or short stature. Therefore, our cohort is representative of a population of Caucasian children presenting with idiopathic short stature and delayed puberty. However it has some limitations, inherent to the retrospective design of the study: low sample size of some groups, variable tests performed, different laboratory assays, use of different regimens of sex steroids either as priming before GHST either as a growth promoting therapy.

Along with the fact that during puberty, the increase of sex steroids circulating levels is associated with an increased GH pulse amplitude, higher IGF-I concentrations, and increased anterior pituitary size [18 19], most [3 6 7], but not all [9 20], studies reported that priming with sex steroids prior to GHST increases GH peaks in response to provocation in both normal and short children, improving diagnostic accuracy of GHST from 90% to 95% [6]. Although, as already mentioned, since 2016 priming before GHST has been recommended by Pediatric Endocrine Society Guidelines, its use in clinical practice is still limited, considered by some clinicians an unphysiological method leading to only a transient increase in GH peak and masking peri-pubertal GHD [10], possibly exposing patients to the risk of side effects, and due to the lack of robust evidence on its diagnostic advantages and long-term outcomes. In fact, so far, only one study explored final height in 50 untreated peripubertal boys with subnormal unprimed but normal primed stimulated GH peaks, reporting a normal FH that layed within the TH range, regardless of the priming protocol used [21].

Auxological outcomes of our GHD cohort (either/P^+ or/P^-) are in line with previous data for Caucasian patients reporting SDS FH ranging between -0.7 and -1.11 and ΔSDS FH-TH between -0.17 and -0.6 [22] [23]. Growth response to rhGH has been previously found to correlate with anthropometric variables at the start of treatment (SDS IH, chronological age, bone age, pubertal status), severity of GHD and genetic potential [24]; however, the results of our study indicate the need for including also the use of priming in models predicting the response to rhGH.

Interestingly, although the overall proportion of re-confirmed idiopathic GHD at final height was similar to that reported in other papers [4] [25] [26] we found a trend towards a higher proportion of permanent GHD in primed compared to unprimed GHD patients. We believe that, although this comparison only approaches the significance threshold, it is of much interest and could become statistically significant by increasing the sample size. Taken together, these results indicate the poor diagnostic accuracy of unprimed GHST in the peri-pubertal phase, likely due to the physiological transient blunting of GH response to stimulation in this period of life [27].

IGF-1 values were not useful to differentiate between pre/peripubertal CDGP and GHD in the initial diagnostic process, nor to differentiate true and permanent GHD after retesting from transient forms of IGHD. This could be possibly related to the poor sensitivity of IGF-1 and its fluctuations in relation to pubertal and nutritional status.

Interestingly, no correlation was found between anterior pituitary size and reconfirmed GHD at retesting, possibly due to the concept that a borderline small anterior pituitary gland is difficult to differentiate from low-normally sized pituitary glands given the very few studies available calculating the anterior pituitary volume/size for chronological age and gender [28] [29] and its physiological variations along with pubertal status [30] .

An additional aim of our study was to evaluate the effects of LDSS given as a growth-promoting treatment in CDGP patients after GH deficiency was excluded.

In keeping with results of previous studies [31] [32] [33] our data indicate that, if left untreated, patients with CDGP fail to achieve their genetic potential, and this is even more evident when priming before GHST is performed. This result may be possibly due to the fact that the CDGP cohort diagnosed with the use of priming may have included some partial/mild forms of GHD, displaying normal transient GH peaks under primed stimuli. Even if there is no consensus on the treatment protocol to adopt for CDGP children, we have demonstrated that regardless of priming before GHST, administering variable schemes of LDSS is a safe and well tolerated approach and may

favour the achievement of height potential, avoiding the risk of premature closure of bone cartilages secondary to an excessive bone age maturation, as previously suggested in smaller cohorts. [34] [35]

Although an objective evaluation was not performed, in our patients LDSS treatment proved to be a safe approach to improve the psycho-social wellness of CGDP patients.

Data evaluating long-term auxological outcomes for females are scanty, especially in CDGP, which typically shows a male preponderance.[31]

In our study we failed to find any significant gender differences between CDGP and GHD, possibly due to the low female representation in our cohort.

We acknowledge that larger prospective studies, particularly in females, would be of great interest to confirm these data and to prove whether LDSS would have a definite advantage on FH in girls. Moreover, a specific focus on optimisation of priming schemes to adopt before GHST would be of great relevance to clinicians, but this may need a larger prospective randomized study, where specific dosages and formulations are taken into account.

Only one female from the whole cohort turned out to be affected by hypogonadotropic hypogonadism, thus ideally confirming the clinical impression that the great majority of children affected by true hypogonadotropic hypogonadism escape retardation of linear growth and do not present with short stature.[36]

In conclusion, in pre or peri-pubertal short subjects, priming with sex steroids prior to provocative testing for GH reserve seems to improve the diagnostic accuracy of provocative testing for GH reserve, potentially avoiding cumbersome and less effective rhGH treatment in children with a false positive GHD diagnosis following an unprimed GHST. Indeed, we documented the highest degree of success, as well as the highest proportion of reconfirmed GHD at retesting after attainment of final height in GHD subjects diagnosed upon a primed GHST. Moreover, administration of LDSS in CDGP patients after exclusion of GHD proved to be effective in improving their auxological outcomes. Further studies are however needed to confirm these data in larger populations, to evaluate gender differences in growth and/or response to LDSS in CDGP patients and to compare the efficacy of different priming regimens in improving the diagnostic accuracy of GHST.

i. References

1. Murray PG, Dattani MT, Clayton PE. Controversies in the diagnosis and management of growth hormone deficiency in childhood and adolescence. *Arch Dis Child*. 2016;101(1):96-100.

2. Delemarre-van de Waal HA, van Coeverden SC, Rotteveel J. Hormonal determinants of pubertal growth. *J Pediatr Endocrinol Metab*. 2001;14 Suppl 6:1521-1526.

3. Marin G, Domené HM, Barnes KM, Blackwell BJ, Cassorla FG, Cutler GB. 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*. 1994;79(2):537-541.

4. Loche S, Bizzarri C, Maghnie M, et al. Results of early reevaluation of growth hormone secretion in short children with apparent growth hormone deficiency. *J Pediatr*. 2002;140(4):445-449.

5. Grimberg A, DiVall SA, Polychronakos C, et al. Guidelines for Growth Hormone and Insulin-Like Growth Factor-I Treatment in Children and Adolescents: Growth Hormone Deficiency, Idiopathic Short Stature, and Primary Insulin-Like Growth Factor-I Deficiency. *Horm Res Paediatr*. 2016;86(6):361-397.

6. MartÍnez AS, Domené HM, Ropelato MG, et al. Estrogen Priming Effect on Growth Hormone (GH) Provocative Test: A Useful Tool for the Diagnosis of GH Deficiency [1]. *J Clin Endocrinol Metab*. 2000;85(11):4168-4172.

7. Molina S, Paoli M, Camacho N, Arata-Bellabarba G, Lanes R. Is Testosterone and Estrogen Priming Prior to Clonidine Useful in the Evaluation of the Growth Hormone Status of Short Peripubertal Children? *J Pediatr Endocrinol Metab*. 2008;21(3).

8. Müller G, Keller A, Reich A, et al. Priming with Testosterone Enhances Stimulated Growth Hormone Secretion in Boys with Delayed Puberty. *J Pediatr Endocrinol Metab*. 2004;17(1).

9. Soliman A, Adel A, Sabt A, Elbukhari E, Ahmed H, Sanctis V. Does priming with sex steroids improve the diagnosis of normal growth hormone secretion in short children? *Indian J Endocrinol Metab*. 2014;18(7):80.

10. Lazar L, Phillip M. Is Sex Hormone Priming in Peripubertal Children prior to Growth Hormone Stimulation Tests Still Appropriate? *Horm Res Paediatr*. 2010;73(4):299-302.

11. Stanhope R, Buchanan CR, Fenn GC, Preece MA. Double blind placebo controlled trial of low dose oxandrolone in the treatment of boys with constitutional delay of growth and puberty. *Arch Dis Child*. 1988;63(5):501-505.

12. Dye AM, Nelson GB, Diaz-Thomas A. Delayed Puberty. *Pediatr Ann*. 2018;47(1):e16-e22.

13. Stalman SE, Hellinga I, Wit JM, Hennekam RCM, Kamp GA, Plötz FB. Growth failure in adolescents: etiology, the role of pubertal timing and most useful criteria for diagnostic workup. *J Pediatr Endocrinol Metab*. 2016;29(4):465-473.

14. Zachmann M, Prader A, Kind HP, Häfliger H BH. Testicular volume during adolescence. Cross-sectional and longitudinal studies. *Helv Paediatr Acta*. 1974;29:61-72.

15. Marshall WA TJ. Variations in pattern of pubertal changes in boys. *Arch Dis Child*. 1970;45:13-23.

16. Cacciari E, Milani S, Balsamo A, et al. Italian cross-sectional growth charts for height, weight and BMI (2 to 20 yr). *J Endocrinol Invest*. 2006;29(7):581-593.

17. Milner GR, Levick RK, Kay R. Assessment of bone age: a comparison of the Greulich and Pyle, and the Tanner and Whitehouse methods. *Clin Radiol*. 1986;37(2):119-121.

18. Mauras N, Rogol AD, Haymond MW, Veldhuis JD. Sex steroids, growth hormone, insulin-like growth factor-1: neuroendocrine and metabolic regulation in puberty. *Horm Res*. 1996;45(1-2):74-80.

19. Kerrigan JR, Rogol AD. The impact of gonadal steroid hormone action on growth hormone secretion during childhood and adolescence. *Endocr Rev*. 1992;13(2):281-298.

20. Wilson DM, Dotson RJ, Neely EK, Cohen P, Hintz RL, Rosenfeld RG. Effects of estrogen on growth hormone following clonidine stimulation. *Am J Dis Child*. 1993;147(1):63-65.

21. Gonc EN, Kandemir N, Ozon A, Alikasifoglu A. Final heights of boys with normal growth hormone responses to provocative tests following priming. *J Pediatr Endocrinol Metab*. 2008;21(10):963-971.

22. Reiter EO, Price DA, Wilton P, Albertsson-Wikland K, Ranke MB. Effect of growth

hormone (GH) treatment on the near-final height of 1258 patients with idiopathic GH deficiency: analysis of a large international database. *J Clin Endocrinol Metab*. 2006;91(6):2047-2054.

23. Pozzobon G, Partenope C, Mora S, Garbetta G, Weber G, Barera G. Growth hormone therapy in children: predictive factors and short-term and long-term response criteria. *Endocrine* 2019; 66 (3): 614-621.

24. Kristrom B, Karlberg J, Albertsson-Wikland K. Prediction of the growth response of short prepubertal children treated with growth hormone. *Acta Pædiatrica*. 1995;84(1):51-57.

25. Tauber M, Moulin P, Pienkowski C, Jouret B, Rochiccioli P. Growth hormone (GH) retesting and auxological data in 131 GH-deficient patients after completion of treatment. *J Clin Endocrinol Metab*. 1997;82(2):352-356.

26. Bonfig W, Bechtold S, Bachmann S, et al. Reassessment of the optimal growth hormone cut-off level in insulin tolerance testing for growth hormone secretion in patients with childhood-onset growth hormone deficiency during transition to adulthood. *J Pediatr Endocrinol Metab*. 2008;21(11):1049-1056.

27. Vuralli D, Gonc EN, Ozon ZA, Alikasifoglu A, Kandemir N. Clinical and laboratory parameters predicting a requirement for the reevaluation of growth hormone status during growth hormone treatment: Retesting early in the course of GH treatment. *Growth Horm IGF Res*. 2017;34:31-37.

28. Yadav P, Singhal S, Chauhan S, Harit S. MRI Evaluation of Size and Shape of Normal Pituitary Gland: Age and Sex Related Changes. *J Clin Diagnostic Res*. 2017;11(12)TC01-TC04.

29. Han X, Xiu J, Huang Z, et al. Three-dimensional magnetic resonance volumetry of the pituitary gland is effective in detecting short stature in children. *Exp Ther Med*. 2014;8(2):551-556.

30. Doraiswamy PM, Potts JM, Axelson DA, et al. MR assessment of pituitary gland morphology in healthy volunteers: age- and gender-related differences. *AJNR Am J Neuroradiol*. 13(5):1295-1299.

31. Poyrazoğlu Ş, Günöz H, Darendeliler F, Saka N, Bundak R, Baş F. Constitutional delay of growth and puberty: From presentation to final height. *J Pediatr Endocrinol Metab*. 2005;18(2):171-179.

32. Crowne EC, Shalet SM, Wallace WH, Eminson DM, Price DA. Final height in girls with untreated constitutional delay in growth and puberty. *Eur J Pediatr*. 1991;150(10):708-712.

33. Crowne EC, Shalet SM, Wallace WH, Eminson DM, Price DA. Final height in boys with untreated constitutional delay in growth and puberty. *Arch Dis Child*. 1990;65(10):1109-1112.

34. Martin MM, Martin ALA, Mossman KL. Testosterone treatment of constitutional delay in growth and development: effect of dose on predicted versus definitive height. *Acta Endocrinol (Copenh)*. 1986;113(4_Suppl):S147-S152.

35. Zachman M, Stüde S PA. Short-term testosterone treatment at bone age of 12 to 13 years does not reduce adult height in boys with constitutional delay of growth and adolescence. *Helv Paediatr Acta*. 1987;42:21-28.

36. Reinehr, T., Hoffmann, E., Rothermel, J., Lehrian, T. J., & Binder, G. (2019). Characteristic dynamics of height and weight in preschool boys with constitutional delay of growth and puberty or hypogonadotropic hypogonadism. *Clin Endocrinol*. 2019; *91*(3): 424–431

Figure legends

Figure 1. *SDS FH mean and related 95% confidence interval for each group. The significant p-values of Tukey adjustment head to head comparisons are reported; SDS= Standard Deviation Score; FH=Final Height; CDGP = constitutional delay of growth and puberty; GHD=growth hormone deficiency; CDGP$^-$/P$^+$= untreated CDGP diagnosed with priming; CDGP$^+$/P$^+$ = CDGP diagnosed with priming treated with low dose sex steroids (LDSS); GHD/P$^+$ = GHD diagnosed with priming, treated with rhGH; CDGP$^-$/P$^-$ = untreated CDGP diagnosed without priming; GHD/P$^-$ = GHD diagnosed without priming, treated with rhGH*

Figure 2. *ΔSDS FH-TH mean and related 95% confidence interval for each group. The significant p-values of Tukey adjustment head to head comparisons are reported; SDS= Standard Deviation Score; Δ SDS FH-TH = difference between SDS of final height and target height; CDGP = constitutional delay of growth and puberty; GHD=growth hormone deficiency; CDGP$^-$/P$^+$= untreated CDGP diagnosed with priming; CDGP$^+$/P$^+$ = CDGP diagnosed with priming treated with low dose sex steroids (LDSS); GHD/P$^+$ = GHD diagnosed with priming, treated with rhGH; CDGP$^-$/P$^-$ = untreated CDGP diagnosed without priming; GHD/P$^-$ = GHD diagnosed without priming, treated with rhGH*

Figure 3. *ΔSDS FH-IH mean and related 95% confidence interval for each group. The significant p-values of Tukey adjustment head to head comparisons are reported; SDS= Standard Deviation Score; Δ SDS FH-IH = difference between SDS of final height and initial height; CDGP = constitutional delay of growth and puberty; GHD=growth hormone deficiency; CDGP$^-$/P$^+$= untreated CDGP diagnosed with priming; CDGP$^+$/P$^+$ = CDGP diagnosed with priming treated with low dose sex steroids (LDSS); GHD/P$^+$ = GHD diagnosed with priming, treated with rhGH; CDGP$^-$/P$^-$ = untreated CDGP diagnosed without priming; GHD/P$^-$ = GHD diagnosed without priming, treated with rhGH.*

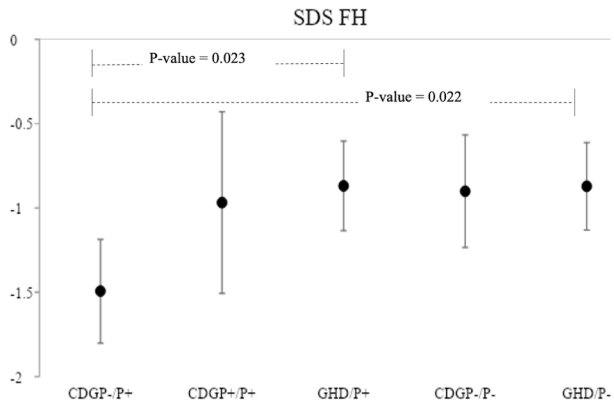
Figure 4. *Proportions of success degree and related 95% confidence interval for each group.*
The significant p-values of False Discovery Rate adjustment head to head comparisons are reported. CDGP = constitutional delay of growth and puberty; GHD = growth hormone deficiency; CDGP$^-$/P$^+$ = untreated CDGP diagnosed with priming; CDGP$^+$/P$^+$ = CDGP diagnosed with priming treated with low dose sex steroids (LDSS); GHD/P$^+$ = GHD diagnosed with priming, treated with rhGH; CDGP$^-$/P$^-$ = untreated CDGP diagnosed without priming; GHD/P$^-$ = GHD diagnosed without priming, treated with rhGH.

Figure 5. *Proportions of degree of success/no success after grouping patients for treatment, regardless of priming adoption before GHST (growth hormone stimulation test) as a diagnostic tool. Comparisons among groups were performed by means of Chi-square test.*
CDGP = constitutional delay of growth and puberty; GHD = growth hormone deficiency; LDSS = low dose sex steroids.

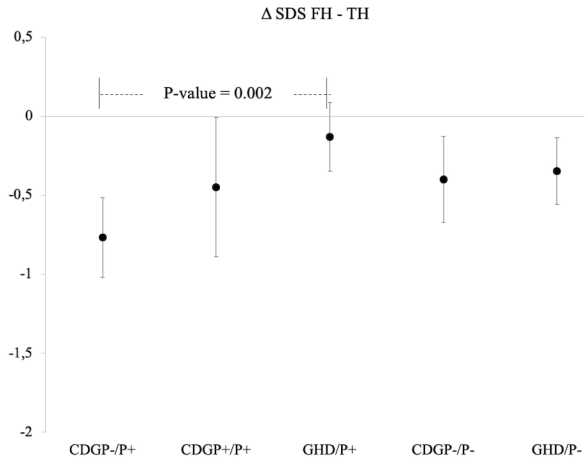
Table 1. Description of the study population. Comparisons between groups were carried out excluding group CDGP$^+$/P$^-$ because composed by 2 patients only.

	CDGP$^-$/P$^+$ (N=34)	CDGP$^+$/P$^+$ (N=12)	GHD/P$^+$ (N=51)	CDGP$^-$/P$^-$ (N=29)	CDGP$^+$/P$^-$ (N=2)	GHD/P$^-$ (N=56)	P-value*
Characteristics at cohort entry							
Chronological age, years *Median [IQ]*	12.60 [11.50 to 14.00]	13.50 [13.10 to 14.10]	13.00 [11.90 to 14.50]	12.60 [11.10 to 14.00]	13.75 (0.64)	11.10 [10.00 to 13.00]	<0.0001ǂ
Male *N (%)*	18 (53.0%)	11 (92.0%)	38 (75.0%)	15 (52.0%)	2 (100.0%)	28 (50.0%)	0.010‡
BMI SDS *Median [IQ]*	-0.85 [-1.50 to -0.20]	-1.25 [-1.66 to -0.55]	-1.15 [-1.86 to -0.29]	-0.81 [-1.65 to 0.30]	-0.55 [-2.64 to 1.55]	-0.39 [-1.45 to 0.49]	0.122ǂ
Puberty *N (%)*							
- G1	9 (26.5%)	6 (50.0%)	17 (33.3%)	7 (24.1%)	2 (100%)	17 (30.4%)	
- G2	9 (26.5%)	5 (41.7%)	21 (41.2%)	8 (27.6%)	0 (0%)	12 (21.4%)	0.202‡
- B1	11 (32.3%)	0 (0.0%)	8 (15.7%)	11 (37.9%)	0 (0%)	17 (30.4%)	
- B2	5 (14.7%)	1 (8.3%)	5 (9.8%)	3 (10.3%)	0 (%)	11 (19.6%)	
IGF-1 SDS *Mean (SD)*	-0.99 (0.97)	-1.50 (1.14)	-1.51 (1.34)	-1.40 (1.07)	-1.27 (.)	-1.39 (1.23)	0.368†
IH SDS *Median [IQ]*	-2.10[-2.60 to -1.90]	-2.43 [-3.00 to -2.00]	-2.43 [-3.00 to -2.00]	-2.00 [-2.80 to -1.62]	-1.52[-2.06 to -0.97]	-2.15 [-2.47 to -1.80]	0.017ǂ
TH SDS, cm *Mean (SD)*	-0.76 (0.66)	-0.37 (0.71)	-0.68 (0.83)	-0.53 (0.58)	-0.55 (0.96)	-0.62 (0.80)	0.539†
Δ IH-TH SDS, cm *Median [IQ]*	-1.53 [-1.75 to -1.10]	-1.98[-2.40 to -1.00]	-1.98[-2.40 to -1.00]	-1.50[-2.10 to -1.20]	-0.97 [-1.10 to -0.83]	-1.60[-2.09 to -0.84]	0.045ǂ
Characteristics at final visit							
BMI SDS *Median [IQ]*	-0.77 (1.30)	-0.86 (1.17)	-0.56 (1.16)	-0.70 (1.05)	-1.58 (1.42)	-0.35 (1.06)	0.363†
FH SDS, cm *Mean (SD)*	-1.5 (0.88)	-0.86 (1.07)	-0.81 (0.93)	-0.9 (1.13)	-0.44 (1.41)	-0.93 (0.71)	0.011†
FH-IH Δ SDS *Mean (SD)*	0.79 (0.98)	1.57 (0.68)	1.68 (0.89)	1.15 (0.87)	1.08 (0.64)	1.18 (0.82)	0.0001†
FH-TH Δ SDS, *Mean (SD)*	-0.74 (0.99)	-0.48 (0.89)	-0.12 (0.72)	-0.37 (0.95)	0.11 (0.45)	-0.31 (0.77)	0.025†
Degree of success *N (%)*	22 (65%)	9 (75%)	46 (90%)	18 (62%)	2 (100%)	48 (86%)	0.006‡

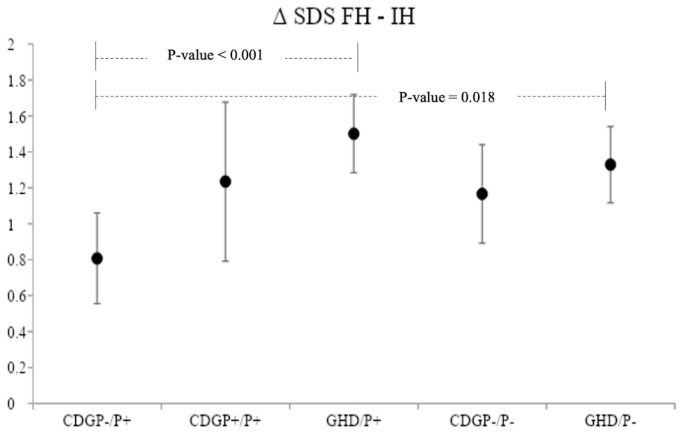
IQ= Interquartile range; *ǂ Kruskal-Wallis test; †Anova model; ‡Chi-square test



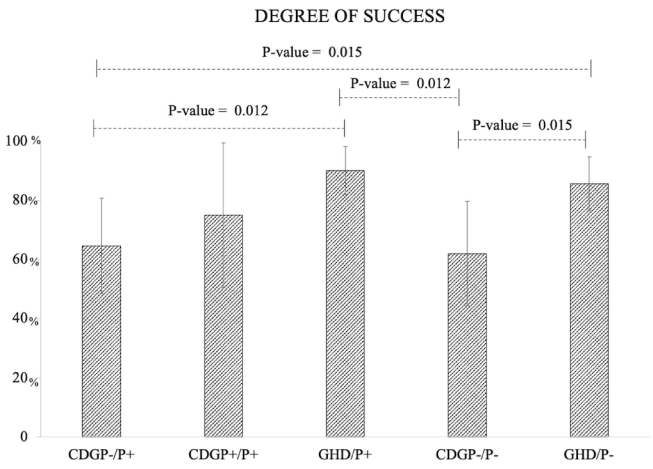
cen_14337_f1.tiff



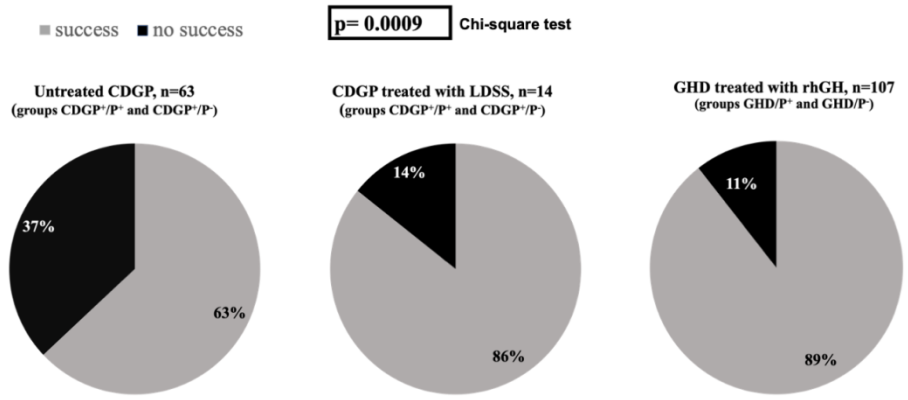
cen_14337_f2.tiff



cen_14337_f3.tiff



cen_14337_f4.tiff



cen_14337_f5.tiff