

## ACTH stimulation tests for the diagnosis of adrenal insufficiency: Systematic review and meta-analysis

Naykky Singh Ospina, MD,<sup>1,2,3</sup> Alaa Al Nofal, MD,<sup>1,4</sup> Irina Bancos, MD,<sup>3</sup>, Asma Javed, MD,<sup>4</sup> Benkhadra Khalid, MD,<sup>1,2</sup> Ekta Kapoor, MD,<sup>5</sup>, Aida N Lteif, MD,<sup>4</sup> Neena Natt, MD,<sup>3</sup>, M. Hassan Murad, MD, MPH<sup>1,2</sup>

1. Evidence-Based Practice Research Program, Mayo Clinic, Rochester, MN.; 2. Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, MN, USA; 3. Division of Endocrinology, Diabetes, Metabolism, and Nutrition, Mayo Clinic, Rochester, MN, USA; 4. Division of Pediatric Endocrinology and Metabolism, Mayo Clinic, Rochester, MN, USA; 5. Division of General Internal Medicine, Mayo Clinic, Rochester, MN, USA

**Context:** The diagnosis of adrenal insufficiency is clinically challenging and often requires ACTH stimulation tests.

**Objective:** To determine the diagnostic accuracy of the high (250 mcg) and low (1 mcg) dose ACTH stimulation tests in the diagnosis of adrenal insufficiency.

**Methods (Data sources, Study Selection, Data Extraction):** We searched 6 databases through February 2014. Pairs of independent reviewers selected studies and appraised the risk of bias. Diagnostic association measures were pooled across studies using a bivariate model.

**Data Synthesis:** For secondary adrenal insufficiency, we included 30 studies enrolling 1,209 adults and 228 children. High and low dose ACTH stimulation tests had similar diagnostic accuracy in adults and children using different peak serum cortisol cut-offs. In general, both tests had low sensitivity and high specificity resulting in reasonable likelihood ratios for a positive test (Adults: High dose 9.1, Low dose 5.9; Children: High dose 43.5, Low dose 7.7), but a fairly suboptimal likelihood ratio for a negative test (Adults: High dose 0.39, Low dose 0.19; Children: High dose 0.65, Low dose 0.34). For primary adrenal insufficiency, we included 5 studies enrolling 100 patients. Data were only available to estimate the sensitivity of high dose ACTH stimulation test (92%; 95% CI: 81%–97%).

**Conclusion:** Both high and low dose ACTH stimulation tests had similar diagnostic accuracy. Both tests are adequate to rule in, but not rule out, primary and secondary adrenal insufficiency. Our confidence in these estimates is low to moderate due to the likely risk of bias, heterogeneity and imprecision.

**A**drenal insufficiency is a life-threatening disorder characterized by failure of adrenal cortisol production either from adrenal disease (primary adrenal insufficiency, PAI) or deficiency of adrenocorticotropic hormone (ACTH) (secondary adrenal insufficiency, SAI) (1, 2). Prompt diagnosis is important since adequate hormonal replacement therapy is lifesaving (1, 3–5). Even with early diagnosis and institution of therapy, patients with the di-

agnosis of adrenal insufficiency have higher mortality (6, 7), decreased quality of life (QOL) (8, 9) and increased risk of adrenal crisis (10, 11).

Adrenal insufficiency may present with nonspecific symptoms (eg, fatigue, weight loss, nausea, or/and loss of appetite) resulting in a potential delay in diagnosis. In a cross-sectional study of 216 patients with both primary and secondary adrenal insufficiency, 47% had symptoms

for more than 1 year before diagnosis, and 20% had symptoms for more than 5 years before diagnosis. The correct diagnosis was established during the initial medical encounter in only 15% of patients (12).

Once adrenal insufficiency is suspected, biochemical testing is required to confirm the diagnosis (1). The initial step in evaluation is the measurement of baseline morning serum cortisol and an ACTH stimulation test. The insulin hypoglycemia test (insulin tolerance test) is considered the gold standard for the diagnosis of SAI. This test may not be possible in all situations as it requires medical supervision and can be unsafe in patients with history of seizures, cardiac disease, or the elderly (1, 13). The single-dose overnight metyrapone stimulation test is another confirmatory dynamic test that has been used in the past for the diagnosis of adrenal insufficiency. Through its inhibition of 11- $\beta$ -hydroxylase, metyrapone results in decreased cortisol levels with subsequent feedback stimulation of ACTH and accumulation of the pre-enzyme block substrate 11-deoxycortisol. This test has a similar diagnostic performance to the insulin hypoglycemia test and is a potential alternative when there is a contraindication to the insulin hypoglycemia test (13, 14).

The insulin induced hypoglycemia test and the single dose overnight metyrapone tests are expensive, cumbersome and have potential significant side effects compared to the ACTH stimulation tests. These latter tests assess the serum cortisol response to acute ACTH stimulation with either a 250  $\mu$ g dose (high or standard dose) or 1  $\mu$ g dose (low dose) (1, 13).

The objective of this systematic review and meta-analysis was to compare the diagnostic accuracy of the high and low dose ACTH stimulation tests in patients with either primary or secondary adrenal insufficiency.

## Materials and Methods

### Eligibility criteria

Inclusion criteria for eligible studies were predefined in a study protocol. We included observational and randomized studies that assessed the diagnostic accuracy of high and low dose ACTH stimulation tests for the diagnosis of PAI or SAI when compared to a gold standard. In cases of PAI the gold standard included clinical features, serum cortisol, serum ACTH levels and follow up. In SAI, both the insulin tolerance test and metyrapone test were considered gold standards. Exclusion criteria included case series (uncontrolled studies), review studies, and studies that evaluated patients with critical illness; patients with expected secondary adrenal insufficiency due to exogenous steroid use (eg, patients with autoimmune diseases treated with steroids, patients with asthma) or steroid therapy not discontinued before adrenal insufficiency testing (**with no restriction regarding time of discontinuation**).

### Search Strategy

We conducted a comprehensive search of several databases without language restriction from each database's earliest inception to February 28th, 2014. The databases included Ovid Medline In-Process & Other Non-Indexed Citations, Ovid MEDLINE, Ovid EMBASE, Ovid Cochrane Central Register of Controlled Trials, Ovid Cochrane Database of Systematic Reviews, and Scopus. The search strategy was designed and conducted by an experienced librarian with input from the study's principal investigator (MHM). Controlled vocabulary supplemented with keywords was used to search for adrenal insufficiency. The details of the search are available in the supplementary material. Cross referencing with previously published systematic reviews and contacting content experts were also performed to supplement the electronic search.

Working independently and in duplicate, the reviewers screened the available abstracts (NSO, AA, IB, AJ, BK, EK). Articles in full text were then retrieved and were reviewed independently and in duplicate for eligibility. Disagreements between reviewers were resolved by consensus.

### Data extraction for systematic review

Working independently and in duplicate, data from the included studies were extracted using a standardized data extraction sheet, including baseline information about included studies and the number of patients with true positive, true negative, false positive and false negative results. In cases where the required data were not present in the published manuscript, authors were contacted for additional information (4 authors were contacted with response obtained from 1 author).

### Data extraction for quality of study reporting

Critical appraisal of the included studies was performed independently and in duplicate following the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) instrument. This includes the assessment of the risk of bias and applicability in the following domains: patient selection, index test, reference standard, and flow and timing. This tool includes signaling questions to help reviewers assess the risk of bias. One domain of the tool evaluates patient selection and the methods used for enrolling patients (eg, consecutive or random sample) and the appropriateness of exclusion criteria. Another domain evaluates the index test and whether it was interpreted without knowledge of the standard reference. A domain about the reference standard evaluates whether the interpretation of the reference standard was performed without knowing the results of the index test. Finally, the domain of flow and timing focuses on knowing when the reference standard was performed and in how many patients (15, 16). Cases in which the reviewers' assessment of the risk of bias differed were resolved by consensus.

### Statistical Analysis

Diagnostic estimates from included studies were pooled by fitting a two-level mixed logistic regression model with independent binomial distributions for the true positives and true negatives. These distributions were conditional on the sensitivity and specificity in each study. We also used a bivariate normal model for the logit transforms of sensitivity and specificity between studies (17, 18). The analysis was done using STATA version 13, StataCorp., College Station, TX. Heterogeneity between the studies was assessed using the  $I^2$  statistic. We report

sensitivity, specificity, likelihood ratios and diagnostic odds ratios, with 95% confidence intervals (CIs).

## Results

### Search Results

The results of the systematic search are shown in Figure 1. The systematic search identified 1284 potentially rele-

vant references of which 35 studies were included (30 in SAI, five in PAI).

### Risk of Bias

Using the QUADAS-2 instrument, all included studies had moderate risk of bias as shown in Supplemental Figure 1. This conclusion is mainly driven by unclear or inappropriate patient selection and referral bias leading to high prevalence. Otherwise, the studies had low risk of bias in

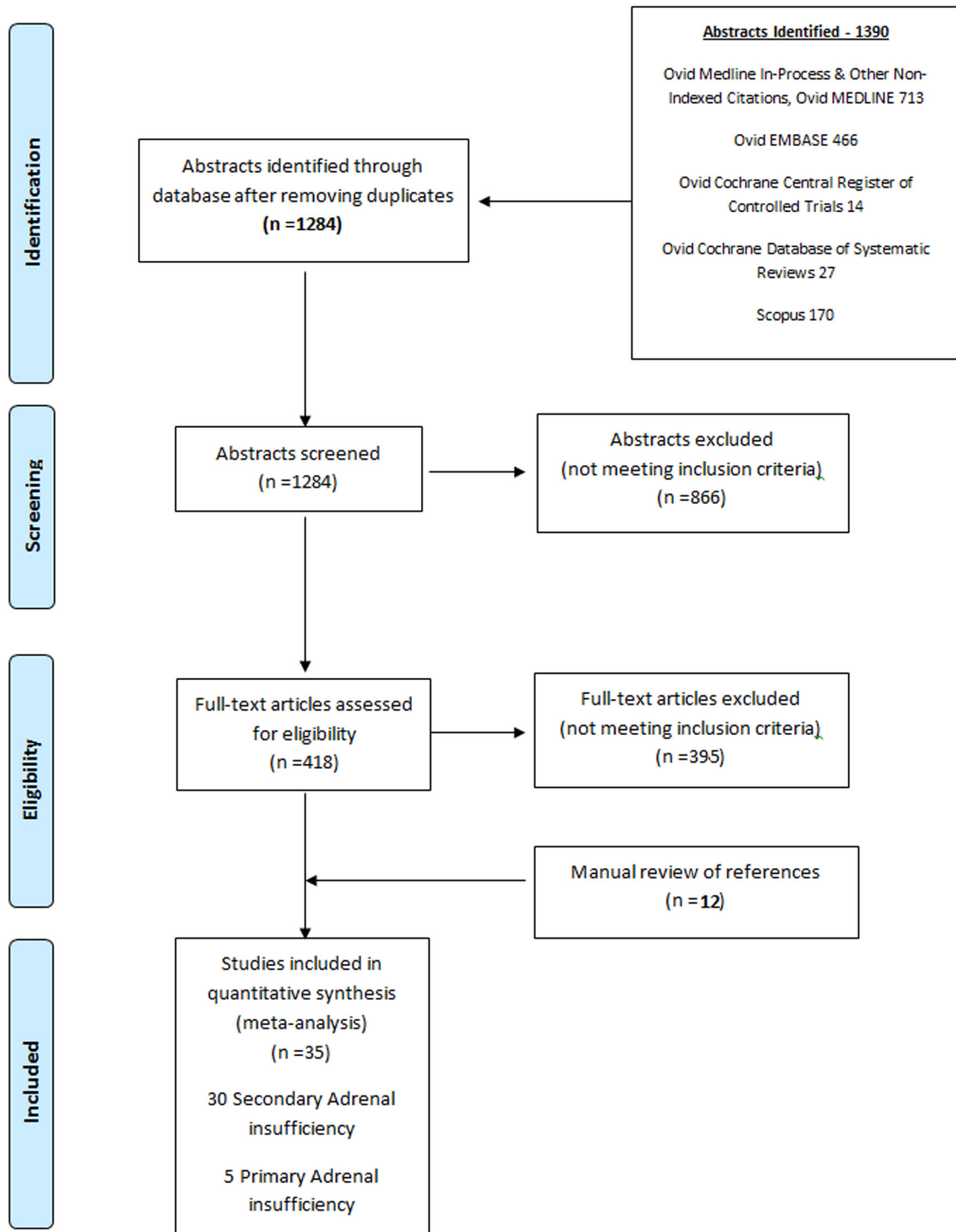


Figure 1. Study selection

the domains of index test, reference standard, and flow and timing.

### Secondary adrenal insufficiency

We identified 30 studies (19–48) assessing the diagnostic performance of the ACTH stimulation test in patients with suspected SAI. Supplemental Table 1 and 2 summarize the characteristics of these studies that enrolled adults and children; respectively. These studies enrolled a total of 1437 patients with a prevalence of SAI of 36% (35% in adults and 38% in children). Most studies administered ACTH intravenously (IV).

We included studies that defined whether the test was positive or negative based on predefined cutoffs that the serum cortisol level had to exceed at any time after ACTH administration, “peak cortisol level”. Other studies used a specific time (30 minutes or 60 minutes) to assess for this predefined serum cortisol value to determine whether the test was positive or negative. The distribution of the included studies in terms of test used and cutoff is as follows:

The overall analysis for the accuracy of high dose ACTH stimulation test in adults included 29 datasets (19, 21–23, 25–29, 31, 33–40, 42, 44–46, 48). Six studies were included in the analysis of high dose ACTH in adults using 500 nmol/L at 30 minutes as cutoff (22, 25, 28, 34, 39, 44); 14 studies used a 500 nmol/L peak serum cortisol value as cutoff (19, 21, 23, 26, 29, 33, 36–38, 40, 42, 45, 46, 48) and eight a serum cortisol cutoff of 550 nmol/L (21, 23, 31, 33, 35, 38, 45, 48).

The overall analysis for the accuracy of the low dose ACTH stimulation test in adults included 19 datasets (19, 20, 23–25, 29, 35, 37, 38, 40, 43, 45, 46, 48). Eleven studies used a 500 nmol/L peak serum cortisol value for the low dose ACTH stimulation test in adults (19, 20, 23, 29, 37, 38, 40, 43, 45, 46, 48); and six a peak serum cortisol level of 550 nmol/L as the cutoff value (23, 35, 38, 43, 45, 48).

The overall analysis of the low dose ACTH stimulation test in children included 5 datasets (30, 32, 41, 47). Three studies evaluated the low dose ACTH stimulation test in children with a peak cortisol of 500 nmol/L (32, 41, 47) and 2 a peak cortisol level of 550 nmol/L as the cutoff (30, 41). The overall analysis of the high dose ACTH stimulation test in children included 4 datasets (30, 41, 47). Two studies evaluated the high dose ACTH in children using a peak of 500 nmol/L (41, 47) and 2 studies with a peak cortisol of 550 nmol/L (30, 41).

### Diagnostic performance in SAI

The diagnostic performance for the high dose and low dose ACTH stimulation test in adults and children according to three different test cutoffs are summarized in Table 1 and 2. Summary receiving operator characteristics curves are in Figures 2 and 3 for low and high dose; respectively. Studies were excluded if patients on long acting steroid were included or due to the lack of a predefined gold standard, reported equivocal results for the gold stan-

**Table 1.** Meta-analysis results – ACTH stimulation tests for the diagnosis of secondary adrenal insufficiency

	Estimate	95% CI
<b>Adult High dose ACTH Stimulation Test</b>		
Sensitivity	0.64	0.52–0.73
Specificity	0.93	0.89–0.96
Likelihood ratio for positive test	9.1	5.7–14.6
Likelihood ratio for negative test	0.39	0.30–0.52
Diagnostic odds ratio	23	13–42
<b>Adult Low dose ACTH Stimulation Test</b>		
Sensitivity	0.83	0.75–0.89
Specificity	0.86	0.78–0.91
Likelihood ratio for positive test	5.9	3.8–8.9
Likelihood ratio for negative test	0.19	0.13–0.29
Diagnostic odds ratio	30	18–50
<b>Children High dose ACTH Stimulation Test</b>		
Sensitivity	0.36	0.10–0.73
Specificity	0.99	0.81–0.99
Likelihood ratio for positive test	43.5	1–1891.2
Likelihood ratio for negative test	0.65	0.36–1.15
Diagnostic odds ratio	67	1–4152
<b>Children Low dose ACTH Stimulation Test</b>		
Sensitivity	0.69	0.28–0.93
Specificity	0.91	0.63–0.98
Likelihood ratio for positive test	7.7	1.3–44.8
Likelihood ratio for negative test	0.34	0.10–1.18
Diagnostic odds ratio	23	2–313

**Table 2.** ACTH stimulation tests for the diagnosis of secondary adrenal insufficiency based on cortisol cut off

Adults									
High Dose ACTH test					Low Dose ACTH Test				
Cortisol Cut off (nmol/liter)	LR +	LR -	Diagnostic Odds Ratio	Number of studies	LR +	LR -	Diagnostic Odds Ratio	Number of studies	P value (for difference)
500 – 30 min	6.3 (2.5, 16)	0.32 (0.20, 0.51)	20 (5, 75)	6	NR	NR	NR	NR	NA
500 - peak	12.4 (6.7, 23.0)	0.48 (0.32, 0.72)	26 (11,60)	14	7.1 (4.3, 11.6)	0.21 (0.13, 0.33)	34 (17, 68)	11	0.631
550 - peak	6.4 (3.4, 12)	0.36 (0.21, 0.61)	18 (8,43)	8	3.8 (1.5, 9.4)	0.23 (0.11, 0.49)	16 (6, 40)	6	0.855
Children									
High Dose ACTH test					Low Dose ACTH Test				
500 - peak	15.96 (2.12–120.04)	0.37 ( 0.01–12.95)	40.67 (1.1–1424.1)	2	18.3 ( 2.04, 164.73)	0.31 (0.5, 1.9)	93.63 ( 14.6–620.1)	3	0.686
550 - peak	6.1 ( 1.09–34.17)	0.78 ( 0.58–1.06)	7.96 (1.2–51.4)	2	4.3 (2.65, 7.06)	0.2 ( 0.02, 1.92)	24.8 (1.73, 356.9)	2	0.494

NR, not reported; NA, not applicable;

**LR + ( likelihood ratio of a positive test); LR – (likelihood ratio of a negative test)**

Heterogeneity values (I<sup>2</sup>):

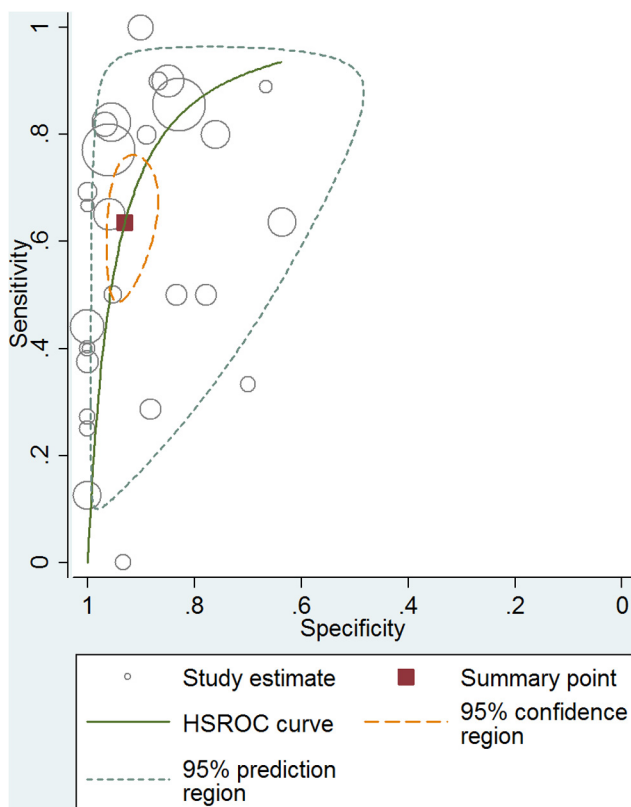
Adults:

- High dose 30 min cut off: 32%; High dose 500 peak cut off:90%; High dose 550 peak cut off: 81%
- Low dose 500 peak cut off: 88%; Low dose 550 peak cut off: 93% Children
- High dose 500 peak cut off: 60%; High dose 550 peak cut off: 0%
- Low dose 500 peak cut off: 0%; Low dose 550 peak cut off: 66%

dard or used a gold standard that was not compatible with the inclusion criteria (14, 49–60).

In general, both tests had low sensitivity and high specificity resulting in reasonable likelihood ratios for a positive test (Adults: High dose 9.1, Low dose 5.9; Children: High dose 43.5, Low dose 7.7), but a fairly suboptimal likelihood ratio (LR) for a negative test (Adults: High dose 0.39, Low dose 0.19; Children: High dose 0.65, Low dose

0.34). Both high dose and low dose tests had moderate accuracy overall (Diagnostic odds ratios ranging from 23 to 67) primarily due to the low sensitivity. However, there was no statistically significant difference between accuracy of the high dose and the low dose tests when comparing diagnostic odds ratios. The analysis was associated with significant heterogeneity; which is common in diagnostic meta-analysis. A summary of the meta-analysis re-



**Figure 2.** Receiver operator curve- High dose ACTH stimulation test for secondary adrenal insufficiency

sults are shown in Table 1 and 2. The receiver operator curve for the high and low dose ACTH stimulation test in adults are found in Figure 2 and 3, respectively (61).

### Primary adrenal insufficiency

We identified 5 studies (62–66) investigating the diagnostic performance of the high dose ACTH stimulation test for the diagnosis of PAI. The characteristics of these studies are summarized in Supplemental Table 3.

### Diagnostic performance in PAI

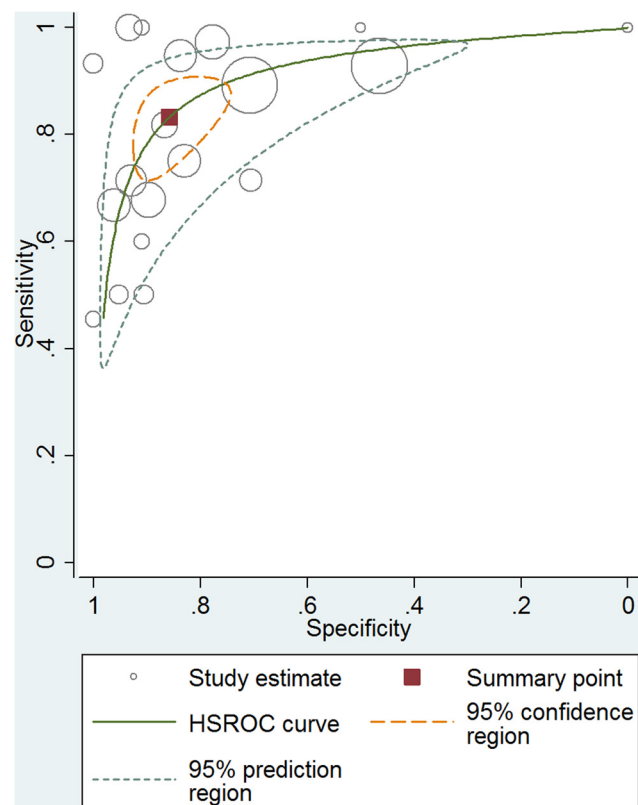
Data were insufficient to estimate specificity, likelihood and diagnostic odds ratios. Only the sensitivity (the rate of a positive test among patients with confirmed PAI) was estimable and was 92%; (95% CI, 81%–97%).

## Discussion

This systematic review and meta-analysis aimed at identifying the diagnostic accuracy of ACTH stimulation test in patients with PAI and SAI. We demonstrated that both high and low dose stimulation tests had similar diagnostic accuracy in SAI. Both tests in general had moderate accuracy due to low sensitivity. Therefore, they are more helpful in ruling in the condition, when positive. However, they are not as reliable in ruling out the condition, when

negative. We demonstrated overall consistency of accuracy measures across different peak cortisol cutoffs and in children and adults. Data in PAI are insufficient to estimate diagnostic accuracy and one can only conclude that the high dose test had high sensitivity of 92%. Many of these PAI patients may have had congenital adrenal hyperplasia, however, the available studies did not provide data to distinguish these patients and allow estimation of diagnostic accuracy measures specific to them. The quality of evidence (confidence in estimates) generated from this analysis is moderate in PAI (due to heterogeneity) and low to moderate in SAI (due to heterogeneity and increased risk of bias).

Two previous systematic reviews attempted to evaluate the diagnostic accuracy of ACTH stimulation tests (67, 68). Dorin and colleagues reported high sensitivity (97.5%) and specificity (96.5%) for the high dose ACTH stimulation test in the diagnosis of primary adrenal insufficiency. However, they included studies in which healthy volunteers and persons without endocrine disease were used as a reference. We did not find any studies that assessed the performance accuracy of the high dose ACTH test in patients with suspected PAI and, therefore, are only able to report the sensitivity based on studies that included patients with known disease. Data from such cohorts exaggerate diagnostic accuracy measures (compared to the



**Figure 3.** Receiver operator curve – Low dose ACTH stimulation test for secondary adrenal insufficiency

optimal study design that includes patients with suspected disease).

Dorin and colleagues noted a positive LR of 11.5 and a negative LR of 0.45 for the high dose ACTH stimulation test (at a set specificity of 95%) for evaluating SAI, which is comparable with our results. We found no statistically significant difference between the diagnostic performance of the high dose vs low dose ACTH stimulation test for the diagnosis of SAI, which is consistent with previous reports (67). Our results are in contrast to the finding of Kazlauskaitė and colleagues (68) who performed a systematic review based on patient level data and reported better performance of the 30 minute cortisol values obtained during low dose ACTH stimulation test when compared to the high dose ACTH stimulation test, even when excluding patients with steroid use from the analysis. Differences in methods (patient level data) and number of included studies (13) should be taken into consideration when comparing the results of this meta-analysis to prior reports.

The limitations of the current available literature are mostly related to significant variability in: (i) the pretest probability of the diagnosis of adrenal insufficiency in the included populations, (ii) the use of different cortisol assays (mostly radio-immunoassays in the included studies), (iii) different cutoff values for the interpretation of the test results (time of measurement and value) in both the index test (ACTH stimulation test) and the gold standard (insulin tolerance test and/or metyrapone test). In addition, technical differences should also be considered in future studies in which the diagnostic performance of the different doses of ACTH stimulation tests are evaluated such as the preparation of the 1 mcg dose of ACTH and the length of tubing used for administration (69). These differences are reflected in the significant level of heterogeneity that we encountered between studies and the wide confidence intervals for some of the estimates.

In addition, the quality assessment of the included studies showed a moderate risk for bias due to patient selection and concern of applicability of the results due to the performance and interpretation of the index test.

Despite these limitations, we believe the results of our study provide interesting insights for the diagnostic performance of ACTH stimulation studies in diagnosing adrenal insufficiency. First, when considering the diagnosis, physicians should have an understanding of the pretest probability of disease. This is important since the presented likelihood ratios of both the high and low dose ACTH stimulation test, suggest that although helpful, these tests are not perfect and can be misleading in some cases. Second, knowledge of the limitations of the test and possible responsible factors (cortisol assay used, time and cut off used for interpretation) should be considered dur-

ing the medical decision making process. The use of gold standard tests might be needed when the results of the ACTH stimulation tests are equivocal or when the test is negative in the setting of high clinical suspicion. For example, in a patient with history of pituitary disease who presents with fatigue and deficiency of other pituitary hormones, most clinicians would be highly suspicious of SAI (high risk for SAI). As shown, in Supplemental Figure 2 panel A, a negative test in that patient would not decrease the likelihood of disease to a level at which most physicians would be comfortable excluding SAI.

**On the other hand, in a patient with fatigue without any signs or risk factors for SAI and an equivocal serum morning cortisol (3–18 mcg/dL) (low risk for SAI) a negative result will significantly decrease the probability of disease (Supplemental Figure 2, panel B).** Unfortunately, there are no validated tools to establish a reliable pretest probability for adrenal insufficiency and this only depends on clinical experience.

A taskforce from the Endocrine Society will provide the clinical context and interpretation to our findings.

## Conclusion

Both high and low dose ACTH stimulation tests have similar diagnostic accuracy. Both tests are adequate to rule in, but not rule out, primary and secondary adrenal insufficiency. Our confidence in these estimates is low-moderate due to the risk of bias of the included studies, heterogeneity and imprecision.

## Acknowledgments

We thank Larry J. Prokop, MLS for his help in designing and executing the search strategy.

Address all correspondence and requests for reprints to: Mohammad Hassan Murad, MD, MPH, Division of Preventive, Occupational and Aerospace Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN 55 905, United States, E-mail: Murad.Mohammad@mayo.edu Phone: (507) 266–1829 Fax: (507) 284–0909.

Disclosure Summary: the authors have nothing to disclose.

**Acknowledgment:** Dr's Singh Ospina and Al Nofal equally contributed to this study.

This work was supported by .

## References

1. Bancos I, Hahner S, Tomlinson J, Arlt W. 2014 Diagnosis and management of adrenal insufficiency. *Lancet Diabetes Endocrinol*
2. Charmandari E, Nicolaidis NC, Chrousos GP. Adrenal insufficiency. *Lancet*. 2014;383:2152–2167.

3. Lovas K, Husebye ES. High prevalence and increasing incidence of Addison's disease in western Norway. *Clin Endocrinol (Oxf)*. 2002; 56:787–791.
4. Regal M, Paramo C, Sierra SM, Garcia-Mayor RV. Prevalence and incidence of hypopituitarism in an adult Caucasian population in northwestern Spain. *Clin Endocrinol (Oxf)*. 2001;55:735–740.
5. Arlt W, Allolio B. Adrenal insufficiency. *Lancet*. 2003;361:1881–1893.
6. Bensing S, Brandt L, Tabaroj F, Sjoberg O, Nilsson B, Ekblom A, Blomqvist P, Kampe O. Increased death risk and altered cancer incidence pattern in patients with isolated or combined autoimmune primary adrenocortical insufficiency. *Clin Endocrinol (Oxf)*. 2008; 69:697–704.
7. Burman D, Morrison GD. Too Many Infant Deaths. *Br Med J*. 1963;2:1419–1420.
8. Benson S, Neumann P, Unger N, Schedlowski M, Mann K, Elsenbruch S, Petersenn S. Effects of standard glucocorticoid replacement therapies on subjective well-being: a randomized, double-blind, crossover study in patients with secondary adrenal insufficiency. *Eur J Endocrinol*. 2012;167:679–685.
9. Hahner S, Loeffler M, Fassnacht M, Weismann D, Koschker AC, Quinkler M, Decker O, Arlt W, Allolio B. Impaired subjective health status in 256 patients with adrenal insufficiency on standard therapy based on cross-sectional analysis. *J Clin Endocrinol Metab*. 2007; 92:3912–3922.
10. Hahner S, Loeffler M, Bleicken B, Drechsler C, Milovanovic D, Fassnacht M, Ventz M, Quinkler M, Allolio B. Epidemiology of adrenal crisis in chronic adrenal insufficiency: the need for new prevention strategies. *Eur J Endocrinol*. 2010;162:597–602.
11. White K, Arlt W. Adrenal crisis in treated Addison's disease: a predictable but under-managed event. *Eur J Endocrinol*. 2010;162: 115–120.
12. Bleicken B, Hahner S, Ventz M, Quinkler M. Delayed diagnosis of adrenal insufficiency is common: a cross-sectional study in 216 patients. *Am J Med Sci*. 2010;339:525–531.
13. de Miguel Novoa P, Vela ET, Garcia NP, Rodriguez MM, Guerras IS, Martinez de Salinas Santamaria Mde L, Maso AA. Guidelines for the diagnosis and treatment of adrenal insufficiency in the adult. *Endocrinol Nutr* 61 Suppl. 2014;1:1–35.
14. Fiad TM, Kirby JM, Cunningham SK, McKenna TJ. The overnight single-dose metyrapone test is a simple and reliable index of the hypothalamic-pituitary-adrenal axis. *Clin Endocrinol (Oxf)*. 1994; 40:603–609.
15. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155:529–536.
16. Swiglo BA, Murad MH, Schunemann HJ, Kunz R, Vigersky RA, Guyatt GH, Montori VM. A case for clarity, consistency, and helpfulness: state-of-the-art clinical practice guidelines in endocrinology using the grading of recommendations, assessment, development, and evaluation system. *J Clin Endocrinol Metab*. 2008;93:666–673.
17. Harbord R. 2008 Stata module for meta-analysis of diagnostic accuracy. In: Statistical Software Components. April 2008 ed: Boston College Department of Economics
18. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics*. 2007;8:239–251.
19. Abdu TA, Elhadd TA, Neary R, Clayton RN. 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*. 1999;84:838–843.
20. Ambrosi B, Barbetta L, Re T, Passini E, Faglia G. The one microgram adrenocorticotropin test in the assessment of hypothalamic-pituitary-adrenal function. *Eur J Endocrinol*. 1998;139:575–579.
21. Ammari F, Issa BG, Millward E, Scanlon MF. A comparison between short ACTH and insulin stress tests for assessing hypothalamo-pituitary-adrenal function. *Clin Endocrinol (Oxf)*. 1996; 44:473–476.
22. Bangar V, Clayton RN. How reliable is the short synacthen test for the investigation of the hypothalamic-pituitary-adrenal axis? *Eur J Endocrinol*. 1998;139:580–583.
23. Cho HY, Kim JH, Kim SW, Shin CS, Park KS, Jang HC, Kim SY. Different cut-off values of the insulin tolerance test, the high-dose short Synacthen test (250 mug) and the low-dose short Synacthen test (1 mug) in assessing central adrenal insufficiency. *Clin Endocrinol (Oxf)*. 2014;81:77–84.
24. Choi CH, Tiu SC, Shek CC, Choi KL, Chan FK, Kong PS. Use of the low-dose corticotropin stimulation test for the diagnosis of secondary adrenocortical insufficiency. *Hong Kong Med J*. 2002;8:427–434.
25. Courtney CH, McAllister AS, Bell PM, McCance DR, Leslie H, Sheridan B, Atkinson AB. Low- and standard-dose corticotropin and insulin hypoglycemia testing in the assessment of hypothalamic-pituitary-adrenal function after pituitary surgery. *J Clin Endocrinol Metab*. 2004;89:1712–1717.
26. Cunningham SK, Moore A, McKenna TJ. Normal cortisol response to corticotropin in patients with secondary adrenal failure. *Arch Intern Med*. 1983;143:2276–2279.
27. Deutschbein T, Unger N, Mann K, Petersenn S. Diagnosis of secondary adrenal insufficiency: unstimulated early morning cortisol in saliva and serum in comparison with the insulin tolerance test. *Horm Metab Res*. 2009;41:834–839.
28. Ferrante E, Morelli V, Giavoli C, Mantovani G, Verrua E, Sala E, Malciodi E, Bergamaschi S, Profka E, Cairoli E, Palmieri S, Chiodini I, Lania AG, Spada A, Peccoz PB. Is the 250 mug ACTH test a useful tool for the diagnosis of central hypoadrenalism in adult patients with pituitary disorders? *Hormones (Athens)*. 2012;11:428–435.
29. Giordano R, Picu A, Bonelli L, Balbo M, Berardelli R, Marinazzo E, Corneli G, Ghigo E, Arvat E. Hypothalamus-pituitary-adrenal axis evaluation in patients with hypothalamo-pituitary disorders: comparison of different provocative tests. *Clin Endocrinol (Oxf)*. 2008; 68:935–941.
30. Gonc EN, Kandemir N, Kinik ST. Significance of low-dose and standard-dose ACTH tests compared to overnight metyrapone test in the diagnosis of adrenal insufficiency in childhood. *Horm Res*. 2003;60:191–197.
31. Jackson RS, Carter GD, Wise PH, Alagband-Zadeh J. 1994 Comparison of paired short Synacthen and insulin tolerance tests soon after pituitary surgery. *Ann Clin Biochem* 31 ( Pt 1):46–49.
32. Kamrath C, Boehles H. The low-dose ACTH test does not identify mild insufficiency of the hypothalamic-pituitary-adrenal axis in children with inadequate stress response. *J Pediatr Endocrinol Metab*. 2010;23:1097–1104.
33. Kehlet H, Blichert-Toft M, Lindholm J, Rasmussen P. Short ACTH test in assessing hypothalamic-pituitary-adrenocortical function. *Br Med J*. 1976;1:249–251.
34. Lindholm J, Kehlet H, Blichert-Toft M, Dinesen B, Riishede J. Reliability of the 30-minute ACTH test in assessing hypothalamic-pituitary-adrenal function. *J Clin Endocrinol Metab*. 1978;47:272–274.
35. Maghnie M, Uga E, Temporini F, Di Iorgi N, Secco A, Tinelli C, Papalia A, Casini MR, Loche S. Evaluation of adrenal function in patients with growth hormone deficiency and hypothalamic-pituitary disorders: comparison between insulin-induced hypoglycemia, low-dose ACTH, standard ACTH and CRH stimulation tests. *Eur J Endocrinol*. 2005;152:735–741.
36. Mukherjee JJ, de Castro JJ, Kaltsas G, Afshar F, Grossman AB, Wass JA, Besser GM. A comparison of the insulin tolerance/glucagon test with the short ACTH stimulation test in the assessment of the hypothalamo-pituitary-adrenal axis in the early post-operative period after hypophysectomy. *Clin Endocrinol (Oxf)*. 1997;47:51–60.
37. Nasrallah MP, Arafah BM. The value of dehydroepiandrosterone



- sulfate measurements in the assessment of adrenal function. *J Clin Endocrinol Metab.* 2003;88:5293–5298.
38. Nye EJ, Grice JE, Hockings GI, Strakosch CR, Crosbie GV, Walters MM, Torpy DJ, Jackson RV. Adrenocorticotropin stimulation tests in patients with hypothalamic-pituitary disease: low dose, standard high dose and 8-h infusion tests. *Clin Endocrinol (Oxf).* 2001;55:625–633.
  39. Orme SM, Peacey SR, Barth JH, Belchetz PE. Comparison of tests of stress-released cortisol secretion in pituitary disease. *Clin Endocrinol (Oxf).* 1996;45:135–140.
  40. Rasmuson S, Olsson T, Hagg E. A low dose ACTH test to assess the function of the hypothalamic-pituitary-adrenal axis. *Clin Endocrinol (Oxf).* 1996;44:151–156.
  41. Rose SR, Lustig RH, Burstein S, Pitukcheewanont P, Broome DC, Burghen GA. Diagnosis of ACTH deficiency. Comparison of overnight metyrapone test to either low-dose or high-dose ACTH test. *Horm Res.* 1999;52:73–79.
  42. Schmiegelow M, Feldt-Rasmussen U, Rasmussen AK, Lange M, Poulsen HS, Muller J. Assessment of the hypothalamo-pituitary-adrenal axis in patients treated with radiotherapy and chemotherapy for childhood brain tumor. *J Clin Endocrinol Metab.* 2003;88:3149–3154.
  43. Soule S, Van Zyl Smit C, Parolis G, Attenborough S, Peter D, Kinvig S, Kinvig T, Coetzer E. The low dose ACTH stimulation test is less sensitive than the overnight metyrapone test for the diagnosis of secondary hypoadrenalism. *Clin Endocrinol (Oxf).* 2000;53:221–227.
  44. Stewart PM, Corrie J, Seckl JR, Edwards CR, Padfield PL. A rational approach for assessing the hypothalamo-pituitary-adrenal axis. *Lancet.* 1988;1:1208–1210.
  45. Talwar V, Lodha S, Dash RJ. Assessing the hypothalamo-pituitary-adrenocortical axis using physiological doses of adrenocorticotrophic hormone. *Qjm.* 1998;91:285–290.
  46. Tordjman K, Jaffe A, Trostanetsky Y, Greenman Y, Limor R, Stern N. Low-dose (1 microgram) adrenocorticotrophin (ACTH) stimulation as a screening test for impaired hypothalamo-pituitary-adrenal axis function: sensitivity, specificity and accuracy in comparison with the high-dose (250 microgram) test. *Clin Endocrinol (Oxf).* 2000;52:633–640.
  47. Weintrob N, Sprecher E, Josefsberg Z, Weininger C, Aurbach-Klipper Y, Lazard D, Karp M, Pertzalan A. Standard and low-dose short adrenocorticotropin test compared with insulin-induced hypoglycemia for assessment of the hypothalamic-pituitary-adrenal axis in children with idiopathic multiple pituitary hormone deficiencies. *J Clin Endocrinol Metab.* 1998;83:88–92.
  48. Dokmetas HS, Colak R, Kelestimur F, Selcuklu A, Unluhizarci K, Bayram F. A comparison between the 1-microg adrenocorticotropin (ACTH) test, the short ACTH (250 microg) test, and the insulin tolerance test in the assessment of hypothalamo-pituitary-adrenal axis immediately after pituitary surgery. *J Clin Endocrinol Metab.* 2000;85:3713–3719.
  49. Dluhy RG, Himathongkam T, Greenfield M. Rapid ACTH test with plasma aldosterone levels. Improved diagnostic discrimination. *Ann Intern Med.* 1974;80:693–696.
  50. Gandhi PG, Shah NS, Khandelwal AG, Chauhan P, Menon PS. Evaluation of low dose ACTH stimulation test in suspected secondary adrenocortical insufficiency. *J Postgrad Med.* 2002;48:280–282.
  51. Gerritsen RT, Vermes I. 1997 The short Synacthen test: with 1 microgram or 250 micrograms ACTH? *Ann Clin Biochem.* 34 ( Pt 1):115–116.
  52. Hartzband PI, Van Herle AJ, Sorger L, Cope D. Assessment of hypothalamic-pituitary-adrenal (HPA) axis dysfunction: comparison of ACTH stimulation, insulin-hypoglycemia and metyrapone. *J Endocrinol Invest.* 1988;11:769–776.
  53. Hurel SJ, Thompson CJ, Watson MJ, Harris MM, Baylis PH, Kendall-Taylor P. The short Synacthen and insulin stress tests in the assessment of the hypothalamic-pituitary-adrenal axis. *Clin Endocrinol (Oxf).* 1996;44:141–146.
  54. Kong MF, Jeffcoate W. Eighty-six cases of Addison's disease. *Clin Endocrinol (Oxf).* 1994;41:757–761.
  55. Mayenknecht J, Diederich S, Bahr V, Plockinger U, Oelkers W. Comparison of low and high dose corticotropin stimulation tests in patients with pituitary disease. *J Clin Endocrinol Metab.* 1998;83:1558–1562.
  56. Kane KF, Emery P, Sheppard MC, Stewart PM. Assessing the hypothalamo-pituitary-adrenal axis in patients on long-term glucocorticoid therapy: the short synacthen versus the insulin tolerance test. *Qjm.* 1995;88:263–267.
  57. Kehlet H, Binder C. Adrenocortical function and clinical course during and after surgery in unsupplemented glucocorticoid-treated patients. *Br J Anaesth.* 1973;45:1043–1048.
  58. Lindholm J, Kehlet H. Re-evaluation of the clinical value of the 30 min ACTH test in assessing the hypothalamic-pituitary-adrenocortical function. *Clin Endocrinol (Oxf).* 1987;26:53–59.
  59. Shankar RR, Jakacki RI, Haider A, Lee MW, Pescovitz OH. Testing the hypothalamic-pituitary-adrenal axis in survivors of childhood brain and skull-based tumors. *J Clin Endocrinol Metab.* 1997;82:1995–1998.
  60. Suliman AM, Smith TP, Labib M, Fiad TM, McKenna TJ. The low-dose ACTH test does not provide a useful assessment of the hypothalamic-pituitary-adrenal axis in secondary adrenal insufficiency. *Clin Endocrinol (Oxf).* 2002;56:533–539.
  61. Bossuyt P DC, Deeks J, Hyde C, Leeflang M, Scholten R. 2013 Chapter 11: Interpreting results and drawing conclusions. In: Deeks JJ BP, Gatsonis C ed. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy: The Cochrane Collaboration*
  62. Gonzalez-Gonzalez JG, De la Garza-Hernandez NE, Mancillas-Adame LG, Montes-Villarreal J, Villarreal-Perez JZ. A high-sensitivity test in the assessment of adrenocortical insufficiency: 10 microg vs 250 microg cosyntropin dose assessment of adrenocortical insufficiency. *J Endocrinol.* 1998;159:275–280.
  63. Nelson JC, Tindall DJ, Jr. 1978 A comparison of the adrenal responses to hypoglycemia, metyrapone and ACTH. *Am J Med Sci* 275:165–172.
  64. Oelkers W, Diederich S, Bahr V. Diagnosis and therapy surveillance in Addison's disease: rapid adrenocorticotropin (ACTH) test and measurement of plasma ACTH, renin activity, and aldosterone. *J Clin Endocrinol Metab.* 1992;75:259–264.
  65. Soule S. Addison's disease in Africa—a teaching hospital experience. *Clin Endocrinol (Oxf).* 1999;50:115–120.
  66. Speckart PF, Nicoloff JT, Bethune JE. Screening for adrenocortical insufficiency with cosyntropin (synthetic ACTH). *Arch Intern Med.* 1971;128:761–763.
  67. Dorin RI, Qualls CR, Crapo LM. Diagnosis of adrenal insufficiency. *Ann Intern Med.* 2003;139:194–204.
  68. Kazlauskaite R, Evans AT, Villabona CV, Abdu TA, Ambrosi B, Atkinson AB, Choi CH, Clayton RN, Courtney CH, Gonc EN, Maghnie M, Rose SR, Soule SG, Tordjman K. Corticotropin tests for hypothalamic-pituitary-adrenal insufficiency: a metaanalysis. *J Clin Endocrinol Metab.* 2008;93:4245–4253.
  69. Wade M, Baid S, Calis K, Raff H, Sinaii N, Nieman L. Technical details influence the diagnostic accuracy of the 1 microg ACTH stimulation test. *Eur J Endocrinol.* 2010;162:109–113.