

Limited usefulness of the test of spontaneous growth hormone (GH) nocturnal secretion as a screening procedure in diagnosing GH deficiency in children with short stature

Joanna Smyczyńska¹, Renata Stawerska¹, Andrzej Lewiński^{1,2}, Maciej Hilczer^{1,3}

¹ Department of Endocrinology and Metabolic Diseases, Polish Mother's Memorial Hospital – Research Institute, Lodz, Poland

² Department of Endocrinology and Metabolic Diseases, Medical University of Lodz, Lodz, Poland

³ Department of Pediatric Endocrinology, Medical University of Lodz, Lodz, Poland

Smyczyńska J, Stawerska R, Lewiński A, Hilczer M. Limited usefulness of the test of spontaneous growth hormone (GH) nocturnal secretion as a screening procedure in diagnosing GH deficiency in children with short stature. *Ann Agric Environ Med.* 2014; 21(4): 893–897. doi: 10.5604/12321966.1129954

Abstract

Introduction and objective. In Poland, the assessment of nocturnal GH secretion has gained the status of screening test; however, this procedure is not included in international recommendations. The aim of the study was to assess the accuracy and predictive value of the test of nocturnal GH secretion as a screening procedure in diagnosing GHD, and to check the adequacy of the cut-off value for GH peak in this test on the level of 10 ng/ml.

Materials and methods. The analysis comprised the data of 1,000 children with short stature. In all the patients, GH secretion was assessed in a screening test (after falling asleep) and in 2 stimulating tests (reference tests), with simultaneous assessment of IGF-I secretion before stimulating tests. The indices of screening test accuracy, likelihood ratios and predictive values were assessed. The cut-off level of GH peak after falling asleep, ensuring its 95% sensitivity, was calculated in ROC curve analysis.

Results. Sensitivity of the screening test was 70.4%, while the specificity – 61.2%, positive likelihood ratio – 1.842, negative likelihood ratio – 0.482, positive predictive value – 0.462, negative predictive value – 0.812. The sensitivity of the test of GH secretion after falling asleep is too low with respect to the requirements for screening test. The ROC curve analysis showed 95% sensitivity for the screening test on the level of 19.0 ng/ml; however, with a very low specificity – below 25%, thus making this test completely useless as a screening procedure.

Conclusions. The obtained results strongly contradict the opinion that the assessment of GH secretion after falling asleep should be a screening test in diagnosing GHD in children with short stature.

Key words

short stature, growth hormone deficiency, screening test, stimulating tests

INTRODUCTION

Growth hormone (GH) deficiency (GHD) is an important cause of short stature in children and a standard indication to the therapy with recombinant human GH (rhGH). Insulin-like growth factor (IGF-I) is a main mediator of GH action, especially in its growth-promoting effects. The classification of disorders of GH secretion and function has recently been updated, and GHD is considered as one of the forms of secondary IGF-I deficiency [1,2]. Thus, the significance of IGF-I assessment in diagnosing GHD has recently increased. However, according to previous classifications, the most important diagnostic step was confirmation of the decreased GH secretion, either in stimulating tests (classic form of GHD) or only of spontaneous secretion after falling asleep, despite normal results of GH stimulating tests (neurosecretory dysfunction – NSD).

In Poland, the first step in diagnosing GHD in children with short stature is an assessment of nocturnal GH secretion [3]. Since this test has gained the status of a screening

procedure, its normal result may be sufficient reason to reject a short child, not only from rhGH therapy, but also from further assessment of GH secretory status. The procedure for screening test in question involves the measurement of GH concentrations in samples obtained every 30 minutes during 2 hours of sleep, starting 1 hour after falling asleep.

Unfortunately, there is no concordance with respect to the significance of the assessment of nocturnal GH secretion between the requirements of the Polish Therapeutic Programme of rhGH therapy for GH-deficient children [3] and the statements of international experts [4,5]. According to the quoted recommendations, the assessment of nocturnal GH secretion should not be a standard diagnostic test and may be performed only for scientific purposes.

It seems very important to be sure that the results of all the above-mentioned procedures (i.e., an assessment of nocturnal GH secretion, the results of GH stimulating tests and measurement of IGF-I serum concentration) demonstrate reciprocal compatibility, especially, if any of them is to be recommended as a screening test.

The aim of the study was to assess the accuracy and predictive value of the test of GH secretion after falling asleep as a screening procedure in diagnosing GHD, and also to check the adequacy of the cut-off value for GH peak in that test at the level of 10 ng/ml.

Address for correspondence: Andrzej Lewiński, Department of Endocrinology and Metabolic Diseases, Medical University of Lodz, Polish Mother's Memorial Hospital – Research Institute, Rzgowska 281/289, 93-338 Lodz, Poland
E-mail: alewin@csk.umed.lodz.pl

Received: 15 May 2013; accepted: 14 June 2013

MATERIALS AND METHODS

The retrospective analysis included the data of 1,000 children (644 boys, 356 girls), aged 10.9 ± 3.4 years (mean \pm SD) with short stature, diagnosed in the Department of Endocrinology and Metabolic Diseases, Polish Mother's Memorial Hospital – Research Institute in Lodz, Poland, in the years 2005–2012. In all the patients, the screening test of GH secretion after falling asleep was performed according to the current requirements (GH concentration assessed in 5 blood samples collected every 30 minutes, starting 1 hour after falling asleep; the highest value from all samples was considered as the test result). Next, independently of the result of the screening test, the GH secretion was assessed in 2 standard stimulating tests (test with clonidine in the dose of 0.15 mg/m² orally, and with glucagon in the dose of 30 µg/kg m.c., *i.m.*, not exceeding 1 mg). The results of both stimulating tests were interpreted together on the ground of the highest GH concentration in both tests. The cut-off value for normal and decreased GH secretion for all the tests was set at the level of 10.0 ng/ml. Further, in all the patients, IGF-I serum concentration was measured in the blood samples collected before the first stimulating test in the morning hours. The results were expressed as IGF-I SDS for age and gender, according to the reference data provided by the Diagnostic Systems Laboratories. The result of this test was considered to be normal if IGF-I SDS was over -1.0, as suggested by Cianfarani et al. [6].

The children with chronic diseases, including other hormonal disorders, malnutrition or obesity, acquired GHD, genetic syndromes or other known health problems, were excluded from the study.

The concentration of GH was measured by the two-site chemiluminescent enzyme immunometric assay (hGH IMMULITE, DPC). Both IGF-I and IGF-BP-3 concentrations were assessed by a solid-phase, enzyme-labelled chemiluminescent immunometric assays, (IMMULITE, DPC).

The study was approved by the local Ethics Committee in Polish Mother's Memorial Hospital – Research Institute in Lodz, Poland.

RESULTS

Accuracy of the screening test. The assumption of the presented study was that assessment of GH peak in 2 stimulating tests is the reference procedure in diagnosing GHD. Thus, the accuracy of the screening test should be analyzed with respect to the results of stimulating tests. The results of screening test were classified as:

- true positive (TP): decreased GH peak in screening test, confirmed by stimulating tests, $n = 227$ (22.7%);
- false positive (FP): decreased GH peak in screening test, but normal in stimulating tests, $n = 262$ (26.2%);
- false negative (FN): normal GH peak in screening test while decreased in stimulating tests, $n = 96$ (9.6%);
- true negative (TN): normal GH peak in screening test and in stimulating tests, $n = 415$ (41.5%).

The following indices of screening test accuracy were calculated:

1. Sensitivity and specificity of the screening test

Sensitivity (SENS) – the proportion of TP results of the screening test to all the positive results of stimulating tests:

$$SENS = \frac{TP}{TP + FN} = \frac{227}{322} = 0.705 \text{ (70.5\%)}$$

Specificity (SPEC) – the proportion of TN results of the screening to all the negative results of stimulating tests:

$$SPEC = \frac{TN}{TN + FP} = \frac{415}{678} = 0.612 \text{ (61.2\%)}$$

2. Likelihood ratios of the screening test

Positive likelihood ratio (LR+) – proportion of the patients with positive result of both the reference and assessed tests (*i.e.*, those with decreased GH peak both in stimulating and screening tests) to the patients with negative result of reference test, despite the positive result of the assessed tests (*i.e.*, those with normal GH peak in stimulating tests, despite decreased GH peak in screening test), according to the formula:

$$(LR+) = \frac{SENS}{1 - SPEC} = \frac{0.705}{1 - 0.612} = 1.817$$

Negative likelihood ratio (LR-) – proportion of the patients with negative result of both the reference and assessed tests (*i.e.*, those with normal GH peak both in stimulating and screening tests) to the patients with negative result of reference test, despite the positive result of the assessed tests (*i.e.*, those with decreased GH peak in stimulating tests, despite normal GH peak in screening test), according to the formula:

$$(LR-) = \frac{1 - SENS}{SPEC} = \frac{1 - 0.705}{0.612} = 0.482$$

3. Predictive values of the screening test

Positive predictive value (PPV) – the proportion of TP results of screening test to all the positive results of that test:

$$PPV = \frac{TP}{TP + FP} = \frac{227}{489} = 0.462$$

Negative predictive value (NPV) – the proportion of TN results of screening test to all the negative results of that test:

$$NPV = \frac{TN}{TN + FN} = \frac{415}{511} = 0.812$$

Thus, the incidence of GHD, diagnosed on the basis of the results of stimulating tests, was 46.2% in the group of patients with positive result of screening test and 18.8% in those with negative result of the screening test (*i.e.*, among the patients in whom GHD should be theoretically excluded by the normal result of the screening test).

Adequacy of the cut-off value of GH peak in the screening test. The next part of the analysis was performed in order to

ensure that the cut-off value for GH peak in the screening test at the level of 10.0 ng/ml was optimal. It had to be assumed that sensitivity of the screening test in diagnosing the disease (*i.e.*, its sensitivity with reference to the results of reference tests) should be at least 95%, as is required for all the screening tests. The cut-off value of GH peak after falling asleep providing the 95% sensitivity of this test with respect to the results of GH stimulating tests, was estimated by the receiver operating characteristic (ROC) curve analysis.

According to the results of ROC analysis, the cut-off value for screening test of GH secretion after falling asleep that provided its 95% sensitivity with respect to the diagnosis of GHD – established on the basis of decreased GH peak in stimulating tests – should be established at the level of 19.1 ng/ml; however, such an increase in the cut-off level for screening test simultaneously decreased its specificity to 21.9% only. The ROC curve for the analysis is presented in Figure 1.

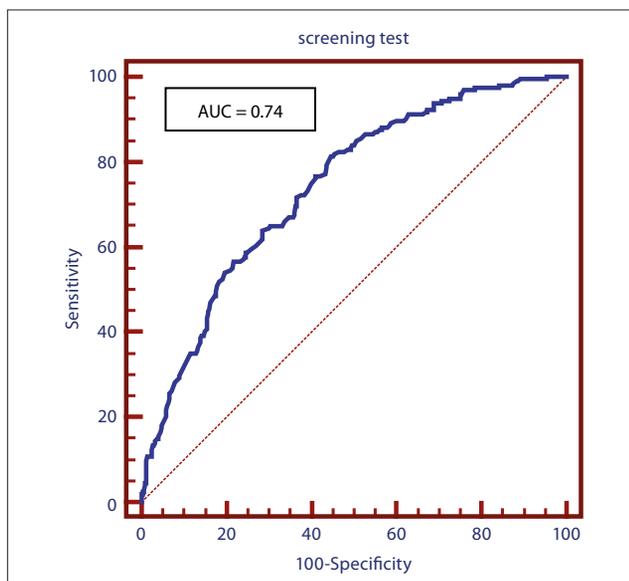


Figure 1. ROC curve for screening test with the cut-off GH peak at the level of 10.0 ng/ml

Taking into account the relatively high incidence of FP results of GH stimulating tests (*i.e.*, decreased GH peak in non GH-deficient child) [7], the diagnosis of GHD was considered certain if both GH peak in stimulating tests and IGF-I secretion were decreased. In turn, both normal GH peak in stimulating tests and normal IGF-I secretion were necessary to exclude GHD. The patients with discordant results of the above-mentioned tests were excluded from this part of the analysis.

In the total group of 1,000 patients, the diagnosis of GHD was undoubtedly confirmed by decreased GH peak in stimulating tests and decreased IGF-I concentration in 193 cases, while excluded on the grounds of normal GH peak in stimulating tests and normal IGF-I concentration in 325 patients. It is worth mentioning that the results of GH stimulating tests and of IGF-I assessment led to divergent diagnostic conclusions in the remaining 482 patients.

According to the results of the ROC curve analysis for the group of 518 patients with certain diagnoses, the cut-off value for screening test of GH secretion after falling asleep that provides its 95% sensitivity should be established on

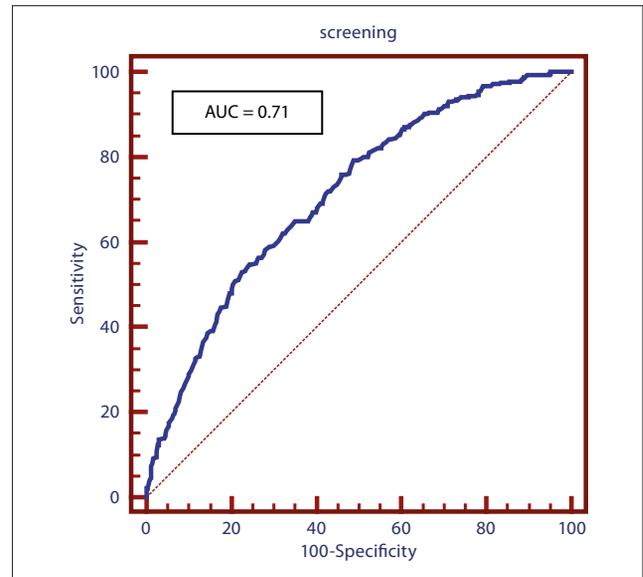


Figure 2. ROC curve for screening test with the cut-off GH peak at the level of 19.0 ng/ml

almost the same level as in previous analysis, *i.e.* 19.0 ng/ml. The specificity of the screening test in its latter variant was 24.9% only. The ROC curve for the latter analysis is presented in Figure 2.

DISCUSSION

It is generally assumed that any screening test should demonstrate at least 95% sensitivity. This condition was not fulfilled by the screening test of nocturnal GH secretion, as its sensitivity with respect to GH peak in stimulating tests as the reference tests is only 70%. In almost 30% of the patients, the result of the screening test was normal, despite decreased GH peak in stimulating tests. Thus, the strict adherence to the principle that normal result of screening test would waive further diagnosis, and could lead to the undiagnosing of 30% of children with GHD. In fact, currently, it is impossible to qualify the child with normal result of the screening test to GH therapy in terms of the therapeutic programme, even if standard diagnostic procedures confirm GHD. The basis for the introduction of the assessment of nocturnal GH secretion into routine practice was the study of Radetti et al. [8] who proved the effectiveness of GH therapy in patients with decreased spontaneous nocturnal GH secretion, but not in those with normal spontaneous GH secretion (even if GH peak in stimulating tests was decreased). To-date, however, it is not recommended to assess only spontaneous GH secretion. Moreover, it has been explicitly stated that the tests of spontaneous GH secretion might not identify all cases of GHD [5].

The likelihood ratios (LRs), calculated for the screening test, were also too low to justify the performance of this test. It should be recalled that LR_s over 5 and below 0.2 demonstrate the usefulness of the diagnostic procedure, while LR_s from 0.5–2.0 argue against the usefulness of that test. In the presented study (LR₊) value was 1.817. This observation should be interpreted in such way that the result of screening test does not practically change the probability of the disease [9].

Next, the PPV equal to 0.462 means that that diagnosis has been confirmed in less than 50% of all the patients with the positive result of screening test. On the other hand, NPV equal to 0.86 means that in 86% patients with negative (normal) result of screening test, GHD is excluded. The last information, however, should be interpreted from the clinical point of view in such a way that 14% of the patients with negative result of screening test is GH-deficient. In all the examined group of 1,000 children, decreased GH peak in stimulating tests, despite normal result of screening test, was observed in almost 10% (96 cases).

In Poland, assessment of GH secretion after falling asleep was introduced a few years ago as a screening procedure in diagnosing GHD in children, in agreement with national recommendations [10]. In that statement, the results presented by Obara-Moszyńska [11] were quoted as the basis for the implementation of these recommendations. In the next study, Obara-Moszyńska et al. [12] assessed the diagnostic usefulness of spontaneous GH secretion test and different pharmacological tests of GH secretion, and observed much higher GH peaks in children after falling asleep than in stimulating tests. Unfortunately, the authors assumed that pituitary insufficiency might be diagnosed after falling asleep by GH surge below 10.0 ng/ml; in fact, treating the screening test as the reference test. The observations made in the current study also point at higher GH secretion after falling asleep than in stimulating test; however, conversely to the quoted authors, the authors of the presented study are convinced that this situation should lead to the increase of the cut-off value of GH peak in screening test.

In our study, the incidence of FN results of screening test was almost 10%. In the previously quoted study by Obara-Moszyńska et al. [12], the test of spontaneous nocturnal GH secretion was performed in 56 patients. In 10 of them, GH peak in this test was normal, despite the results of GH stimulating tests pointing at the pituitary insufficiency. According to the statement of the authors, this was enough to rule out GHD. In the opinion of the authors of the current study, the negative result (*i.e.*, the result considered as normal) of the test of GH secretion after falling asleep, does not constitute in its present form a sufficient basis for the exclusion of GHD.

The increase of the cut-off value of GH peak in screening test up to 19.0 ng/ml, ensuring its adequate sensitivity (95%), is connected with the significant decrease of its specificity (up to less than 25%). In other words, the proportion of children with FP result of screening test seems to be too high to justify the execution of this test in all short children, as only 25% of them might not be subjected to further diagnostic procedures, including stimulating tests.

The lack of concordance between the spontaneous and stimulated GH secretion was reported more than 20 years ago by Donaldson et al. [13] and Bercu et al. [14]. Next, the high incidence of FP results of the two stimulating tests in the child who, in fact, is not GH-deficient, was found in our previous study [7].

Another crucial issue is to prove the good reproducibility of any procedure recommended as a screening test. The poor reproducibility of the results of repeated GH stimulating tests with pharmacological agents has been well documented [15, 16], and also observed in previous studies by the authors of this study [7, 17]. Moreover, the lack of concordance between GH peak after falling asleep and in stimulating tests has

been demonstrated in recent research by the authors of this study [18].

The most important problem seems to be the poor correlation between the results of two consecutive tests of nocturnal GH secretion, performed in the same patient, documented in a previous study [19] in which the authors of the current research interestingly found very similar results in the paper published by Cacciari et al. in 1992 [20].

Summing up, the results of the presented study, as well as of previous reports by the authors, favour the opinion that the test for spontaneous nocturnal GH secretion should be performed only for scientific and not diagnostic purposes [4, 5], at least until GH stimulating tests remain the standard diagnostic procedures. The obtained results strongly contradict the view that the assessment of nocturnal GH secretion should be a screening test in diagnosing GHD in children with short stature.

Acknowledgments

This study was financially by statutory funds from the Medical University of Lodz (503/1-1-107-03/503-01, 503/1-107-04/503-01) and Polish Mother's Memorial Hospital - Research Institute, Lodz, Poland

REFERENCES

1. Wit JM, Ranke M, Kelnar CJH. ESPE Classification of Paediatric Endocrine Diagnoses. *Horm Res.* 2007; 68 (Suppl. 2): 1-120.
2. Savage MO, Burren CP, Rosenfeld RG. The continuum of growth hormone - IGF-I axis defects causing short stature: diagnostic and therapeutic challenges. *Clin Endocrinol.* 2010; 72: 721-728.
3. Attachment to Disposition No. 19 of the President of the National Health Fund 59/2011/DGL of October 10, 2011 (in Polish).
4. GH Research Society. Consensus guidelines for the diagnosis and treatment of growth hormone deficiency in childhood and adolescence: summary statement of the GH Research Society. *J Clin Endocrinol Metab.* 2000; 85: 3990-3993.
5. Webb EA, Dattani MT. Diagnosis of Growth Hormone Deficiency. *Endocr Dev.* 2010; 18: 55-66.
6. Cianfarani S, Liguori A, Germani D. IGF-I and IGFBP-3 assessment in the management of childhood onset growth hormone deficiency. *Endocr Dev.* 2005; 9: 66-75.
7. Hilczer M, Smyczyńska J, Lewiński A. Limitations of clinical utility of growth hormone stimulating tests in diagnosing children with short stature. *Endocr Regul.* 2006; 40: 69-75.
8. Radetti G, Buzi F, Cassar W, Paganini C, Stacul E, Maghnie M. Growth hormone secretory pattern and response to treatment in children with short stature followed to adult height. *Clin Endocrinol (Oxf).* 2003; 59: 27-33.
9. Brożek J, Jaeschke R, Leśniak W, Gajewski P, Mrukowicz J, Bała M. Evaluation of information about diagnostic method. In: Gajewski P, Jaeschke R, Brozek M (eds.) *Basics of EBM that is of medicine based on scientific data for doctors and medical students.* *Medycyna Praktyczna* 2008; Edition 1: 97-109.
10. Walczak M, Romer TE, Korman E, Małeczka-Tendera E, Lewiński A, Rymkiewicz-Kluczyńska B, et al. Management of short stature caused by growth hormone deficiency. *Klinika Pediatryczna* 2005; 13: 212-221.
11. Obara-Moszyńska M. Usefulness of selected tests of growth hormone secretion in diagnostics of somatotrophic pituitary insufficiency. Doctoral Thesis. Karol Marcinkowski Medical University, Poznan, 2003 (in Polish).
12. Obara-Moszyńska M, Kędzia A, Korman E, Niedziela M. Usefulness of growth hormone (GH) stimulating tests and IGF-I concentration measurement in GH deficiency diagnosis. *J Ped Endocrinol Metab.* 2008; 21: 569-579.
13. Donaldson DL, Pan F, Hollowell JG, Stevenson JL, Gifford RA, Moore WV. Reliability of stimulated and spontaneous growth hormone levels in identifying the child with low GH secretion. *J Clin Endocrinol Metab.* 1991; 72: 647-652.
14. Bercu BB, Shulman D, Root AW, Spiliotis BE. Growth hormone provocative testing frequently does not reflect endogenous GH secretion. *J Clin Endocrinol Metab.* 1986; 63: 709-716.
15. Loche S, Bizzarri C, Maghine M, Faedda A, Tziala C, Autelli M, Casini MR, Cappa M. Results of early reevaluation of growth hormone

- secretion in children with apparent growth hormone deficiency. *J Pediatr*. 2002; 140: 445–449.
16. Ropelato MG, Martinez A, Heinrich JJ, Bergada C: Reproducibility of growth hormone secretion tests. *J Pediatr Endocrinol Metab*. 1996; 9: 41–50.
17. Hilczer M, Smyczyńska J, Stawerska R, Lewiński A. Stability of insulin like growth factor-I concentration despite divergent results of repeated growth hormone stimulating tests indicates poor reproducibility of test results. *Endocrine Regul*. 2006; 40: 8–16.
18. Smyczyńska J, Stawerska R, Lewiński A, Hilczer M. Growth hormone (GH) peak after falling asleep reflects spontaneous nocturnal GH secretion, however is not corresponding to the results of GH stimulating tests in children with short stature. *Neuro Endocrinol Lett*. 2012; 33: 37–41.
19. Smyczyńska J, Stawerska R, Lewiński A, Hilczer M. Reproducibility of the nocturnal growth hormone (GH) secretion profile – a screening procedure in diagnosing GH deficiency in children – and optimisation of its duration. *Pediatr Endocrinol* 2010; 9 (3): 9–17.
20. Cacciari E, Tassoni P, Parisi G, Pirazzoli P, Zucchini S, Mandini M et al. Pitfalls in diagnosing impaired growth hormone (GH) secretion: retesting after replacement therapy of 63 patients defined as GH deficient. *J Clin Endocrinol Metab*. 1992; 74: 1284–1289.