This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0). To view a copy of the license, visit **https://creativecommons.org/licenses/by-nc/4.0/**

Annals of Clinical and Analytical Medicine
Original Research

A comparison of cortisol responses to the insulin tolerance test and glucagon stimulation test in patients with suspected adrenal insufficiency

Cortisol response in adrenal insufficiency

Damla Tüfekçi[1], İrfan Nuhoglu[2], Hülya Coşkun[3], Savaş Volkan Kişioğlu[4], Ahmet Suat Demir[4], Yasemin Emür Günay[5], Mustafa Koçak[2], Halil Önder Ersöz[6]

[1]Department of Endocrinology and Metabolism, Recep Tayyip Erdogan University Faculty of Medicine, Rize, Türkiye.
[2]Department of Endocrinology and Metabolism, Karadeniz Technical University Faculty of Medicine, Trabzon, Türkiye.
[3]Private Endocrinology Office, Sakarya, Türkiye.
[4]Clinic of Endocrinology and Metabolism, Medical Park Hospital, Trabzon, Türkiye.
[5]Clinic of Endocrinology and Metabolism, Giresun Training and Research Hospital, Giresun, Türkiye.
[6]Clinic of Endocrinology and Metabolism, 7M Hospital, Trabzon, Türkiye.

Abstract

Aim: Evaluation of the hypothalamic-pituitary-adrenal (HPA) axis is highly important as a key determinant of the glucocorticoid replacement decision in patients with suspected adrenal insufficiency. This study compared the cortisol responses to the insulin tolerance test (ITT) and glucagon stimulation test (GST) in patients with suspected adrenal insufficiency.

Methods: One hundred seventeen patients (median age: 41.59, min-max 17.0-73.0 years, 75% female) with suspected adrenal insufficiency were included in this cross-sectional study. The ITT and GST were applied to all patients in a random order following overnight fasting.

Results: The mean peak cortisol level was 11.71 in the ITT and 17.45 in the GST, the difference being statistically significant (t = -11.016, P < .001). Examination of the cross distributions of ITT and GST response status revealed a negative GST response in 55.6% and a positive response in 44.4% of the patients with no ITT response. The mean Δcortisol values were 4.77 in the ITT and 10.88 in the GST, the increase in the cortisol response being significantly higher compared to the ITT (P < .001). In terms of time-dependent changes in cortisol levels, a later-onset and more prolonged cortisol response was observed in the GST.

Conclusion: The study findings revealed that both peak cortisol levels and Δcortisol values detected in the GST were significantly higher than those in the ITT when the two tests were applied to the same patients. The GST emerged as a reliable and sufficient alternative dynamic test compared to the ITT, regarded as the gold standard in the evaluation of adrenal insufficiency.

Keywords

adrenal insufficiency, insulin tolerance test, glucagon stimulation test

DOI: 10.4328/ACAM.50228 Ann Clin Anal Med 2026;17(8):774-778
Received: 03.06.2026 Accepted: 21.07.2026 Published Online: 25.07.2026 Printed: 01.08.2026
Corresponding Author: Damla Tüfekçi, Department of Endocrinology and Metabolism, Recep Tayyip Erdogan University Faculty of Medicine, Rize, Türkiye.
E-mail: drtufekci@gmail.com **P.:** +90 545 553 77 53
Corresponding Author ORCID: https://orcid.org/0000-0001-5928-873X
Other Authors ORCID: İrfan Nuhoglu, https://orcid.org/0000-0003-0650-3242 • Hülya Coşkun, https://orcid.org/0000-0003-2053-8898 • Savaş Volkan Kişioğlu, https://orcid.org/0000-0002-7742-048X • Ahmet Suat Demir, https://orcid.org/0000-0002-2282-3820 • Yasemin Emür Günay, https://orcid.org/0000-0002-0645-2070 • Mustafa Koçak, https://orcid.org/0000-0001-9193-7277 • Halil Önder Ersöz, https://orcid.org/0000-0001-7157-1116

Introduction

Evaluation of the hypothalamic-pituitary-adrenal (HPA) axis is difficult due to both the daily rhythm and the pulsatile nature of anterior pituitary hormone secretion. While morning serum cortisol levels are useful, the majority of patients require dynamic stimulation tests for HPA axis evaluation.[1] Although the insulin tolerance test (ITT) is regarded as the gold standard in evaluating the HPA axis in patients with pituitary disorders, the test is a laborious one requiring medical supervision.[2,3] It is also contraindicated in patients with epilepsy and cerebrovascular or cardiovascular disorders, and must be used with care in elderly patients. Various alternatives to the ITT have been proposed, such as the adrenocorticotropic hormone (ACTH) test and glucagon stimulation test (GST). The GST is an attractive alternative since it can be used to assess the HPA axis.[4]

Accurate diagnosis of secondary adrenal insufficiency is highly important in patients with pituitary gland disease, because the decision to apply glucocorticoid replacement therapy is based on the detection of HPA axis insufficiency.

An inexpensive, more reliable, and practical test to the ITT is needed. A number of studies have compared these tests in healthy individuals and patients with pituitary disorders, but few have compared them directly, in the same individuals. The aim of this study was to investigate whether or not the GST can be employed instead of the ITT in evaluating the HPA axis.

Materials and Methods

One hundred fifty-six adults with suspected adrenal insufficiency (117 women, 39 men) were included in the study. Individuals with (i) major systemic diseases, (ii) contraindications for ITT such as epilepsy or cerebrovascular or cardiovascular diseases, (iii) Cushing's disease, (iv), known diabetes mellitus requiring drug therapy, or (v) estrogen replacement were excluded.

The ITT and GST tests were applied on separate days following overnight fasting, in a random order. All patients were in a euthyroid state when the dynamic tests were performed. Glucocorticoids were discontinued at least 24 hours before the tests. In case of a known potential cause of adrenal insufficiency, such as surgery, the dynamic tests were conducted at least three months after that event.

The ITT was performed with a 0.10 U/kg intravenous bolus injection of regular human insulin (Humulin R, Eli Lilly, Indianapolis, IN, USA). Blood samples for serum cortisol measurements were collected at baseline, at the time of symptomatic hypoglycemia (based on clinical findings and capillary blood glucose levels), and after 30, 60, 90, and 120 minutes. Only patients with biochemically confirmed hypoglycemia (a serum glucose level ≤40 mg/dl) were enrolled. Serum cortisol levels ≥ 18 µg/dL were regarded as an adequate response to the ITT.

The GST was conducted with the subcutaneous injection of 1 mg glucagon (Novo Nordisk, Bagsvaerd, Denmark). Blood samples for cortisol measurements were collected 90, 120, 150, 180, 210, and 240 minutes after the injection.

Serum cortisol levels were determined using the radioimmunoassay method with a commercially available kit (Immunotech SRO, Czech Republic).

Ethical Approval

The study was approved by the Ethics Committee of Karadeniz Technical University Faculty of Medicine.

Statistical Analysis

Normality assumptions by groups were first examined using the Shapiro-Wilk test. The independent two-sample *t*-test was applied in the comparison of normally distributed data between paired groups, and the Mann-Whitney U test for non-normally distributed data. During group comparisons of categorical data, Pearson's chi-square test was applied if all expected observation values were greater than 5, or Fisher's exact test if any of these was smaller than 5 (Howell, 2011).

Since normality was not established, Friedman's test was applied in the comparison of time-dependent changes during the two tests. Significant differences emerging from the Friedman test were then compared using the Bonferroni-corrected Wilcoxon signed-rank test and were presented using a lettering method. A paired two-sample t-test was applied for normally distributed data in the comparison of basal and peak cortisol levels in the ITT and GST, and Wilcoxon's signed-rank test for non-normally distributed data. Relationships between quantitative data meeting normality assumptions were examined using Pearson's correlation coefficient.

The analysis results were presented as mean ± standard deviation, median, minimum, and maximum values for quantitative data and frequency (*n*) and percentage (%) values for categorical data. Significance was set at $P < .05$ for all calculations. All analyses and evaluations were performed on IBM SPSS version 26 software (IBM Corp., released 2019).

Reporting Guidelines

This study was reported in accordance with the STROBE guideline.

Results

Descriptive statistics for the demographic and clinical characteristics of the patients in the study are shown in Table 1.

Women represented 75% of the study group and men 25%, with a mean age of 41.59 years and a mean BMI of 26.87. Mean systolic values at blood pressure measurement were 118.21, and mean diastolic values were 74.81.

Pituitary adenoma was present in 35.1% of the participants, functional adenoma being detected in 56.4% of those. A history of pituitary surgery was present in 19.9% of patients, and of Cushing surgery in 5.8%. No additional disease was present in 73.7% of patients, while hypertension was determined in 13.5%, diabetes mellitus in 10.9%, and chronic artery disease in 1.9%.

The most common symptom was malaise (51.9%), followed by lightheadedness (10.9%) and hypoglycemia (7.1%). No symptom was present in 23.7% of the participants, and steroid use was present in 14.7%.

To determine time-dependent changes in cortisol levels during the ITT and GST, cortisol levels were calculated from existing measurements for each time point, the sample size being n = 156 for all analyses. The Friedman test was applied using observations at each particular time point. Changes over time in cortisol levels were statistically significant for both tests.

ITT cortisol levels exhibited significant changes over time ($P < .001$). The median value at baseline (0 min) was 6.28, rising to 7.7 at 15 min and 8.8 at 30 min. The highest levels were obtained between 30 and 60 min, and were similar at these time points in Bonferroni-corrected analysis. The median cortisol level at 90 min decreased to 8.05, and to 6.91 at 120 min, a value close to baseline. These findings show an early, brief cortisol response in the ITT.

GST cortisol values also exhibited significant changes over time ($P < .001$). The median baseline cortisol value was 5.94, rising to 7.88 at 90 min, 8.46 at 120 min, and 12.75 at 150 min. The highest cortisol levels were obtained between 150 and 210 min, and were similar between these time points at Bonferroni-corrected analysis. The median cortisol level at 240 min decreased to 11. These findings show a later-onset and more prolonged cortisol

Table 1. *Descriptive statistics for demographic and clinical characteristics*

Variables	*n* or Mean ± SD	% or median (min - max)
Sex		
Female	117	75
Male	39	25
Age	41.59 ± 12.24	43 (17-73)
BMI, kg/m^2	26.87 ± 4.92	26 (18-42)
Systolic BP, mmHg	118.21 ± 13.66	120 (90-165)
Diastolic BP, mmHg	74.81 ± 10.54	70 (50-110)
Pituitary adenoma		
No	101	64.7
Present	55	35.3
Functional pituitary adenoma		
No	24	43.6
Present	31	56.4
History of pituitary surgery		
No	125	80.1
Yes	31	19.9
History of cushing surgery		
No	147	94.2
Yes	9	5.8
Additional diseases		
None	115	73.7
Hypertension	21	13.5
Diabetes	17	10.9
CAD	3	1.9
Symptoms		
Malaise	81	51.9
Low blood pressure	6	3.8
Lightheadedness	17	10.9
Hypoglycemia	11	7.1
Nausea, abdominal pain	4	2.6
None	37	23.7
Steroid use		
No	133	85.3
Yes	23	14.7

SD= Standard deviation, Min= Minimum, Max= Maximum BMI= Body mass index, BP= Blood pressure; CAD= Coronary artery disease

response in the GST.

The median baseline cortisol level in the ITT was 6.28, with a median peak value of 11.7. The difference was statistically significant (Z = 10.264, P < .001) (Figure 1).

The mean baseline cortisol level in the GST was 6.58, rising significantly to a peak value of 17.45 (t = -23.079, P < .001) (Figure 1).

The mean peak cortisol level in the ITT was 11.71, compared to 17.45 in the GST, the difference being statistically significant (t = -11.016, P < .001) (Figure 1).

Examination of the cross distributions of ITT and GST response statuses revealed a negative GST response in 55.6% and a positive response in 44.4% of the patients with no ITT response and a

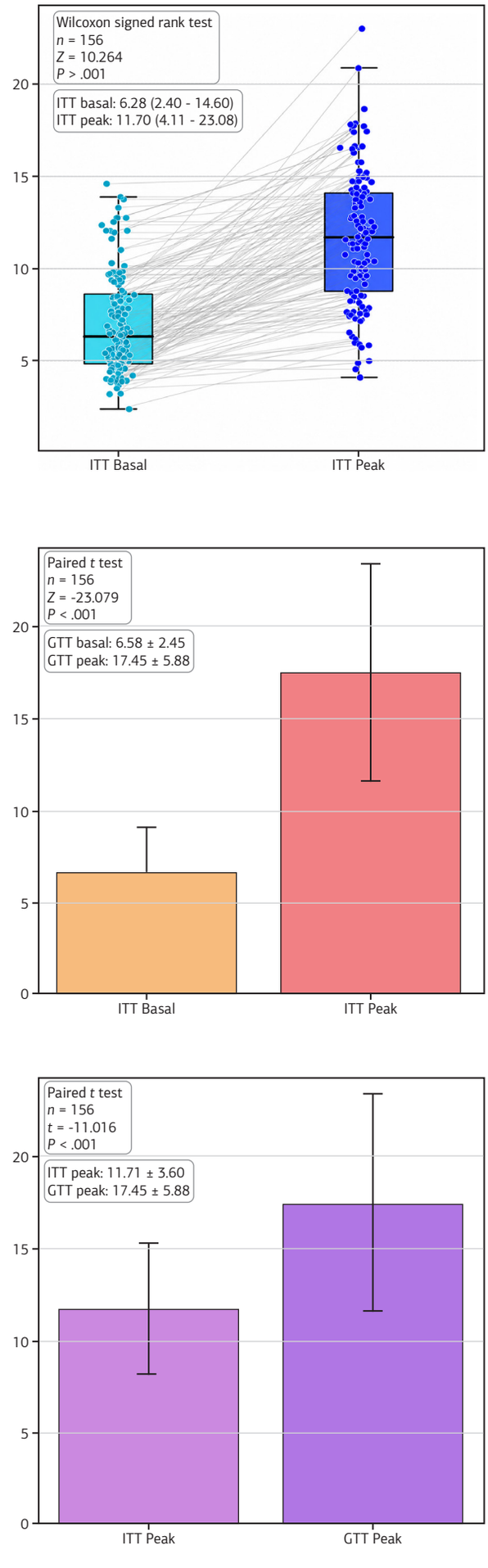


Figure 1. *A comparison of peak and baseline cortisol levels in the ITT and GST*

positive response in 44.4% of the patients with no ITT response. The number of patients responding to the ITT was quite low (n = 3), a negative GST response being determined in 66.7% of these and a positive GST response in 33.3%. No statistically significant association was generally determined between ITT and GST responses (Fisher exact *P* = 1.000).

The relationship between the peak cortisol levels registered during the ITT and GST is shown with a distribution chart in Figure 2. Correlation analysis was performed using Pearson's correlation test. The regression line and 95% confidence interval are shown in the figure. Although the analysis revealed a weak, positive correlation between the two variables, this was not statistically significant (r = 0.120, *P* = .135).

Mean Δcortisol values were 4.77 in the ITT and 10.88 in the GST, the GST exhibiting a significantly higher cortisol response increase than the ITT (*P* < .001) (Table 2).

Median peak cortisol values were 45 in the ITT and 180 in the GST, the difference being statistically highly significant (*P* < .001) (Table 2).

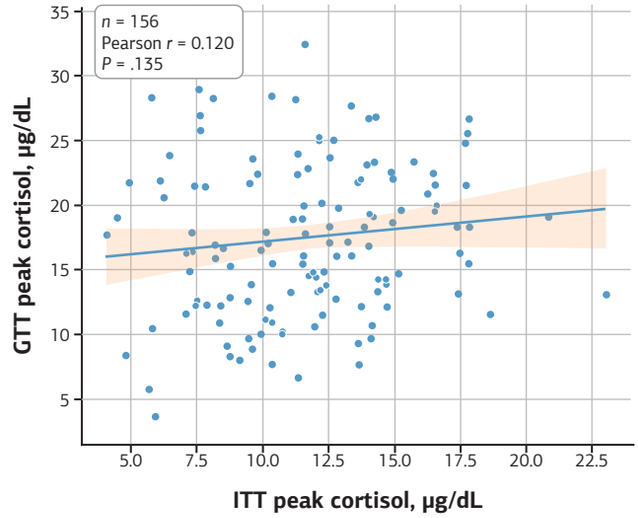


Figure 2. *The relationship between ITT and GST peak cortisol values*

Discussion

Biochemical tests are needed to confirm the diagnosis when adrenal insufficiency is suspected. The first step in the evaluation is the measurement of baseline morning serum cortisol levels. Basal cortisol measurement may not identify adrenal insufficiency in some patients, and several dynamic tests may be required.

Although the ITT is regarded as the gold standard in evaluating the HPA axis in patients with pituitary disorders, the test is a laborious one requiring medical supervision. In addition, it is contraindicated due to potential side-effects in patients with epilepsy and cerebrovascular and cardiovascular diseases, and must be employed with care in elderly patients. A reliable, practical, and less costly test capable of use instead of the ITT is therefore needed. Its various advantages, such as simplicity, easy availability, and fewer side-effects, gave rise to the idea that the GST can be used as a primary stimulation test in patients with relative ITT contraindication. Few studies, involving only limited numbers of patients, have evaluated the effectiveness of the GST and ITT in this important decision. Studies have usually involved patients with detected pituitary pathologies. The present study evaluated a wide and varied population with and without histories of pituitary pathology. Additionally, this study directly compared the GST and ITT in the evaluation of the HPA axis, and investigated whether the GST may be advantageous due to its ease of use.

Some authors have proposed that the GST exhibits a lower effect in terms of HPA axis evaluation, in other words, in terms of the size of ACTH and cortisol secretion.[5]

In contrast, other authors maintain that the ITT and GST are similarly useful in the evaluation of anterior pituitary function.[6,7] The authors of another study reported no difference in mean maximum cortisol concentrations obtained during the GST in comparison to the ITT.[8]

In another study on that subject, all 129 individuals diagnosed with pituitary disease underwent the ITT, a low-dose ACTH test, and the GST. Those authors reported that peak cortisol levels during the ITT were significantly lower than those obtained during both the response to the ITT exhibited inadequate cortisol responses to low-dose ACTH and the GST, the majority (61%) registering an adequate cortisol response.[7] Similarly in the current research, the GST response was negative in the majority (55.6%) of patients exhibiting no response to the ITT, while positive responses were determined in 44.4%.

Despite its longer duration, the GST represents a good alternative to the ITT. Like the ITT, both ACTH and growth hormone (GH) are released in the GST. The cortisol response in the present study also exhibited a later onset and longer duration.

However, limited data are available in the literature regarding modern analyses of the validity of the GST in the evaluation of the HPA axis and comparing its sensitivity and specificity to those of the ITT. Pediatric studies have described the GST as a reliable test of the HPA in children; girls exhibited higher peak cortisol levels on the GST (mean difference approximately 55 nmol/L) and even greater differences following the onset of puberty (median 672 nmol/L compared to 490 nmol/L, *P* = .002).[9] A recent study comparing the two tests in adults reported that arousal magnitudes of both cortisol and GH (evaluated via Δcortisol and ΔGH) were effectively the same during the GST and ITT. The authors therefore concluded that both tests exhibited equally good performance in evaluating cortisol or GH release, and that neither was markedly 'superior' to the other. In the present study, however, both peak cortisol levels and Δcortisol determined during the GST were significantly higher compared to the ITT.

Another previous study showed that the cortisol threshold should be 9.1 µg/dL in healthy adults subjected to the GST.[10] However, the purpose of the present study was not to determine optimal threshold values for cortisol concentrations. Instead, it was to investigate and compare the degree of cortisol release (in other words, Δcortisol) following the relevant stimulations.

The physiological mechanisms by which glucagon stimulates cortisol release have not yet been adequately explained.[4,5] The HPA axis may not be capable of evaluation via a single test in some patients, and more than one test may be required to diagnose adrenal insufficiency, as shown in previous research.[11] The present study showed that the GST may represent an alternative gold standard test to the ITT. Moreover, the GST and ATT were

Table 2. *A comparison of Δcortisol and peak times in the ITT and GST*

	n	ITT	GST	Test statistic	*P*
ΔCortisol	156	4.77 ± 3.46	10.88 ± 5.89	*t* = 11.969	< .001
Peak time, min	156	45 (15-120)	180 (60-240)	*Z* = 10.833	< .001

t = Paired two sample t test, Mean ± SD

A prospective study from 2018 in which the ITT and GST were evaluated in 81 patients reported a significant correlation between the cortisol response to glucagon and the response in the ITT. The GST was regarded as a valid alternative test, due to its low side-effects and consistent results, in situations in which the ITT is contraindicated.[8] In the present study, analysis of peak levels from both tests revealed a weak, positive correlation between the two variables.

Limitations

The principal limitation of this study is the relatively small number of participants. In addition, another confirmatory test, the Synacthen test, could not be performed due to disruptions in drug supplies. Moreover, the very low number of positive cases in the ITT group (n = 7) led to uncertainty in sensitivity and specificity calculations and an inability to reliably establish a receiver operating characteristic curve. No clinically interpretable cut-off value could therefore be determined. More extensive research is now needed in order to be able to assess diagnostic performance from the methodological perspective.

Conclusion

In conclusion, the GST results in a later-onset and longer-lasting response and to marked increases in peak cortisol and Δcortisol values compared to the ITT. We therefore conclude that the GST should not automatically be regarded as an 'inferior' option to the ITT.

Declarations

Animal and Human Rights Statement
All procedures performed in this study were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed Consent
Informed consent was obtained from all subjects involved in the study.

Data Availability
The data supporting the findings of this study are available from the corresponding author on reasonable request.

Conflict of Interest
The authors declare no conflict of interest.

Funding
None.

Author Contributions (CRediT Taxonomy)
Conceptualization: D.T., I.N.
Methodology: D.T.
Software: D.T.
Validation: H.C., M.K., H.O.E.
Formal analysis: D.T.
Investigation: D.T.
Resources: D.T.
Data curation: D.T., A.S.D, S.V.K, Y.E.G.
Writing—original draft preparation: D.T.
Writing—review and editing: H.C.
Visualization: M.K.
Supervision: D.T.
Project administration: D.T.
All authors have read and agreed to the published version of the manuscript.

AI Usage Disclosure
The authors declare that no AI-assisted technologies were used.

Abbreviations
ACTH: Adrenocorticotropic hormone
GST: Glucagon stimulation test
HPA: Hypothalamic-pituitary-adrenal
ITT: Insulin tolerance test

References
1. Grinspoon SK, Biller BM. Clinical review 62: laboratory assessment of adrenal insufficiency. J Clin Endocrinol Metab. 1994;79(4):923-931. doi:10.1210/jcem.79.4.7962298
2. Arlt W, Allolio B. Adrenal insufficiency. Lancet. 2003;361(9372):1881-1893. doi:10.1016/s0140-6736(03)13492-7
3. Schneider HJ, Aimaretti G, Kreitschmann-Andermahr I, Stalla GK, Ghigo E. Hypopituitarism. Lancet. 2007;369(9571):1461-1470. doi:10.1016/s0140-6736(07)60673-4
4. Leong KS, Walker AB, Martin I, Wile D, Wilding J, MacFarlane IA. An audit of 500 subcutaneous glucagon stimulation tests to assess growth hormone and ACTH secretion in patients with hypothalamic-pituitary disease. Clin Endocrinol (Oxf). 2001;54(4):463-468. doi:10.1046/j.1365-2265.2001.01169.x
5. Berg C, Meinel T, Lahner H, Yuece A, Mann K, Petersenn S. Diagnostic utility of the glucagon stimulation test in comparison with the insulin tolerance test in patients following pituitary surgery. Eur J Endocrinol. 2010;162(3):477-482. doi:10.1530/eje-09-0824
6. Toogood A, Brabant G, Maiter D, et al. Similar clinical features among patients with severe adult growth hormone deficiency diagnosed with insulin tolerance test or arginine or glucagon stimulation tests. Endocr Pract. 2012;18(3):325-334. doi:10.4158/ep11146.or
7. Simsek Y, Karaca Z, Tanriverdi F, Unluhizarci K, Selcuklu A, Kelestimur F. A comparison of low-dose ACTH, glucagon stimulation, and insulin tolerance tests in patients with pituitary disorders. Clin Endocrinol (Oxf). 2015;82(1):45-52. doi:10.1111/cen.12528
8. Ach T, Yosra H, Jihen M, et al. Cortisol cut points for the glucagon stimulation test in the evaluation of the hypothalamic-pituitary-adrenal axis. Endocr J. 2018;65(9):935-942. doi:10.1507/endocrj.ej18-0147
9. Borghammar C, Svensson J, Tidblad A, Elfving M. Sex and age differences in cortisol levels during glucagon stimulation test in children. BMC Pediatr. 2025;25(1):440. doi:10.1186/s12887-025-05784-5
10. Karaca Z, Lale A, Tanriverdi F, Kula M, Unluhizarci K, Kelestimur F. The comparison of low- and standard-dose ACTH and glucagon stimulation tests in the evaluation of the hypothalamo-pituitary-adrenal axis in healthy adults. Pituitary. 2011;14(2):134-140. doi:10.1007/s11102-010-0270-3
11. Ospina NS, Al Nofal A, Bancos I, et al. ACTH stimulation tests for the diagnosis of adrenal insufficiency: systematic review and meta-analysis. J Clin Endocrinol Metab. 2016;101(2):427-434. doi:10.1210/jc.2015-1700

Additional Information
Publisher's Note
Bayrakol MP remains neutral with regard to jurisdictional and institutional claims.

How to cite this article:
Damla Tüfekçi, İrfan Nuhoglu, Hülya Coşkun, Savaş Volkan Kişioğlu, Ahmet Suat Demir, Yasemin Emür Günay, Mustafa Koçak, Halil Önder Ersöz. A comparison of cortisol responses to the insulin tolerance test and glucagon stimulation test in patients with suspected adrenal insufficiency. Ann Clin Anal Med 2026;17(8):774-778 doi:10.4328/ACAM.50228