Corticotropic Tests for Hypothalamic-Pituitary Adrenal Insufficiency: A Meta-Analysis

Short Title: Meta-analysis: Tests for HPAI

Consortium for evaluation of corticotropin test in hypothalamic-pituitary adrenal insufficiency

Affiliations
1John H. Stroger Jr. Hospital of Cook County & Rush Medical College, Chicago, IL; 2Walsall Manor Hospital NHS Trust, Walsall, UK; 3University of Milan & IRCCS Policlinico San Donato, Italy; 4Regional Centre for Endocrinology and Diabetes, Royal Victoria Hospital, Belfast, UK; 5Queen Elizabeth Hospital, Hong Kong SAR; 6Keele University and North Staffordshire Hospital, Stoke on Trent, UK; 7Hacettepe University, Ankara, Turkey; 8IRCCS G. Gaslini, University of Genova, Italy; 9Cincinnati Children’s Hospital Medical Center & University of Cincinnati, OH; 10Christchurch Hospital, New Zealand; 11Institute of Endocrinology, Metabolism and Hypertension, Tel Aviv Sourasky Medical Center, Israel

Corresponding Author/ Reprint requests

Rasa Kazlauskaite MD MSc
1900 W. Polk St, Room 806
Chicago, IL 60612
rasa_kazlauskaite@rush.edu
(773) 505-7667, Fax: (312) 864-9735

Keywords: Adrenal insufficiency, cortisol, corticotropin stimulation test, HPA insufficiency

Disclosure summary:
The authors have nothing to disclose.

Number of words (text): 3915
Number of words (abstract): 240
Number of tables: 1
Number of figures: 3
Abstract

Context: Diagnostic value of tests for detecting hypothalamic-pituitary adrenal insufficiency (HPAI) is controversial.

Objective: To compare standard-dose and low-dose corticotropin tests for diagnosing HPAI.

Data Sources: PubMed database from 1966 to 2006 for studies reporting diagnostic value of standard-dose or low-dose corticotropin tests, with patient-level data obtained from original investigators.

Study Selection: Eligible studies had more than 10 patients; all subjects were evaluated because of suspicion for chronic HPAI; and patient-level data was available. We excluded studies with no accepted reference standard for HPAI (insulin hypoglycemia or metyrapone test), if test subjects were in the intensive care unit, or if only normal healthy subjects were used as controls.

Data Extraction: We constructed receiver operator characteristic (ROC) curves using patient-level data from each study, and then merged results to create summary ROC curves, adjusting for study size and cortisol assay method. Diagnostic value of tests was measured by calculating area under the ROC curve (AUC) and likelihood ratios.

Data Synthesis: Patient-level data from 13 of 23 (57%) studies (679 subjects) were included in the meta-analysis. The areas under the ROC curves were: low-dose corticotropin test 0.92 (95% CI: 0.89-0.94), and standard-dose corticotropin test 0.79 (95% CI: 0.74-0.84). Among patients with paired data (7 studies, 254 subjects), diagnostic value of low-dose corticotropin test was superior to standard-dose test, AUC 0.94 and 0.85, respectively (P<0.001).

Conclusions: Low-dose corticotropin test was superior to standard-dose test for diagnosing chronic HPAI, however it has technical limitations.

Abbreviations (also defined in text): LDCT – low-dose corticotropin stimulation test; SDCT – standard-dose corticotropin stimulation test; HPA axis – hypothalamic-pituitary-adrenal axis; HPAI – hypothalamic-pituitary adrenal insufficiency; AUC – area under the curve; ROC curve – receiver-operator characteristic curve; CI – confidence interval.
Chronic hypothalamic-pituitary adrenal insufficiency (HPAI) is characterized by impaired adrenocorticotropin (ACTH) secretion as a result of disease or injury to the pituitary or hypothalamus or from prolonged exogenous glucocorticoid administration that exceeds physiological doses.

The reference tests for establishing the integrity of the hypothalamic–pituitary–adrenal (HPA) axis require assessing the response to either a strong stimulus (e.g., insulin-induced hypoglycemia) or an interruption of the negative feedback from cortisol (overnight metyrapone test). However, these reference tests have major drawbacks. The insulin tolerance test is contraindicated in the elderly and those with a history of seizures or cardiovascular disease and requires continuous physician supervision to monitor for serious adrenergic or neurological symptoms (1). The overnight metyrapone test carries a risk of adrenal crisis, and errors can occur from other drugs affecting metyrapone clearance (2). Thus, there is a great clinical demand for alternative tests that are quicker, cheaper, and safer.

The rationale for using the corticotropin analog stimulation test is the assumption that in chronic endogenous corticotropin deficiency, acute responsiveness of the adrenal **zona fasciculata** is diminished and fails to mount an adequate cortisol response (3). We examined the published literature for evidence on two corticotropin analog stimulation tests—using either a standard (250 micrograms) or low-dose (1 microgram) of the corticotropin analog. The primary objective of our quantitative meta-analysis was to compare the performance of the standard-dose and low-dose corticotropin tests in diagnosing HPAI (defined by results of the insulin tolerance test or overnight metyrapone test). The second objective was to determine how best to incorporate the tests in clinical practice, especially in relation to the basal cortisol level.

**Methods**

**Data collection:** We searched the PubMed (www.PubMed.gov) database from 1966 to 2006 for articles with keywords “adrenal insufficiency” and “diagnosis” and limited the search to human studies published in the English language. We selected studies with at least 10 subjects with suspected HPAI and required that the disease be verified with either the insulin tolerance test or metyrapone test. We then tried to contact the principal investigator of each relevant study to request their patient-level data on the following variables: results of the integrated HPA axis reference test, baseline cortisol value, and cortisol values after the standard-dose and low-dose corticotropin tests.

**Population characteristics:** To be eligible for inclusion, adult and pediatric subjects had to be suspected of HPAI from disease or injury to the pituitary or hypothalamus or from prolonged exogenous glucocorticoid administration in supra-physiological doses. Patients had to be affected by hypothalamic-pituitary disease for at least 4 weeks in order to exclude acute hypothalamic or pituitary disorders. A normal sleep–wake cycle was required (or assumed, if no information). We did not include data from studies done in critical care settings.

In actual clinical practice, testing for HPAI is performed when there is some suspicion for HPAI. We have therefore investigated the performance of the tests in this at-risk population in order to avoid the problem of spectrum bias (4), which occurs when tests are evaluated among patients who are different from the ones who will be tested in practice. Thus, we excluded from analysis those subjects who were described as normal healthy control subjects (those recruited as “healthy volunteers” and not due to suspected pituitary disease based on signs/symptoms or imaging).
The most common reason for testing (43%) was because of a pituitary macroadenomas (before or after surgical or radiation treatment). The second most common reason was treatment of other intracranial tumors (23%); 14% of patients were tested due to pituitary disease other than pituitary macroadenoma; 13% were tested due to other pituitary hormone deficiencies (growth hormone or panhypopituitarism). Prolonged supraphysiological glucocorticoid treatment prompted HPAI evaluation in 7% of total patients, but in only 8 patients were there paired data on LDCT and SDCT (5-7). The only study (5) that tested patients solely due to suspected glucocorticoid-induced HPAI was analyzed separately (see Figure 1: Kane), but this study did not have paired SDCT and LDCT data.

**Reference test:** The diagnosis of HPAI was based on an abnormal response to one of the two reference standards for evaluating the integrity of the HPA axis: insulin tolerance test or overnight metyrapone test. We relied on individual study investigators to correctly dichotomize the reference test results into HPAI present or absent.

**Cortisol assay:** Cortisol assays are not standardized and vary across hospitals and studies (8-10). The cortisol assay methods used in various studies included radioimmunoassay or commercially available immunometric methods. To adjust for cortisol assay variability, first, we excluded 5 studies that used the fluorescence polarization assay, which is less specific than more modern methods (6) and has never been utilized to evaluate LDCT. Second, as plasma cortisol is known to be consistently higher than serum cortisol levels, we converted plasma values (used in 2 studies (6, 11)) to their expected serum value by multiplying by the published correction factor of 0.877 (7).

We also investigated whether the results are sensitive to assay type by reassessing the SDCT and LDCT comparisons using the same thresholds for all studies (weighted mean cortisol thresholds) rather than individual study cortisol thresholds (see statistical analysis below).

**Standard-dose corticotropin stimulation test:** One of the two available synthetic corticotropin analogs—cosyntropin (Cortrosyn, Amphastar Pharmaceuticals, Inc.) or tetracosactrin (Synacthen, Novartis Pharma, Switzerland)—was administered intravenously at a dose of 250 micrograms, and serum cortisol levels were obtained at baseline and at least once after injection (most commonly at 30 or 60 minutes).

A dose of 250 mcg (0.25 mg) of Cosyntropin or Tetracosactrin is equivalent to 25 USP units of corticotropin, indicating the equivalence of Synacthen and Cortrosyn formulations. For brevity, we use the term “corticotropin” for both analogs, while acknowledging that Synacthen and Cortrosyn are synthetic corticotropin analogs, different from the native ACTH molecule.

**Low-dose corticotropin stimulation test:** The low-dose test was performed in the morning with patients fasting. One of the two synthetic corticotropin analogs (cosyntropin or tetracosactrin) was administered intravenously in a 1 microgram dose, after being prepared using the method of Dickstein and colleagues (8). Serum cortisol was measured at baseline and at 30 minutes post-injection (except for 1 study, which measured it at 20 minutes). We excluded studies that used a different dose or different protocol.

**Basal cortisol:** All studies measured serum cortisol between 8AM and 10AM after an overnight fast (basal cortisol). The database from one study (9) did not provide basal cortisol.

**Statistical analysis:** We conducted data analysis using Stata statistical software, version 10.0 (StataCorp LP, College Station, TX).
To compare performance of LDCT and SDCT (and basal cortisol), we used receiver-operator-characteristic (ROC) curve analysis. From each study’s data, we calculated the area under the ROC curve (AUC) with 95% confidence intervals (Figure 1). The same methods were used to compare test performance at different time points for measuring stimulated cortisol after corticotropin injection (usually, 30 minutes or 60 minutes).

For each study, dependent on data availability, we conducted paired analyses of the areas under the ROC curves for the 3 test pairs—SDCT and LDCT, basal and SDCT, and basal and LDCT—to evaluate their relative performance in predicting results of the integrated HPA axis testing (reference standard). The AUC is independent of measurement units, thus overall comparison of SDCT and LDCT is not influenced by type of cortisol assay.

In order to combine data across studies, we categorized cortisol response into 3 intervals (high, indeterminate and low likelihood of HPAI), using standard criteria applied consistently to all studies. Two thresholds were defined for each test: the threshold below which cortisol values had high likelihood of HPAI (likelihood ratio [LR] > 9; rule-in threshold), and the threshold above which cortisol values had a low likelihood of HPAI (LR < 0.15; rule-out threshold). Cortisol values between these thresholds (LR 0.15–9) defined the interval with indeterminate likelihood of HPAI.

Likelihood ratios were calculated as the ratio of two probabilities: the probability of the test result among patients with HPAI, divided by the probability of the same test result among patients without HPAI.

The use of categorized cortisol responses recalibrates individual study results to common metric, therefore permitting paired comparisons of the ROC areas for SDCT, LDCT and basal cortisol, stratified by study and cortisol assay.

Comparison of corticotropin stimulation tests using summary cortisol thresholds.
We combined results across studies by calculating a weighted average of the cortisol values defining the rule-in threshold and the rule-out threshold for each test (Table 1). The weights were based on study sample size. Using these new summary thresholds, regardless of cortisol assay method, we re-categorized the data from all studies and combined results in summary ROC curves. This analysis was performed to determine whether the results of SDCT and LDCT performance comparisons were sensitive to cortisol assay type (Figure 2B).

Corticotropin-stimulated cortisol in subjects with indeterminate basal cortisol.

To address the research goal of defining an optimal testing strategy, we tested the sequential testing strategy of first measuring basal cortisol and then measuring a stimulated cortisol in only those subjects with an intermediate basal value. We used paired ROC curve analyses, adjusted for study size and cortisol assay (as described above), to compare LDCT and SDCT in subjects with indeterminate basal cortisol results (LR 0.15-9 for HPAI).

Optimal testing strategy

Based on results of meta-analyses, we derived an optimal testing strategy algorithm (Figure 3). The basal cortisol and LDCT summary thresholds described in Figure 3 were based on the mean cortisol values weighted by study sample size (last row of Table 1).

We used summary thresholds for the two sequential tests to define 5 subgroups: 1) low basal cortisol; 2) high basal cortisol; 3) indeterminate basal cortisol and low
stimulated cortisol; 4) indeterminate cortisol and indeterminate stimulated cortisol; and 5) indeterminate basal cortisol and high stimulated cortisol. We calculated the expected probability of HPAI, with 95% confidence intervals, within each of the 5 subgroups.

**Results**

We were able to analyze patient-level data from 13 published studies, with 7 studies having paired data comparing results of the low-dose and standard-dose corticotropin stimulation tests in the same patients (Table 1). The prevalence of HPAI in the study samples ranged from 18 percent to 58 percent, with a mean of 33 percent (Table 1).

Cortisol testing methods varied from individual radioimmunoassay kits (6, 7, 11, 14-17) to immunometric test kits from various manufacturers (5, 18-22). The lack of a standard cortisol assay method (8-10) explains some of the variability in diagnostic cortisol thresholds reported across studies.

**Standard-dose corticotropin test**

After standard-dose corticotropin stimulation, there was variability across studies in the optimal timing for measuring cortisol response; however, in no study was there a statistically significant difference in diagnostic discrimination at 30 minutes, 60 minutes, or at peak response. In 6 studies, there was a trend favoring 60-minute measurements; in 6 studies the peak cortisol appeared best; and in 4 studies, 30-minute testing appeared best. In our analyses of standard-dose testing, we used 30-minute serum cortisol values and used peak cortisol in the studies where 30-minute values were not available (7, 15, 17).

Combining results from 10 studies of standard-dose stimulation (346 subjects), 30-minute cortisol values of less than 16 μg/dl (440 nmol/l) best predicted a normal reference test result (ruling out HPAI). Intermediate values—16 to 30 μg/dl—were diagnostically indeterminate. The area under the ROC curve for these categorized test results was 0.82 (95% CI: 0.78–0.86).

In paired analyses of the standard-dose stimulation test and the basal cortisol test (9 studies, 302 subjects), diagnostic discrimination was similar: area under the ROC curve was 0.79 for the SDCT and 0.80 for the basal cortisol test (P=0.45).

**Low-dose corticotropin test**

After low-dose corticotropin stimulation, 30-minute cortisol measurements had superior test characteristics in 3 studies (18, 19, 21), while 2 studies found 20-minute values to be best (although not statistically different from 30-minute values) (20, 22). In our analyses of LDCT, we used 30-minute cortisol values (5, 11, 14, 16, 18-22), or, if not available, then the 20-minute (10) or peak values (11).

A meta-analysis of the 11 studies (589 subjects) using the LDCT found that values less than 16 μg/dl (440 nmol/l) best predicted HPAI, while values greater than 22 μg/dl (600 nmol/l) best predicted a normal reference test result. The area under the ROC curve for these diagnostic thresholds was 0.94 (95% CI: 0.90–0.94).

In paired analyses (10 studies, 545 subjects), the LDCT was superior to the basal cortisol test in overall diagnostic discrimination: area under the ROC curve 0.80 for basal cortisol and 0.92 for LDCT (P=0.01).

**Comparison of low-dose and standard-dose corticotropin tests**

In the 7 studies with paired 30-minute cortisol data for both tests (254 subjects), the LDCT had a larger area under the ROC curve compared to the SDCT: 0.94 vs. 0.85 (P < 0.001), after adjusting for type of cortisol assay (Figure 2C). Excluding the 8
Among subjects with indeterminate basal cortisol values (5 to 13 μg/dl, or 138 to 365 nmol/l) (paired data from 6 studies), the LDCT was more discriminating than the SDCT in diagnosing HPAI: area under the ROC curve 0.94 for low-dose and 0.75 for standard-dose test (P < 0.001).

**Basal cortisol test**

In a meta-analysis of 12 studies (635 subjects), a basal cortisol less than 5 μg/dl (138 nmol/l) best predicted HPAI, while values greater than 13 μg/dl (365 nmol/l) best predicted normal HPA axis testing. The area under the ROC curve was 0.79 (95% CI: 0.75–0.82).

**Discussion**

This meta-analysis demonstrates that the low-dose corticotropin stimulation test is superior to the standard-dose test in diagnosing hypothalamic-pituitary adrenal insufficiency. Because all study subjects were ambulatory and presumably had normal sleep–wake cycles, these findings may not generalize to hospital settings or patients with acute illnesses.

Our results differ from the meta-analysis of Dorin and colleagues (12), who reported similar operating characteristics for the LDCT and SDCT. There are 3 possible reasons for the discrepant results. First, the meta-analyses differed in the studies that were included. All studies included in our meta-analysis, except three (5, 16, 17), were also included in the meta-analysis by Dorin and colleagues. The three exceptions were published after Dorin’s literature search. Dorin also included 12 studies that we decided not to include for the following reasons: 5 studies (published before 1990) tested cortisol levels used the fluorescence polarization assay, which is less cortisol specific than more modern assay methods (6) and has never been used to evaluate LDCT; 2 studies evaluated subjects in the early post-operative period; and 5 studies (3 with paired comparisons of LDCT and SDCT) were considered eligible for inclusion in our meta-analysis, but we were unable to obtain patient-level data. Among the 3 eligible studies with paired comparisons that were not included, one did not recruit consecutive patients (13); the second (14) reported superiority of SDCT, but no data was provided regarding statistical significance (their study also included subjects with glucocorticoid-induced HPAI), and the third (15) demonstrated results favoring LDCT, however the results were not statistically significant.

A second reason is that the meta-analyses differed in the study subjects that were included. Whereas the Dorin meta-analysis used summary data reported in the published articles, we obtained patient-level data from study investigators and could therefore apply patient-level eligibility criteria. For example, subjects were not eligible in our meta-analysis if they were healthy control subjects without any suspicion of HPAI (in order to reduce the risk of spectrum bias (4)), or if they had pituitary surgery within 4 weeks of testing.

A third reason for the discrepancy is the difference in analytic methods. Because we had access to patient-level data, we were able to adjust for cortisol assay and sample size and also able to reanalyze data to define two (rule-in and rule-out) cortisol thresholds per test, instead of relying on the single cortisol threshold available in published reports.

In the unadjusted analysis (Figure 2A), the low-dose and standard-dose tests perform similarly—the area under the ROC curve is slightly better for the LDCT, but the difference is probably clinically unimportant. However, when the analysis is adjusted for study size and cortisol assay, the superiority of the LDCT is more
dramatic (Figure 2C) and likely to be clinically relevant. Using summary cortisol thresholds, regardless of cortisol assay method, we have found that the results of SDCT and LDCT performance comparisons were not sensitive to cortisol assay type (Figure 2B). Nevertheless, there may be clinical settings where SDCT is more appropriate to diagnose HPAI, especially if the quality of the low-dose testing protocol cannot be assured.

Based on our findings, we suggest a 3-step approach for evaluating patients with possible HPAI (Figure 3). The first step is measuring a morning basal cortisol. If the results are not convincingly normal or abnormal (basal cortisol level falls in indeterminate range of 5-13 mcg/dl), then the second step is performing a low-dose corticotropin stimulation test. If this test is indeterminate and there are no contraindications to integrated HPA axis testing, we suggest the third step of an insulin hypoglycemia test or metyrapone test. This three-step approach will accurately diagnose the majority of patients, but because it is not perfect, there will still be an important role for clinical judgment, especially regarding use of glucocorticoid supplementation during extreme stress. For convenience, in appropriate clinical circumstances, the first and second steps (basal and LDCT or SDCT) can be done at the same clinical visit to reduce the number of visits for testing.

Gleeson and colleagues (16) published a retrospective study of 10 years of observational data on patients undergoing SDCT. The clinical diagnosis of HPAI was ascertained by an unblinded assessment of the clinical course depicted in medical records. Although the authors report a 97% negative predictive value for SDCT, the data does not allow calculations of sensitivity, specificity or accuracy. The impressive negative predictive value might be partly explained by a low prevalence of newly diagnosed HPAI (19%), which is lower than in most studies included in our meta-analysis. The lower prevalence might be a result of reference test bias, as the clinical diagnosis of milder HPAI might be easily missed without the aid of integrated HPA axis testing. In addition, there is a significant risk of selection bias because 38 percent of patients were excluded from analysis.

Agha and colleagues (28) reported that only 3 to 8 percent of patients with an intermediate response to SDCT (30 minute cortisol, 18.5–23 μg/dl or 510–635 nmol/l), developed signs of adrenal insufficiency over an average follow-up of 4.2 years. The implication is that a clearly negative standard-dose test result (>23 μg/dl or > 635 nmol/l) should perform even better than the intermediate results and effectively rule-out HPAI. However, no data are provided regarding patients with either positive or negative standard-dose test results, and there is no assessment of the overall prevalence of adrenal insufficiency in the entire test population. In fact, for certain cortisol assay methods (5), a 30 minute cortisol value in the range studied by Agha (18.5–23 μg/dl or 510–635 nmol/l) would be classified as a negative test, with a low likelihood of HPAI (Table 1). Unfortunately, the data from Agha’s study do not allow the calculation of test characteristics for positive or negative SDCT results.

Due to the lack of cortisol assay standardization and other reasons for measurement variability, the error in measuring cortisol can be up to 6 μg/dl (165 nmol/l), thus caution is advised when making clinical decisions based on cortisol values close to threshold values. In addition to high variability in the cortisol thresholds, another concern is that a test with a low likelihood of HPAI does not exclude the possibility of future HPAI, especially if there is progression of hypothalamic-pituitary disease or radiation therapy. Therefore, longitudinal assessments may be necessary.

The LDCT has not been validated in patients with acute illnesses, abnormal sleep–wake
cycles, or acute hypothalamic–pituitary disorders (e.g., within one month of pituitary surgery). In addition, all studies of the LDCT that were included in this meta-analysis were conducted in the morning with the patients fasting. Afternoon cortisol values tend to be lower by 1 to 1.5 μg/dl (28-58 nmol/l) (13, 29), and the effect of eating or drinking is uncertain. We also have no information on how the LDCT would perform among patients with low serum protein levels, as circulating cortisol is highly protein bound.

There are several technical details to performing a low-dose test that must be rigorously addressed to avoid false-positive test results (falsely low 30-minute stimulated cortisol value). Currently, there are two acceptable corticotropin analogs that can be used—cosyntropin (Cortrosyn, Amphastar Pharmaceuticals) or tetracosactrin (Synacthen, Novartis Pharma)—supplied in vials containing 250 micrograms of powder. Preparing the 1 microgram dose requires a several-step process of first reconstituting with 250 ml of normal saline and then using a 1 ml aliquot (1 microgram) for intravenous injection. There are additional steps for minimizing adherence of the medication to plastic tubing (17). In addition, the timing of cortisol sampling after low-dose corticotropin administration is very important (we recommend collecting the blood sample 20-30 minutes after corticotropin analog administration), as later sampling may result in false-positive results (18). Thus, low-dose testing should be performed only by personnel knowledgeable of the multiple steps required for preparation and administration. If the quality of administering a 1 microgram dose is suspect, then we recommend using the standard dose of 250 micrograms (reconstituted with 1 ml of sterile diluent) and measuring serum cortisol 30 minutes after intravenous injection. A result less than 16 μg/dl (440 nmol/l)—which is the same threshold used for LDCT—strongly suggests HPAI. However, with standard-dose testing, the 30-minute cortisol value must be greater than 30 μg/dl (833 nmol/l) to be reasonably confident in ruling out HPAI.

Research has suggested (32, 33) that dehydroepiandrosterone sulphate (DHEA-S) blood levels might also help with assessing the HPA axis, particularly when the results of the LDCT are close to either of the two threshold values.

In establishing the diagnosis of HPAI, the accepted reference standard is an abnormal insulin tolerance test or metyrapone test. Both tests, however, can be unreliable. The average intra-subject variability in peak cortisol response to insulin-induced hypoglycemia is 8 to 12 percent (19), but in men with hypopituitarism it can vary by 42 percent (20). Healthy control subjects have been known to “fail” this test. Neither of these reference tests has been validated by assessing predictive accuracy, that is, the ability to predict adrenal crisis.

A limitation of this meta-analysis is that we could not include data of four studies that had published paired results of LDCT and SDCT, because we were unable to obtain the patient-level data (24-26, 32). One study favored LDCT (15), and another study (14) apparently favored SDCT, however the latter study had an incomplete ROC to estimate the magnitude of the difference. The third study (21) found no difference in performance of the two tests. The fourth study (13) used non-consecutive patients for testing and is therefore not pertinent to our meta-analysis.

In conclusion, the performance of the low-dose corticotropin stimulation test was superior to the standard-dose test for evaluating hypothalamic-pituitary adrenal insufficiency. However LDCT must be used by personnel knowledgeable of the multiple steps required for preparation and administration. We describe a three-step testing strategy that proposes use of basal cortisol and the low-dose stimulation test before proceeding to one of the reference
tests for hypothalamic-pituitary adrenal insufficiency.

REFERENCES


22. **Abdu TA, Elhadd TA, Neary R, Clayton RN** 1999 Comparison of the low dose short synacthen test (1 microg), the conventional dose short synacthen test (250 microg), and the insulin tolerance test for assessment of the hypothalamo-pituitary-adrenal axis in patients with pituitary disease. J Clin Endocrinol Metab 84:838-43


Table and Figure legends

**Table 1.** Study characteristics.

**Figure 1.** Receiver operator characteristic (ROC) curves (A. paired and B. unpaired data) for low-dose and standard-dose corticotropin stimulation tests (30-minute cortisol levels) in diagnosing hypothalamic-pituitary adrenal insufficiency.

**Figure 2.** Receiver operator characteristic (ROC) curves for low-dose and standard-dose corticotropin stimulation tests (30-minute cortisol levels) in diagnosing hypothalamic-pituitary adrenal insufficiency: (A) unadjusted for study size and cortisol assay; (B) after categorizing cortisol values, but without other adjustments (arrows represent mean cortisol values across the studies; in $\mu$g/dl (multiply by 27.57 to convert tonmol/l); (C) after categorizing cortisol values and adjusting for cortisol assay.

**Figure 3.** Optimal testing strategy for evaluating patients with possible hypothalamic-pituitary adrenal insufficiency.
### Table 1.

<table>
<thead>
<tr>
<th>Study</th>
<th>Basal cortisol (early AM)</th>
<th>Standard-dose stimulated cortisol (30 minute/peak)</th>
<th>Low-dose stimulated cortisol (20-30 minute)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n HPAI / n total</td>
<td>HPAI† (μg/dl)*</td>
<td>No HPAI‡ (μg/dl)*</td>
</tr>
<tr>
<td>Abdu (22)</td>
<td>13/42</td>
<td>≤3.2</td>
<td>≥12.2</td>
</tr>
<tr>
<td>Courtney (23)</td>
<td>11/41</td>
<td>≤3.0</td>
<td>≥11.4</td>
</tr>
<tr>
<td>Gonc (24)</td>
<td>11/29</td>
<td>≤5.7</td>
<td>≥9.0</td>
</tr>
<tr>
<td>Maghnie (19)</td>
<td>14/24</td>
<td>≤9.1</td>
<td>≥15.0</td>
</tr>
<tr>
<td>Mayenknecht (9)</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Talwar (11)</td>
<td>13/24</td>
<td>≤10.0</td>
<td>≥13.5</td>
</tr>
<tr>
<td>Tordjman (21)</td>
<td>19/62</td>
<td>≤4.8</td>
<td>≥17.9</td>
</tr>
</tbody>
</table>

**Studies with unpaired data**

| Ambrosi (25)   | 12/57           | ≤3.6          | ≥9.3           | n/a             | n/a          | n/a            | 12/57           | ≤16.6          | ≥26.8          |
| Ammari (26)    | 17/30           | ≤7.4          | ≥14.7          | 17/30           | ≤17.0         | ≥33.9          | n/a             | n/a          |
| Choi (27)      | 36/72           | ≤4.8          | ≥13.3          | n/a             | n/a          | n/a            | 36/72           | ≤15.3          | ≥18.9          |
| Kane (5)       | 9/22            | ≤4.4          | ≥8.6           | 9/22            | ≤12.2         | ≥15.7          | n/a             | n/a          |
| Rose (10)      | 42/158          | ≤3.2          | ≥12.5          | 14/38           | ≤16.0         | ≥39.0          | 28/120          | ≤17.5          | ≥20.5          |
| Soule (28)     | 13/74           | ≤3.5          | ≥18.0          | n/a             | n/a          | n/a            | 13/74           | ≤17.0          | ≥24.5          |

**Mean**

<table>
<thead>
<tr>
<th>33% (95% CI)</th>
<th>210/635</th>
<th>≤5§</th>
<th>13§</th>
<th>40%</th>
<th>140/346</th>
<th>16§</th>
<th>33%</th>
<th>193/586</th>
<th>≤16§</th>
<th>22§</th>
</tr>
</thead>
<tbody>
<tr>
<td>201.7 (4.7-5.3)</td>
<td>(12.9-13.6)</td>
<td>(15.2-16.4)</td>
<td>(29.9-32.3)</td>
<td>15.2-16.0)</td>
<td>(20.9-21.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

HPAI – hypothalamic-pituitary adrenal insufficiency
n/a – no data available

* Serum cortisol in μg/dl, to convert to nmol/l multiply by 27.56.
† Threshold below which cortisol values had a LR > 9 (rule-in threshold)
‡ Threshold above which cortisol values had a LR < 0.15 (rule-out threshold)
§ Mean, weighted by size of the studies.
¶ SDCT and LDCT were performed on different subsets of patients (no paired data)
¶¶ Included patients with glucocorticoid-induced HPAI
C

LDCT (AUC 0.94, 95% CI 0.91-0.96)

SDCT (AUC 0.84, 95% CI 0.80-0.88)

P < 0.001
Suspected hypothalamic-pituitary disorder in ambulatory patient with no acute illness

**Basal cortisol***

- **< 5 µg/dl** (138 nmol/l)
  - Probability HPAI > 92%
  - (95% CI 75-99)
- **5–13 µg/dl** (138–365 nmol/l)
  - Probability HPAI 40%
  - (95% CI 31-47)
- **> 13 µg/dl** (365 nmol/l)
  - Probability HPAI < 9%
  - (95% CI 3-18)

**Low-dose corticotropin stimulation test (30-minute cortisol)**

- **< 16 µg/dl** (440 nmol/l)
  - Probability HPAI > 83%
  - (95% CI 67-94)
- **16–22 µg/dl** (440–600 nmol/l)
  - Probability HPAI 33%
  - (95% CI 21-48)
- **> 22 µg/dl** (600 nmol/l)
  - Probability HPAI < 5%
  - (95% CI 1-18)

**Insulin tolerance test or overnight metyrapone test**

- Abnormal
  - Glucocorticoid stress supplementation and replacement
- Normal
  - No glucocorticoid supplementation, unless high clinical suspicion†

---

* A normal DHEA-S blood level makes HPAI highly unlikely (27) and therefore supports the finding of a normal basal cortisol level.
† This algorithm is valid only for chronic HPAI (longer than 4-6 weeks duration). The corticotropin analog stimulation tests have no proven validity in suspected acute or subacute HPAI (less than 4-6 weeks). Insulin tolerance test or overnight metyrapone test can be used to make a diagnosis of acute or subacute HPAI. In addition, even when testing with a corticotropin analog indicates good adrenal responsiveness, the HPA integrity may need reassessment over time to assess for progression of hypothalamic-pituitary dysfunction.