



GHRH + arginine test and body mass index: do we need to review diagnostic criteria for GH deficiency?

V. Gasco¹ · D. Cuboni¹ · E. Varaldo¹ · F. Bioletto¹ · A. M. Berton¹ · C. Bona¹ · N. Prencipe¹ · E. Ghigo¹ · M. Maccario¹ · S. Grottoli¹

Received: 19 November 2022 / Accepted: 27 March 2023 / Published online: 16 April 2023
© The Author(s) 2023

Abstract

Introduction The proportion of patients with low GH response to provocative tests increases with the number of other pituitary hormone deficiencies, reason why in panhypopituitary patients GH stimulation tests may be unnecessary to diagnose GH deficiency (GHD)

Purpose To re-evaluate the diagnostic cut-offs of GH response to GHRH + arginine (ARG) test related to BMI, considering the patients' pituitary function as the gold standard for the diagnosis of GHD.

Methods The GH responses to GHRH + ARG were studied in 358 patients with history of hypothalamic-pituitary disease. GHD was defined by the presence of at least 3 other pituitary deficits (n = 223), while a preserved somatotrophic function was defined by the lack of other pituitary deficits and an IGF-I SDS ≥ 0 (n = 135). The cut-off with the best sensitivity (SE) and specificity (SP), was identified for each BMI category using the ROC curve analysis. To avoid over-diagnosis of GHD we subsequently searched for the cut-offs with a SP $\geq 95\%$.

Results The best GH cut-off was 8.0 $\mu\text{g/l}$ (SE 95%, SP 100%) in lean, 7.0 $\mu\text{g/l}$ (SE 97.3%, SP 82.8%) in overweight, and 2.8 $\mu\text{g/l}$ (SE 84.3%, SP 91.7%) in obese subjects. The cut-off with a SP $\geq 95\%$ was 2.6 $\mu\text{g/l}$ (SE 68.5%, SP 96.6%) in overweight and 1.75 $\mu\text{g/l}$ (SE 70.0%, SP 97.2%) in obese subjects.

Conclusions This is the first study that evaluates the diagnostic cut-offs of GH response to GHRH + ARG related to BMI using a clinical definition of GHD as gold standard. Our results suggest that with this new approach, the GHRH + ARG cut-offs should be revised to avoid GHD over-diagnosis.

Keywords GHRH + arginine · GH deficiency · Obesity · Diagnosis · Sensibility · Specificity

Introduction

Adult GH deficiency (GHD) is a pathological condition that may occur during childhood or adult life, resulting from genetic or congenital disorders of pituitary development, as well as be secondary to central nervous system (CNS) injury including tumors, surgery, trauma, radiation or infiltrative diseases [1, 2]. This condition is characterized by altered glucose and lipid metabolism, derangement of body composition, premature atherosclerosis, osteoporosis,

impaired quality of life, and increased mortality [3–6]. GH replacement therapy improves most of these abnormalities [7, 8]. However, to date, a reduction in cardiovascular mortality during GH treatment remains to be demonstrated, and considering the high cost of the replacement treatment [9] and its potential long-term risks, it is essential to establish the correct diagnosis so that appropriate GH replacement is offered only to adults who are truly GH-deficient.

Adult GHD diagnosis depends on the demonstration of a reduced peak serum GH level in response to one or more GH stimulation tests [2, 10, 11].

Obesity is a state of functional GHD, with decreased spontaneous secretion, pulses, and half-life of GH [12–14]. Moreover, a decreased GH responsiveness to all stimulation tests has been demonstrated in obesity as well as in subjects with abdominal adiposity [15–18]. Therefore, one of the warnings in interpreting the results of most GH stimulation

✉ S. Grottoli
silvia.grottoli@unito.it

V. Gasco
valentina.gasco@unito.it

¹ Department of Medical Science, Division of Endocrinology, Diabetes and Metabolism, University of Turin, Turin, Italy

tests in adults is the high prevalence of obesity in the general population, as well as in patients with acquired hypothalamic-pituitary disease [19].

Insulin tolerance test (ITT) is the gold standard for the diagnosis of adult GHD. At the same time, GH-releasing hormone (GHRH) in combination with GH-releasing peptide-6 (GHRP-6) or arginine (ARG) are recognized as equally faithful tests [2, 10, 11], and glucagon stimulation is considered of good diagnostic value [20]. To date all these tests have clear cut-offs and normative BMI based reference values [21–24]. Recently, the macimorelin test has been demonstrated to be safe, effective, highly reproducible [25, 26] and therefore in December 2017 the FDA (https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/205598s0001bl.pdf) and in January 2019 the EMA (<https://www.ema.europa.eu/en/medicines/human/EPAR/ghryvelin-previously-macimorelin-aeterna-zentaris>) approved oral macimorelin for use as a diagnostic test for adult GHD.

However, it should be considered that, unlike previous tests, to date macimorelin has no specific cut-offs for overweight or obese patients. Lastly, it should be considered that all previous studies have evaluated the GH response to various stimulus tests by comparing the response of patients who are certainly GHD with normal subjects [22–24, 27, 28] and/or by using another stimulation test as gold standard [21, 25–29]. This approach is exposed to possible bias. In particular, comparing patients with normal subjects, even if matched by sex, age and BMI, is at risk of overestimating the diagnostic accuracy of the cut-offs thus identified [30]. Moreover, comparing a particular test with another diagnostic test, considered the gold standard, is at risk of overestimating or underestimating the diagnosis in a similar way to what the gold standard itself does. The proportion of patients with low GH response to provocative tests increases with the number of other pituitary hormone deficiencies and several studies involving panhypopituitary patients have shown that under certain circumstances GH stimulation tests may be unnecessary to diagnose GHD [2, 10, 11]. Aim of this study was to re-evaluate the diagnostic cut-offs of GH response to GHRH + ARG test related to BMI. To this aim the patients' pituitary function was considered as the gold standard for the diagnosis or exclusion of GHD.

Materials and methods

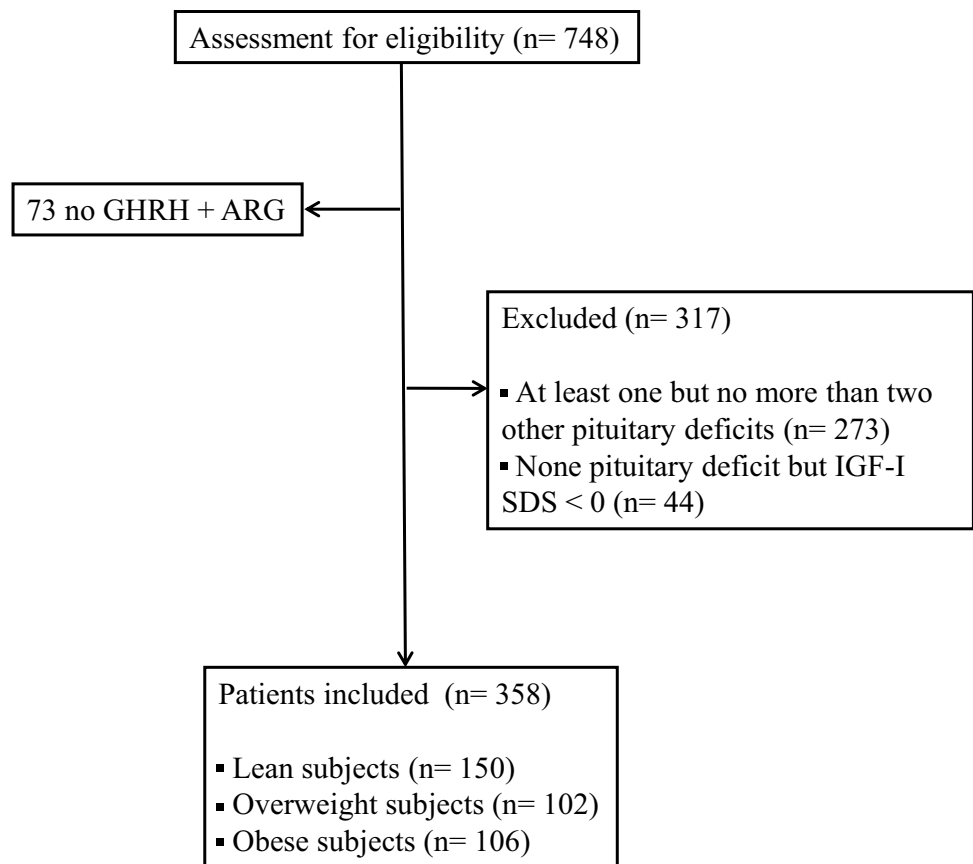
Patient selection and data collection

We retrospectively analyzed the data of 748 patients with history of pituitary disease, referred to the Neuroendocrinology Clinic of our Center from 01.01.2016 to 31.12.2018 for the evaluation of GH secretion. The patients' pituitary function was considered as the gold standard for the diagnosis

or exclusion of GHD; in particular GHD was defined by the presence of at least 3 other pituitary deficits, while a preserved somatotrophic function was defined by the lack of other pituitary deficits with an IGF-I standard deviation score (SDS) ≥ 0 . Seventy-three patients never performed a GHRH + ARG test and were therefore excluded. Two hundred and seventy-three patients were excluded because they presented at least one but no more than two other pituitary deficits except the assumed GHD; 44 patients were excluded because, even if they did not have any pituitary deficit, they presented an IGF-I SDS < 0 . A consort flow diagram of the study population is outlined in Fig. 1. Therefore, the final study population consisted in 358 patients (218 M, 140 F; age [mean \pm SD]: 45.1 ± 17.2 years; BMI: 27.2 ± 5.8 kg/m²), subdivided in relation to BMI in 150 normal weight, 102 overweight, and 106 obese subjects. Patients' clinical characteristics are reported in Table 1. We studied the GH response to GHRH (1 μ g/kg i.v. at 0 min) + ARG (0.5 g/kg infused i.v. from 0 to +30 min) (sampling every 15 min from +30 to +60 min) as previously described [24]. IGF-I levels were measured in all patients. All subjects gave their informed consent to the processing of their data. The study was approved by the local Ethics Committee and was in accordance with the principles of the Declaration of Helsinki.

Analytical methods

Serum GH levels (μ g/l) were measured in duplicate by IRMA method (IRMA GH, Beckman Coulter, Czech Republic). The IRMA assay of GH is a sandwich-type assay. The kit utilizes mouse monoclonal antibodies directed against two different epitopes of the molecule. The antibodies recognize the 22 kDa monomer, the dimer and GH bound to its binding protein. The calibrators are balanced out on the international standard WHO 2nd IS 98/574 in human serum. The sensitivity of the assay was 0.033 μ g/l. The inter- and intra-assay coefficients of variation (CV) were 9.0–14.0% and 2.4–6.5%, respectively. Serum IGF-I levels (μ g/l) were measured in duplicate by RIA method (SM-C-RIA-CT, DIAsource Immuno Assays, Belgium) after acid-ethanol extraction to avoid interference by binding proteins. The sensitivity of the method was 0.25 μ g/l. The inter- and intra-assay CV were 6.8–14.9% and 4.5–7.0%, respectively. IGF-I levels are expressed both as an absolute value and as a SDS of the mean normal value. The SDS for each subject was calculated in accordance with the published normality data on a population of 547 healthy Italian subjects [31]. All samples from an individual subject were analysed together.

Fig. 1 Consort diagram of the study population

Statistical analysis

Statistical analysis was performed using STATA 17 (Stata-Corp, College Station, Texas, USA).

Baseline patients' characteristics were summarized using mean and standard deviation (SD) for continuous data and percent values for categorical data. Between-group differences were evaluated by Student t test or by Chi-squared test, as appropriate.

We tried to identify the best GH cut-off to GHRH + ARG for the diagnosis of GHD using the Receiver-Operating Characteristic (ROC) curve analysis. The best cut-off in a ROC curve is the closest point to that with the theoretical maximum of both sensibility (SE) and specificity (SP). In order to avoid over-diagnosis of GHD we subsequently searched for the cut-offs with a SP \geq 95%. The diagnostic cut-off points were calculated for the lean (BMI < 25 kg/m²), overweight (BMI 25–29.9 kg/m²), and obese (BMI \geq 30 kg/m²) groups. For each identified cut-off we also calculated the positive predictive value (PPV), the negative predictive value (NPV), the positive likelihood ratio (LHR+), and the negative likelihood ratio (LHR-). A LHR+ greater than 1 indicates that a positive result to the test is associated with the disease. A LHR- less than 1 indicates that a negative result to the test is associated with absence of the disease.

Tests where the likelihood ratios lie close to 1 have little practical significance, as the post-test probability (odds) is little different from the pre-test probability, and as such is used primarily for diagnostic purposes, and not screening purposes. When the LHR+ is greater than 5 or the LHR- is less than 0.2 (i.e. 1/5), then they can be applied to the pre-test probability of a patient having the disease tested for to estimate a post-test probability of the disease state existing [32]. The accuracy of the identified cut-offs was defined as the probability of a person who has the disease and of a person who does not has the disease testing positive and negative, respectively.

Results

Based on patients' pituitary function 223 patients were defined GHD (142 M, 81 F) and 135 patients were defined as not affected by GHD (noGHD; 76 M, 59 F). The two groups differed both for age (GHD vs noGHD: 47.5 \pm 16.2 vs 41.0 \pm 18.1 yrs, $p < 0.001$) and BMI (GHD vs noGHD: 27.8 \pm 5.6 vs 26.3 \pm 6.0 kg/m², $p = 0.01$) (Table 1). An hypothalamic-pituitary tumor was more frequent in GHD than in noGHD patients, while a previous diagnosis of idiopathic childhood onset GHD and a history of traumatic brain injury

Table 1 Clinical features of both the whole cohort of 358 patients with history of hypothalamic-pituitary disease and subdivided according to the presence (GHD) or absence (noGHD) of adult growth hormone deficiency defined by a clinical point of view

	All study population (n=358)	GHD (n=223)	noGHD (n=135)	p
Gender				
Male subjects (%)	218 (60.9)	142 (63.7)	76 (56.3)	0.20
Female subjects (%)	140 (39.1)	81 (36.3)	59 (43.7)	
Age (mean±SD) yrs	45.1 ± 17.2	47.5 ± 16.2	41.0 ± 18.1	<0.001
BMI (mean±SD) kg/m²	27.2 ± 5.8	27.8 ± 5.6	26.3 ± 6.0	0.01
Lean subjects (%)	150 (41.9)	80 (35.9)	70 (51.8)	0.02
Overweight subjects (%)	102 (28.5)	73 (32.7)*	29 (21.5)*	
Obese subjects (%)	106 (29.6)	70 (31.4)*	36 (26.7)*	
Peak GH to GHRH + ARG test (mean±SD) µg/l	13.8 ± 24.4	2.1 ± 3.0	34.8 ± 36.1	< 0.001
IGF-I (mean±SD) µg/l	123.1 ± 91.0	79.2 ± 42.6	240.0 ± 83.5	< 0.001
IGF-I SDS	- 0.7 ± 1.3	- 1.3 ± 0.79	0.92 ± 0.69	< 0.001
Pathology (%)				
Pituitary adenomas	157 (43.9)	121 (54.3)*	36 (26.7)*	
Craniopharingioma + Rathke's cleft cyst	33 (9.2)	32 (14.3)*	1 (0.7)*	
Primary + secondary (due to Sheehan syndrome) empty sella	69 (19.3)	51 (22.9)*	18 (13.3)*	0.04
Idiopathic CO-GHD	17 (4.7)	0 (0)	17 (12.6)	
Traumatic brain injury + ESA	54 (15.1)	6 (2.7)*	48 (35.6)*	
Other	28 (7.8)	13 (5.8)	15 (11.1)	
Treatment of the pituitary pathology (%)				
Neurosurgery alone	155 (43.3)	105 (47.1)*	50 (37.0)*	< 0.001
Radiotherapy alone	6 (1.7)	3 (1.3)	3 (2.3)	
Neurosurgery + radiotherapy	56 (15.6)	51 (22.9)*	5 (3.7)*	
None	141 (39.4)	64 (28.7)	77 (57.0)	

*p < 0.001 GHD vs noGHD

or subarachnoid haemorrhage were more frequent in noGHD than in GHD patients (Table 1). Previous neurosurgical treatment alone or associated with radiotherapy was more frequent in GHD patients (Table 1), while a “wait and see” approach was more common in noGHD (Table 1).

IGF-I levels in noGHD were higher than in GHD when expressed both as mean and SDS levels (240.0 ± 83.5 µg/l vs 79.2 ± 42.6 µg/l, $p < 0.001$; 0.92 ± 0.69 vs $- 1.3 \pm 0.79$, $p < 0.001$) (Table 1), with a clear overlap in the two groups both in IGF-I levels (IGF-I range in noGHD vs GHD: 118–539 µg/l vs 8–268 µg/l) and IGF-I SDS values (IGF-I SDS range in noGHD vs GHD: 0.01–2.98 vs $- 2.71$ –1.85).

The mean GH response to GHRH + ARG in noGHD patients was higher than that recorded in GHD (34.8 ± 36.1 µg/l vs 2.1 ± 3.0 µg/l, $p < 0.001$) (Table 1).

The best GH cut-offs to GHRH + ARG were: (i) 8.0 µg/l in lean subjects with a SE value of 95% and a SP value of 100%; (ii) 7.0 µg/l in overweight subjects with a SE value of 97.3% and a SP value of 82.8%; (iii) 2.8 µg/l in obese subjects with a SE value of 84.3% and a SP value of 91.7%, respectively (Fig. 2, Table 2).

The cut-offs that identify with a SP $\geq 95\%$ those patients who should not undergo rhGH substitution therapy, because more probably not affected by GHD were: (i) 2.6 µg/l in overweight subjects with a SE and SP value of 68.5% and 96.6%, respectively; (ii) 1.75 µg/l in obese subjects with a SE and SP value of 70.0% and 97.2%, respectively (Fig. 2, Table 2). For lean subjects the cut-off identified at the first evaluation (i.e., 8 µg/l) was already characterized by a SP $\geq 95\%$.

The PPV, NPV, the ROC AUC, LHR +, LHR-, and the diagnostic accuracy of these cut-off limits to GHRH + ARG are reported in Table 2.

The current GHRH + ARG cut-offs (i.e., ≤ 11.5 µg/l in lean, ≤ 8.0 µg/l in overweight, and ≤ 4.2 µg/l in obese subjects) applied to our study population clearly overestimated the GHD diagnosis (Table 2).

Side effects

Administration of GHRH + ARG did not cause any relevant side effect except for face flushing observed in 48 patients.

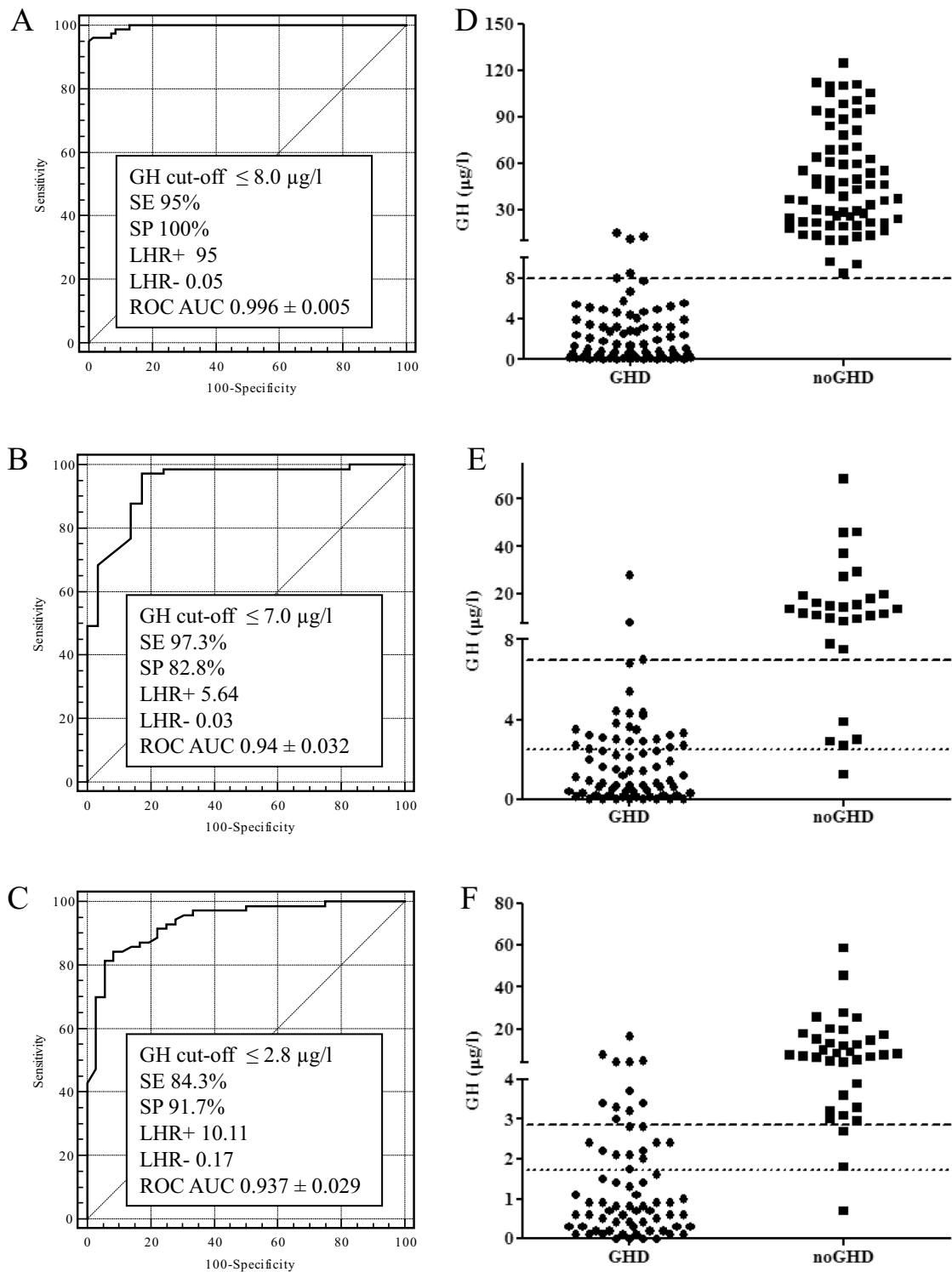


Fig. 2 ROC analysis to identify the best GH cut-off to GHRH+ARG in lean (A), overweight (B) and obese (C) patients. Individual value of peak GH to GHRH+ARG compared both to the GH cut-off with the best SE and the best SP (dashed line) and to the GH cut-off with

a $SP \geq 95\%$ (dotted line) according to ROC analysis in lean (D), overweight (E) and obese (F) patients (SE sensitivity, SP specificity, LHR+ positive likelihood ratio, LHR- negative likelihood ratio, ROC AUC area under the ROC curve)

Table 2 Sensibility (SE), specificity (SP), positive predictive value (PPV), negative predictive value (NPV), area under the ROC curve (ROC AUC) \pm standard deviation (SD), positive likelihood ratio (LHR+), negative likelihood ratio (LHR-), and the diagnostic accu-

racy of the cut-off limits to GHRH + ARG identified with the ROC analysis in lean, overweight and obese patients and compared with the currently recognized cut-offs

References	BMI (kg/m ²)	Cut-off (μ g/l)	SE (%)	SP (%)	PPV (%)	NPV (%)	LHR+	LHR-	ROC AUC (mean \pm SD)	Diagnostic accuracy (%)
This study: best SE and SP and SP \geq 95%	< 25	\leq 8.0	95.0	100	100	94.6	95	0.05	0.996 \pm 0.005	97.3
[24]		\leq 11.5	97.5	92.9	94.0	97.0	13.73	0.03		95.3
This study: best SE and SP	25-29.9	\leq 7.0	97.3	82.8	93.4	92.3	5.64	0.03	0.940 \pm 0.032	93.1
This study: SP \geq 95%		\leq 2.6	68.5	96.6	98.0	54.9	19.86	0.33		76.5
[24]		\leq 8.0	97.3	75.9	91.0	91.7	4.04	0.04		91.2
This study: best SE and SP	\geq 30	\leq 2.8	84.3	91.7	95.2	75.0	10.11	0.17	0.937 \pm 0.029	86.8
This study: SP \geq 95%		\leq 1.75	70.0	97.2	98.0	62.5	25.20	0.31		79.2
[24]		\leq 4.2	95.7	69.4	85.9	89.3	3.13	0.06		86.8

Discussion

To our knowledge this is the first study that identifies the diagnostic cut-offs of GH response to GHRH + ARG related to BMI using a clinical definition of GHD as gold standard.

Our results, based on a considerable group of patients with a history of hypothalamic–pituitary disease, suggest that, with this new approach, the GHRH + ARG cut-offs should be revised. Indeed, our data specifically show that:

- The best cut-off to GHRH + ARG test for the diagnosis of adult GHD is 8.0 μ g/l (SE 95%, SP 100%) in lean, 7.0 μ g/l in overweight (SE 97.3%, SP 82.8%) and 2.8 μ g/l (SE 84.3%, SP 91.7%) in obese subjects.
- However, in order to minimize the bias in GHD diagnosis, by reducing the false positive value (i.e. SP \geq 95%), the cut-off should be further reduced to 2.6 μ g/l (SE 68.5%, SP 96.6%) in overweight, and to 1.75 μ g/l (SE 70.0%, SP 97.2%) in obese subjects.
- The current cut-offs to GHRH + ARG test for the diagnosis of adult GHD overestimate the diagnosis of GHD in either lean, overweight, and obese subjects.

The diagnosis of adult GHD is challenging for the clinician because of the lack of a single biological endpoint. Adult GHD diagnosis depends on the demonstration of a reduced peak serum GH level in response to one or more GH stimulation tests [2, 10, 11]. Testing for adult GHD should only be considered if there is a clinical suspicion of GHD and the intention is to treat if the diagnosis is confirmed [2, 10, 11]. Currently, there is no ideal stimulation test and the decision to consider performing a GH stimulation test to diagnose adult GHD must take into account the validity of the chosen test, its GH cut-offs, the availability of local resources, and the clinician expertise.

All GH stimulation tests are based on the concept that a pharmacological agent stimulates pituitary GH secretion, with peak GH levels detectable by timed frequent serum sampling after administration of the stimulus.

Initially, the diagnostic cut-offs for different stimulus tests were identified by looking for the minimal response observed in a group of normal subjects [33–35]; subsequently new studies have revisited the diagnostic criteria comparing the response to the test in analysis with that of the test considered the gold standard [21, 25–29], often comparing patients with normal subjects [22–24, 27, 28].

All these approaches are exposed to possible bias. In particular, the studies that compare the response of a group of patients already diagnosed for the target condition with the response observed in healthy volunteers, even if matched by sex, age and BMI, are at risk of overestimating the diagnostic accuracy of the cut-offs thus identified and, consequently, the results obtained cannot be applied to the clinical setting. On the other hand, the studies that establish the diagnostic accuracy of a particular test, comparing it with another diagnostic test, are at risk of overestimating or underestimating the diagnosis similarly to what the gold standard itself does.

Taking this into account, our work is the first that tried to identify the best cut-offs to GHRH + ARG test for the diagnosis of adult GHD using only a large group of patients, which is the most numerous reported to date in the literature; furthermore, in order to avoid the bias previously mentioned, we used a clinical criteria as gold standard; in particular we considered the presence of at least 3 other pituitary deficits as equivalent to the diagnosis of GHD and the absence of other hormonal deficits, together with an IGF-I SDS \geq 0, as equivalent to the exclusion of GHD. Indeed, as reported in literature and in current guidelines [2, 36–38], the presence/absence of other pituitary deficits and the IGF-I SDS values are the two most important parameters for predicting a final

diagnosis of GHD. These features provide valuable information that can be used as a proxy for estimating the probability of GHD prior to stimulation tests.

The correct choice of the clinical gold standard seems to be confirmed by the clinical differences between GHD and noGHD observed at baseline: it is not surprising that hypopituitary patients were those who more frequently presented a morpho-structural alteration of the hypothalamic-pituitary region or who more frequently had undergone neurosurgical and/or radiotherapy treatment [39–42]. Moreover, the higher prevalence of overweight and obese subjects in GHD group is in line with the finding that hypopituitary patients are usually characterized by an alteration in body composition similar to that of the metabolic syndrome [43]. Finally, our work confirms that the GH response to GHRH + ARG is significantly more compromised in patients with three or more pituitary hormone deficiencies than in patients without deficiencies [37].

Our work highlights one more time the well-known negative association between the GH response to stimulation tests and the BMI [21, 22, 24, 26, 27, 44] with a clear decline in the identified cut-off in lean, overweight and obese subjects. The distinct and progressive reduction of GH response to any stimulation test has clinical impact when these tests are used to demonstrate the presence of GHD [45]. Indeed, one of the warnings in interpreting the results of all the GH stimulation tests in adults is the high prevalence of obesity in the general population, as well as in patients with hypothalamic-pituitary disease [19]. Even if, to date, GHRH + ARG has already well established cut-off BMI related for the diagnosis of adult GHD, it should be considered that the previous study has evaluated the GH response to GHRH + ARG by comparing the response of patients who are certainly GHD with normal subjects [24].

As shown in Table 2, the new GH cut-off limits to GHRH + ARG, defined as the one with the best SE and SP, are characterized by a diagnostic accuracy that is better than the one of the previous cut-offs; in particular, the identified cut-offs generally improve SP maintaining a near comparable SE in lean and overweight subjects, while SE declines in obese subjects. Moreover, we decided to maximize SP to value $\geq 95\%$ in the three BMI category groups, in order to correctly identify true GHD subjects, avoiding overtreatment in patients not actually GHD, particularly in obese population. With this new approach the cut-offs thus identified obviously reduce the SE values but are characterized by a marked increase in LHR + values ranging between 19.86 in overweight and 95 in lean subjects. The reduced diagnostic accuracy observed with this approach is clearly due to the reduction in the ability to identify some true GHD patients, but, on the other hand we can observe an improvement in the ability to rightly identify subjects that do not need GH replacement therapy.

We are also aware that the definition of a preserved somatotrophic function used in this study (no other pituitary deficits and IGF-I SDS ≥ 0) results in a possible underestimation of isolated GHD cases. However, a previous study [46] has demonstrated that the absence of other pituitary deficiencies, along with IGF-I SDS levels > -0.52 indicates a low pretest probability of GHD; in the current study, we aimed to further emphasize this concept by rising the threshold of IGF-I SDS to values ≥ 0 . This adjustment was made to increase confidence in the clinical definition of no-GHD, ensuring that subjects identified as such truly have a low pretest probability of GHD.

However, considering the cost of rhGH replacement therapy, the potential side effects, and the lack of reliable data on strong end points, we believe that for GHD it is preferable to be sure to exclude from the therapy patients not affected by GHD, rather than not to treat patients with GHD. Considering this, the results of this study allow us to suggest the use of the new cut-offs with the best compromise between SE and SP (i.e. 8.0 $\mu\text{g/l}$, 7.0 $\mu\text{g/l}$, and 2.8 $\mu\text{g/l}$, in lean, overweight and obese subjects, respectively) in patients with a high pretest probability to be GHD, and the use of new cut-offs with SP $\geq 95\%$ (i.e. 8.0 $\mu\text{g/l}$, 2.6 $\mu\text{g/l}$, and 1.75 $\mu\text{g/l}$, in lean, overweight and obese subjects, respectively) in patients with a low pretest probability.

Finally, it must be underlined that the cut-offs identified are acceptable for the assay used in this study; however, it must be emphasized that our study makes it possible to carry out a Bland–Altman plot to identify any correction factors for any other assay.

In conclusion, the results of this study, based on a considerable group of patients with a history of hypothalamic-pituitary disease and on a clinical gold standard, suggest that the GH cut-offs to GHRH + ARG test should be revised in all BMI thresholds, in order to avoid falsely positive diagnoses of severe GHD in adults.

It will need to be evaluated in future studies whether the lack of effectiveness of rhGH therapy on strong endpoints, such as increased mortality, could be partially attributed to the incorrect prescription of therapy to patients who have been identified as having GHD based on current cut-offs, but who do not meet the criteria for GHD according to the new cut-offs proposed by us.

Acknowledgements The authors thank Mrs. Laura Saba and all her nursing staff for their skilful technical assistance provided during the tests. The results of this study were already partially presented as abstracts at the ECE congress held in Milan in May 2022 and the IEM congress held in Naples in September 2022.

Author contributions VG, DC, EV, FB, AMB, CB and NP: collected the anamnestic as well as the biochemical data for each patient; VG performed the data analysis, the figures designing and the manuscript writing; EG, MM and SG. verified the analytical methods and

supervised the manuscript drafting. All authors discussed the results and contributed to the final manuscript.

Funding Open access funding provided by Università degli Studi di Torino within the CRUI-CARE Agreement. This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.

Data availability The data sets generated during and/or analyzed during the current study are not publicly available but are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors have no conflict of interest to declare. Prof. Silvia Grottoli and Dr. Nunzia Prencipe are among the Members of the Editorial Board of the Journal of Endocrinological Investigation.

Ethical approval The protocol was in accordance with the principles of the Declaration of Helsinki, approved by the local Ethical Committee and all the subjects gave their informed consent to the collection of their data.

Research involving human participants and/or animals The study was approved by the local Ethics Committee and was in accordance with the principles of the Declaration of Helsinki.

Informed consent All subjects gave their informed consent to the processing of their data.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Gasco V, Cambria V, Bioletto F et al (2021) Traumatic brain injury as frequent cause of hypopituitarism and growth hormone deficiency: epidemiology, diagnosis, and treatment. *Front Endocrinol*. <https://doi.org/10.3389/fendo.2021.634415>
2. Molitch ME, Clemmons DR, Malozowski S et al (2011) Evaluation and treatment of adult growth hormone deficiency: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 96:1587–1609. <https://doi.org/10.1210/jc.2011-0179>
3. Bioletto F, Prencipe N, Berton AM et al (2022) MRI assessment of cardiac function and morphology in adult patients with growth hormone deficiency: a systematic review and meta-analysis. *Front Endocrinol*. <https://doi.org/10.3389/fendo.2022.910575>
4. Burman P, Mattsson AF, Johannsson G et al (2013) Deaths among adult patients with hypopituitarism: hypocortisolism during acute stress, and de novo malignant brain tumors contribute to an increased mortality. *J Clin Endocrinol Metab* 98:1466–1475. <https://doi.org/10.1210/jc.2012-4059>
5. Simpson H, Savine R, Sönksen P et al (2002) Growth hormone replacement therapy for adults: into the new millennium. *Growth Horm IGF Res* 12:1–33. <https://doi.org/10.1054/ghir.2001.0263>
6. Svensson J, Bengtsson B-A, Rosén T et al (2004) Malignant disease and cardiovascular morbidity in hypopituitary adults with or without growth hormone replacement therapy. *J Clin Endocrinol Metab* 89:3306–3312. <https://doi.org/10.1210/jc.2003-031601>
7. Widdowson WM, Gibney J (2008) The effect of growth hormone replacement on exercise capacity in patients with GH deficiency: a metaanalysis. *J Clin Endocrinol Metab* 93:4413–4417. <https://doi.org/10.1210/jc.2008-1239>
8. Jørgensen AP, Fougner KJ, Ueland T et al (2011) Favorable long-term effects of growth hormone replacement therapy on quality of life, bone metabolism, body composition and lipid levels in patients with adult-onset growth hormone deficiency. *Growth Horm IGF Res* 21:69–75. <https://doi.org/10.1016/j.ghir.2011.01.001>
9. Vance ML (2003) Can growth hormone prevent aging? *N Engl J Med* 348:779–780. <https://doi.org/10.1056/NEJMp020186>
10. Ho KKY (2007) GH Deficiency Consensus Workshop Participants (2007) Consensus guidelines for the diagnosis and treatment of adults with GH deficiency II: a statement of the GH Research Society in association with the European Society for Pediatric Endocrinology, Lawson Wilkins Society, European Society of Endocrinology, Japan Endocrine Society, and Endocrine Society of Australia. *Eur J Endocrinol* 157:695–700. <https://doi.org/10.1530/EJE-07-0631>
11. Yuen KCJ, Biller BMK, Radovick S et al (2019) American association of clinical endocrinologists and american college of endocrinology guidelines for management of growth hormone deficiency in adults and patients transitioning from pediatric to adult care. *EndocrPract* 25:1191–1232. <https://doi.org/10.4158/GL-2019-0405>
12. Veldhuis JD, Iranmanesh A, Ho KK et al (1991) Dual defects in pulsatile growth hormone secretion and clearance subserve the hypsomatotropism of obesity in man. *J Clin Endocrinol Metab* 72:51–59. <https://doi.org/10.1210/jcem-72-1-51>
13. Iranmanesh A, Lizarralde G, Veldhuis JD (1991) Age and relative adiposity are specific negative determinants of the frequency and amplitude of growth hormone (GH) secretory bursts and the half-life of endogenous GH in healthy men. *J Clin Endocrinol Metab* 73:1081–1088. <https://doi.org/10.1210/jcem-73-5-1081>
14. Van Dam EWCM, Roelfsema F, Helmerhorst FH et al (2002) Low amplitude and disorderly spontaneous growth hormone release in obese women with or without polycystic ovary syndrome. *J Clin Endocrinol Metab* 87:4225–4230. <https://doi.org/10.1210/jc.2002-012006>
15. Maccario M, Valetto MR, Savio P et al (1997) Maximal secretory capacity of somatotrope cells in obesity: comparison with GH deficiency. *Int J ObesRelatMetabDisord* 21:27–32. <https://doi.org/10.1038/sj.ijo.0800356>
16. Juul A, Jorgensen JOL (2000) Growth hormone in adults: physiological and clinical aspects. Cambridge University Press, Cambridge
17. Vizner B, Reiner Z, Sekso M (1983) Effect of l-dopa on growth hormone, glucose, insulin, and cortisol response in obese subjects. *Exp Clin Endocrinol* 81:41–48. <https://doi.org/10.1055/s-0029-1210205>
18. Cordido F, Dieguez C, Casanueva FF (1990) Effect of central cholinergic neurotransmission enhancement by pyridostigmine on the growth hormone secretion elicited by clonidine, arginine, or hypoglycemia in normal and obese subjects. *J Clin Endocrinol Metab* 70:1361–1370. <https://doi.org/10.1210/jcem-70-5-1361>

19. Verhelst J, Abs R (2009) Cardiovascular risk factors in hypopituitary GH-deficient adults. *Eur J Endocrinol* 161(Suppl 1):S41–49. <https://doi.org/10.1530/EJE-09-0291>
20. Gordon MB, Levy RA, Gut R, Germak J (2016) Trends in growth hormone stimulation testing and growth hormone dosing in adult growth hormone deficiency patients: results from the answer program. *Endocr Pract* 22:396–405. <https://doi.org/10.4158/EP14603.0R>
21. Gasco V, Ferrero A, Bisceglia A et al (2021) The cut-off limits of growth hormone response to the insulin tolerance test related to body mass index for the diagnosis of adult growth hormone deficiency. *Neuroendocrinology* 111:442–450. <https://doi.org/10.1159/000508103>
22. Kelestimur F, Popovic V, Leal A et al (2006) Effect of obesity and morbid obesity on the growth hormone (GH) secretion elicited by the combined GHRH + GHRP-6 test. *Clin Endocrinol (Oxf)* 64:667–671. <https://doi.org/10.1111/j.1365-2265.2006.02525.x>
23. Dichtel LE, Yuen KCJ, Bredella MA et al (2014) Overweight/Obese adults with pituitary disorders require lower peak growth hormone cutoff values on glucagon stimulation testing to avoid overdiagnosis of growth hormone deficiency. *J Clin Endocrinol Metab* 99:4712–4719. <https://doi.org/10.1210/jc.2014-2830>
24. Corneli G, Di Somma C, Baldelli R et al (2005) The cut-off limits of the GH response to GH-releasing hormone-arginine test related to body mass index. *Eur J Endocrinol* 153:257–264. <https://doi.org/10.1530/eje.1.01967>
25. Garcia JM, Biller BMK, Korbonits M et al (2018) Macimorelin as a diagnostic test for adult gh deficiency. *J Clin Endocrinol Metab* 103:3083–3093. <https://doi.org/10.1210/jc.2018-00665>
26. Garcia JM, Swerdloff R, Wang C et al (2013) Macimorelin (AEZS-130)-stimulated growth hormone (GH) test: validation of a novel oral stimulation test for the diagnosis of adult GH deficiency. *J Clin Endocrinol Metab* 98:2422–2429. <https://doi.org/10.1210/jc.2013-1157>
27. Gómez JM, Espadero RM, Escobar-Jiménez F et al (2002) Growth hormone release after glucagon as a reliable test of growth hormone assessment in adults. *Clin Endocrinol* 56:329–334. <https://doi.org/10.1046/j.1365-2265.2002.01472.x>
28. Popovic V, Leal A, Micic D et al (2000) GH-releasing hormone and GH-releasing peptide-6 for diagnostic testing in GH-deficient adults. *Lancet* 356:1137–1142. [https://doi.org/10.1016/S0140-6736\(00\)02755-0](https://doi.org/10.1016/S0140-6736(00)02755-0)
29. Gasco V, Beccuti G, Baldini C et al (2013) Acylated ghrelin as a provocative test for the diagnosis of GH deficiency in adults. *Eur J Endocrinol* 168:23–30. <https://doi.org/10.1530/EJE-12-0584>
30. Whiting P, Rutjes AWS, Reitsma JB et al (2004) Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med* 140:189–202. <https://doi.org/10.7326/0003-4819-140-3-200402030-00010>
31. Aimaretti G, Boschetti M, Corneli G et al (2008) Normal age-dependent values of serum insulin growth factor-I: results from a healthy Italian population. *J Endocrinol Invest* 31:445–449. <https://doi.org/10.1007/BF03346389>
32. Beardsell I, Bell S, Robinson S et al (2009) Mcem part a: Mcqs. *R Soc Med Press* 2009:460
33. Aimaretti G, Baffoni C, DiVito L et al (2000) Comparisons among old and new provocative tests of GH secretion in 178 normal adults. *Eur J Endocrinol* 142:347–352. <https://doi.org/10.1530/eje.0.1420347>
34. Hoffman DM, O’Sullivan AJ, Baxter RC, Ho KK (1994) Diagnosis of growth-hormone deficiency in adults. *Lancet* 343:1064–1068. [https://doi.org/10.1016/s0140-6736\(94\)90181-3](https://doi.org/10.1016/s0140-6736(94)90181-3)
35. Ghigo E, Aimaretti G, Gianotti L et al (1996) New approach to the diagnosis of growth hormone deficiency in adults. *Eur J Endocrinol* 134:352–356. <https://doi.org/10.1530/eje.0.1340352>
36. Toogood AA, Beardwell CG, Shalet SM (1994) The severity of growth hormone deficiency in adults with pituitary disease is related to the degree of hypopituitarism. *Clin Endocrinol* 41:511–516. <https://doi.org/10.1111/j.1365-2265.1994.tb02583.x>
37. Hartman ML, Crowe BJ, Biller BMK et al (2002) Which patients do not require a GH stimulation test for the diagnosis of adult GH deficiency? *J Clin Endocrinol Metab* 87:477–485. <https://doi.org/10.1210/jcem.87.2.8216>
38. Juul A, Kastrup KW, Pedersen SA, Skakkebaek NE (1997) Growth hormone (GH) provocative retesting of 108 young adults with childhood-onset GH deficiency and the diagnostic value of insulin-like growth factor I (IGF-I) and IGF-binding protein-3. *J Clin Endocrinol Metab* 82:1195–1201. <https://doi.org/10.1210/jcem.82.4.3892>
39. Appelman-Dijkstra NM, Kokshoorn NE, Dekkers OM et al (2011) Pituitary dysfunction in adult patients after cranial radiotherapy: systematic review and meta-analysis. *J Clin Endocrinol Metab* 96:2330–2340. <https://doi.org/10.1210/jc.2011-0306>
40. Bioletto F, Berton AM, Varaldo E et al (2022) Development and internal validation of a predictive score for the diagnosis of central adrenal insufficiency when morning cortisol is in the grey zone. *J Endocrinol Invest*. <https://doi.org/10.1007/s40618-022-01926-z>
41. Schneider HJ, Aimaretti G, Kreitschmann-Andermahr I et al (2007) Hypopituitarism. *Lancet* 369:1461–1470. [https://doi.org/10.1016/S0140-6736\(07\)60673-4](https://doi.org/10.1016/S0140-6736(07)60673-4)
42. Higham CE, Johannsson G, Shalet SM (2016) Hypopituitarism. *Lancet* 388:2403–2415. [https://doi.org/10.1016/S0140-6736\(16\)30053-8](https://doi.org/10.1016/S0140-6736(16)30053-8)
43. Verhelst J, Mattsson AF, Luger A et al (2011) Prevalence and characteristics of the metabolic syndrome in 2479 hypopituitary patients with adult-onset GH deficiency before GH replacement: a KIMS analysis. *Eur J Endocrinol* 165:881–889. <https://doi.org/10.1530/EJE-11-0599>
44. Qu X-D, Gaw Gonzalo IT, Al Sayed MY et al (2005) Influence of body mass index and gender on growth hormone (GH) responses to GH-releasing hormone plus arginine and insulin tolerance tests. *J Clin Endocrinol Metab* 90:1563–1569. <https://doi.org/10.1210/jc.2004-1450>
45. Bonert VS, Elashoff JD, Barnett P, Melmed S (2004) Body mass index determines evoked growth hormone (GH) responsiveness in normal healthy male subjects: diagnostic caveat for adult GH deficiency. *J Clin Endocrinol Metab* 89:3397–3401. <https://doi.org/10.1210/jc.2003-032213>
46. Bioletto F, Parasiliti-Caprino M, Berton AM et al (2021) Development and internal validation of a predictive model for adult GH deficiency prior to stimulation tests. *Front Endocrinol* 12:737947. <https://doi.org/10.3389/fendo.2021.737947>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.