

ORIGINAL**30-Minute Delta Cortisol Post-ACTH Stimulation Test and Proposed Cut-Off Levels for Adrenal Insufficiency Diagnosis**

Worapaka Manosroi¹, Pichitchai Atthakomol², Supawan Buranapin¹, Nipawan Waisayanand¹, Mattabhorn Phimpilai¹, and Natapong Kosachunhanun¹

¹Division of Endocrinology, Department of Internal Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand, ²Department of Orthopedics, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

Abstract: The ACTH stimulation test is used to diagnose adrenal insufficiency (AI). This study evaluated the diagnostic performance of serum delta cortisol from ACTH stimulation tests and determined appropriate cut-off levels of serum 30-minute delta cortisol for the diagnosis of AI, allowing a reduction in the number of 60-minute cortisol tests. A 6-year retrospective study in 471 patients was conducted. The performance of the serum delta cortisol in diagnosing AI was assessed using a multivariable logistic regression model and the area under ROC curves (AuROC). Both serum 30-minute and 60-minute delta cortisol demonstrated equally high diagnostic accuracy for AI (AuROC for LDT: 0.91 vs 0.90; HDT: 0.91 vs 0.92, respectively). The 30-minute delta cortisol test was chosen to develop proposed diagnostic cut-off levels due to its simplicity. The proposed lower cut-off level for 30-minute delta cortisol was $\Delta < 1.8 \mu\text{g/dL}$ for both LDT and HDT. The upper cut-off levels were $\Delta > 11.8 \mu\text{g/dL}$ for LDT and $\Delta > 10.5 \mu\text{g/dL}$ for HDT. These cut-off levels yielded high sensitivity and specificity $> 90\%$. The 30-minute serum delta cortisol using the proposed cut-off levels provides diagnostic performance for AI equal to that of the 60-minute test and is more convenient, requires less time, less invasive and is cost-saving. 67:95-101, February, 2020

Keywords: adrenal insufficiency, serum delta cortisol, ACTH stimulation test

INTRODUCTION

Various measures have been employed in the diagnosis of adrenal insufficiency (AI). The most common procedure in clinical practice is the ACTH stimulation test. According to the current guideline of the Endocrine Society, the standard dose is 250 μg of ACTH (HDT) with serum cortisol measured at 30 minutes and/or 60 minutes (1). Peak cortisol levels below 18 $\mu\text{g/dL}$ (500 nmol/L) at 30 and/or 60 minutes (assay-specific) indicate AI. However, in medical institutions where ACTH is in short supply, using low dose (1 μg) ACTH (LDT) instead of 250 μg may be a viable option (1). If the ACTH stimulation test is not accessible, serum morning cortisol of $< 5 \mu\text{g/dL}$ ($< 140 \text{ nmol/L}$) may be employed as a preliminary test to suggest AI (1).

Multiple studies have proposed simple alternative measurements to diagnose AI, e.g., serum basal cortisol level with an upper level of $> 16.3 \mu\text{g/dL}$ (450 nmol/L) and lower level of $< 3.6 \mu\text{g/dL}$ (100 nmol/L) (2). Other studies have proposed new cut-off values for serum morning cortisol as well as suggesting additional diagnostic variables including the indications for ACTH stimulation testing and a history of glucocorticoid use to help identify the occurrence of AI (3). Other studies have suggested performing only a 30-minute cortisol test after ACTH injection (4, 5).

The peak time for measurement of serum cortisol, at 30 or 60 minutes, using LDT and HDT is still inconclusive (6, 7). Performing only a 30- or a 60-minute cortisol test cannot definitively diagnose AI and can potentially result in over-diagnosis of AI. A

previous study reported that 24% of patients would be misclassified as having AI if only the 30-minute cortisol test were performed because they attained peak serum cortisol at 60 minutes, while with only the 60-minute test, only 9% of patients would be wrongly classified as AI (8). Using the difference between serum basal cortisol at 0 minutes (baseline) and at 30 minutes after ACTH (delta cortisol) stimulation instead of the absolute value of cortisol might provide an accurate diagnosis while reducing the quantity and frequency of blood drawn as well as saving time and reducing costs.

The objectives of our study were a) to compare the diagnostic accuracy of 30 and 60 minutes serum cortisol for the diagnosis of AI and b) to determine the upper and lower cut-off levels of serum 30 minute delta cortisol for the diagnosis of AI categorized by the dose of ACTH used in the stimulation tests.

MATERIALS AND METHODS

This 6-year retrospective study during the period of January 2010-January 2016 was conducted at the Endocrine and Metabolism Unit of Chiang Mai University Hospital, Thailand using data retrieved from electronic hospital medical records. The study protocol was approved by the Faculty of Medicine Ethics Committee, Chiang Mai University. The study protocol, inclusion and exclusion criteria have been previously mentioned and has been used in a previous study (9).

Data collected included patient demographics, indications for administration of ACTH stimulation tests, biochemical data obtained within 3 months prior to or following ACTH stimulation tests and the results of ACTH stimulation tests. When more than one test had been performed on a patient, only the first test was included in the study.

Received for publication October 31, 2019; accepted December 26, 2019.

Address correspondence and reprint requests to Worapaka Manosroi, MD, Division of Endocrinology and Metabolism, Faculty of Medicine, Chiang Mai University Hospital, 110 Intrawarorot Road Soi 2, Si Phum, Amphoe Mueang Chiang Mai, Chiang Mai 50200, Thailand.

ACTH stimulation testing protocol

Following the protocol of our institution, patients who had serum 0800 h cortisol between 3-17.9 µg/dL and were therefore suspected of having AI were given ACTH stimulation tests to confirm the AI diagnosis. This group of patients included those with evidence of pituitary or hypothalamic diseases (e.g. adenoma, hypophysitis, empty sella syndrome, infiltrative disease) based on MRI imaging; those with evidence of adrenal gland diseases identified from CT imaging (e.g. post-adrenalectomy, infiltrative disease, adrenal gland metastasis); those with a history of glucocorticoids use or use of other medications adulterated with glucocorticoids; and those having signs and/or symptoms of AI or biochemical suggestion of AI. Only LDT (1 µg) was used during January 2010-March 2014 as there was an ACTH shortage in Thailand; HDT was employed from April 2014-January 2016. Patients receiving glucocorticoids or other traditional medicine suspected of containing glucocorticoids were instructed to withhold those substances at least 24-48 hours before testing. All tests were performed at the outpatient clinic between 0900 h-1300 h by medical nurses. In the test, serum basal cortisol at 0 minute was determined. That was followed by ACTH administered intravenously, either 1 µg or 250 µg, after which serum cortisol at 30 and 60 minutes were determined. The 1 µg ACTH ampules were prepared by pharmacists under sterile conditions. The 250 µg ampules of ACTH were diluted with normal saline and stored at 2-8°C. Ampules were used within 60 days (1, 10). Serum cortisol assay was performed by electrochemiluminescence immunoassay (ECLIA) using Elecsys® (Roche Diagnostic, Laval, Quebec) with the intra- and inter-assay coefficients of variation for serum cortisol of < 10%.

Definitions

AI was defined as a peak serum cortisol level < 18 µg/dL at either 30 or 60 minutes after ACTH stimulation testing (1). Serum morning cortisol was measured from samples specifically drawn at 0800 h; serum basal cortisol samples were drawn any time before ACTH was administered and prior to the dynamic tests. Serum 30- and 60-minute delta cortisol levels were defined as the difference between the value of basal cortisol (0 minute) and cortisol level at 30 and 60 minutes, respectively. A history of traditional medicine use was described as a history of using any herbal or traditional medicine suspected of being adulterated with glucocorticoids. Signs and symptoms of AI included fatigue, weight loss, syncope, intractable nausea/vomiting and orthostatic hypotension which were documented in the patient's medical chart. Biochemical investigation results suggestive of AI were defined as hyponatremia (serum sodium < 135 mEq/L), hypoglycemia (serum glucose < 55 mg/dL), hyperkalemia (serum potassium > 4.5 mEq/L), hypereosinophilia (total eosinophil levels > 1,500/µL of blood) and lymphocytosis (total lymphocytes > 4,000/µL of blood).

Statistical analysis

Data were analyzed by STATA version 15.1 (Stata Corp., College Station, TX, USA). Categorical variables are presented as counts and percentages; continuous variables as means and standard deviation (SDs). For inferential statistics, categorical variables were analyzed by Fisher exact test and continuous variables by t-test or Mann-Whitney U test. The associations between AI and 30- and 60-minute serum delta cortisol levels was analyzed using a univariable and multivariable logistic regression model and are reported as odds ratios (ORs). The areas under the ROC curve (AuROC) of the model were plotted to discriminate diagnostic performance of each value categorized by the ACTH dosage used in the stimulation tests. Sensitivity,

specificity, positive predictive value (PPV), negative predictive value (NPV), likelihood ratio of positive results (LHR+), likelihood ratio of negative results (LHR-) as well as the number of true positives (TP), false negatives (FN), false positives (FP) and true negatives (TN) are reported for each proposed cut-off point. Proposed cut-off levels were set at intervals of 0.9 µg/dL (25 nmol/L). The cut-off value which gave the highest sensitivity and the lowest number of FN results was classified as the appropriate upper cut-off level while those with the highest specificity with the lowest number of FP results were classified as the appropriate lower cut-off level. The statistical significance level was set as two-tailed with a *P*-value < 0.05.

RESULTS*Baseline characteristics*

Baseline characteristics are depicted in Table 1. Of the 471 patients (226 male and 245 females) included in the study who were at risk for AI, 23.8% (n = 112/471) were identified as having AI. The mean age of the AI group was significantly higher than the normal adrenal response group (non-AI group). The most common indication for ACTH stimulation testing was the presence of AI symptoms (33.7%, n = 159/471). Diagnosis of AI was significantly higher in patients for whom a history of exogenous steroid use, post-pituitary surgery or pituitary tumor was the indication for testing. No significant difference in the prevalence of AI was observed between LDT and HDT. Among the other biochemical values measured, significantly lower serum albumin and serum creatinine and a significantly higher incidence of hyperkalemia were observed in patients with AI than those in the non-AI group.

Demographic and biochemical data categorized by LDT and HDT and by peak serum cortisol at 30 and 60 minutes after ACTH stimulation tests are shown in Supplementary Table 1 and 2. The majority of patients in both the LDT and HDT groups showed peak serum cortisol at 60 minutes after ACTH stimulation testing. A significantly higher number of patients with peak serum cortisol at 60 minutes was found in the HDT group than in the LDT group (91.9%, n = 262/285 and 79.1%, n = 147/186, respectively). In the LDT group, if only a 30-minute cortisol level were performed on the patients with peak cortisol at 60 minutes, 6.8% (n = 10/147) of the patients with normal adrenal function would have been incorrectly classified as having AI. In the HDT group, if only 30-minute cortisol levels were acquired, approximately 13.7% (n = 36/262) of the patients with normal adrenal function would have been misclassified as having AI. The box-plot for serum basal, 30- and 60-minute delta cortisol categorized by LDT and HDT is depicted in Supplementary Figure 1.

Serum delta cortisol levels and predictive performance for adrenal insufficiency

Based on multivariable analysis adjusted for age, sex, serum albumin, cholesterol, creatinine and ACTH dose, serum delta cortisol at both 30 and 60 minutes after ACTH stimulation tests demonstrated a significant association with AI. All ORs were less than 1, which suggests the lower the delta cortisol, the higher the likelihood of having AI (Table 2).

Evaluation of diagnostic performance by AuROC found serum delta cortisol at 60 minutes had similar performance with serum delta cortisol at 30 minutes for both LDT and HDT (p = 0.877 and 0.687, respectively). The AuROCs for 30- and 60-minute delta cortisol were higher than 0.8 for both (Figure 1 and Table 3).

Table 1 Demographic and biochemical factors categorized by adrenal insufficiency (n = 112) and normal adrenal response (n = 359)

	Adrenal Insufficiency (n = 112)	Normal Response (n = 359)	P-value
Age (Mean ± SD) (years)	56.6 ± 17.1	48.6 ± 17.1	< 0.001
Gender, N (%)			
- Male	49 (43.8)	177 (49.3)	
- Female	63 (56.2)	182 (50.7)	0.304
Body Weight (mean ± SD) (kg)	58.2 ± 13.3	61.4 ± 16.1	0.058
Underlying Diseases, N (%)			
- Autoimmune Diseases	18 (16.1)	39 (10.9)	0.140
- Diabetes Mellitus	17 (15.2)	50 (13.9)	0.749
- Hypertension	34 (30.4)	75 (20.9)	0.040
- Coronary Artery Disease	11 (9.8)	12 (3.3)	0.006
- Malignancy	2 (1.8)	10 (2.8)	0.558
- Other underlying diseases	83 (74.1)	231 (64.6)	0.060
Indication for Testing, N (%)			
- Exogenous Steroid Use	54 (48.2)	70 (19.5)	< 0.001
o Prednisolone	17 (24.6)	25 (13.8)	0.041
o Dexamethasone	2 (3.5)	2 (1.2)	0.267
o Herb or Traditional Medicine Use	33 (42.3)	39 (21.1)	< 0.001
- Post-Pituitary Surgery	14 (12.5)	79 (22.0)	0.020
- Pituitary Tumor	8 (7.1)	7 (1.9)	0.006
- Other Pituitary Hormonal Deficiencies	28 (25.0)	111 (30.9)	0.230
- Symptoms and Biochemical Suggestive of Adrenal Insufficiency	37 (33.0)	122 (33.9)	0.853
Baseline SBP (mean ± SD) (mmHg)	120.4 ± 22.6	121.4 ± 19.9	0.653
Baseline DBP (mean ± SD) (mmHg)	71.5 ± 15.9	73.7 ± 12.7	0.137
ACTH Stimulation Dose, N (%)			
- 1 µg	38 (33.9)	148 (41.2)	
- 250 µg	74 (66.1)	211 (58.7)	0.168
Serum Morning Cortisol (mean ± SD) (µg/dL)	7.9 ± 2.9	9.6 ± 3.3	< 0.001
Serum Basal Cortisol (mean ± SD) (µg/dL)	6.2 ± 3.2	10.9 ± 5.4	< 0.001
Cortisol at 30 Min (mean ± SD) (µg/dL)	11.7 ± 3.8	23.9 ± 5.9	< 0.001
Cortisol at 60 Min (mean ± SD) (µg/dL)	6.5 ± 3.5	15.7 ± 5.5	< 0.001
30-Min Delta Cortisol (mean ± SD) (µg/dL)	5.5 ± 2.8	13.1 ± 4.8	< 0.001
60-Min Delta Cortisol (mean ± SD) (µg/dL)	6.5 ± 3.5	15.7 ± 5.6	< 0.001
Serum Albumin (mean ± SD) (g/dL)	3.6 ± 0.8	3.9 ± 0.6	< 0.001
Serum Cholesterol (mean ± SD) (mg/dL)	177.9 ± 58.0	185.6 ± 58.0	0.445
Serum Creatinine (mean ± SD) (mg/dL)	1.1 ± 1.0	0.9 ± 0.7	0.038
Hypoglycemia, N (%)	2 (1.8)	12 (3.3)	0.397
Hyponatremia, N (%)	86 (85.1)	264 (88.3)	0.409
Hyperkalemia, N (%)	117 (32.6)	27 (24.1)	0.089
Hyper eosinophilia, N (%)	2 (2.2)	3 (1.1)	0.420
Lymphocytosis, N (%)	6 (6.7)	17 (6.1)	0.857

SBP : Systolic blood pressure
DBP : Diastolic blood pressure

Supplementary Table 1 Demographic data categorized by low dose and high dose ACTH stimulation tests

Characteristics	Low dose ACTH stimulation test (n = 186)	High dose ACTH stimulation test (n = 285)	P-value
Adrenal insufficiency, N (%)	38 (20.4)	74 (25.9)	0.168
Peak cortisol at 30 minutes, N (%)	39 (20.9)	23 (8.1)	
Peak cortisol at 60 minutes, N (%)	147 (79.1)	262 (91.9)	< 0.001
Age (Mean ± SD) (years)	51.5 ± 17.4	49.8 ± 17.4	0.301
Gender, N (%)			
- Male	65 (34.9)	162 (56.6)	
- Female	121 (65.0)	124 (43.4)	< 0.001
Body Weight (mean ± SD) (kg)	60.2 ± 12.2	60.9 ± 15.1	0.638
Underlying Diseases, N (%)			
- Autoimmune Diseases	29 (15.6)	28 (9.8)	0.057
- Diabetes Mellitus	28 (15.1)	39 (13.7)	0.677
- Hypertension	42 (22.6)	68 (23.7)	0.748
- Coronary Artery Disease	13 (6.99)	10 (3.51)	0.087
- Malignancy	3 (1.6)	9 (3.1)	0.304
- Other underlying diseases	144 (77.4)	170 (59.7)	< 0.001
Indication for Testing, N (%)			
- Exogenous Steroid Use	63 (33.9)	61 (21.3)	0.002
- Post-Pituitary Surgery	35 (18.8)	58 (20.3)	0.696
- Pituitary Tumor	3 (1.6)	12 (4.2)	0.118
- Other Pituitary Hormonal Deficiencies	59 (31.7)	80 (27.9)	0.383
- Symptoms of Adrenal Insufficiency	77 (41.4)	82 (28.6)	0.004
Baseline SBP (mean ± SD) (mmHg)	121.6 ± 20.9	121.0 ± 20.4	0.782
Baseline DBP (mean ± SD) (mmHg)	74.1 ± 11.9	72.7 ± 14.5	0.305
Serum Morning Cortisol (mean ± SD) (µg/dL)	9.3 ± 3.5	9.7 ± 3.2	0.531
Serum Basal Cortisol (mean ± SD) (µg/dL)	10.2 ± 6.1	9.5 ± 4.7	0.166
Cortisol at 30 Min (mean ± SD) (µg/dL)	22.6 ± 7.9	20.1 ± 7.3	< 0.001
Cortisol at 60 Min (mean ± SD) (µg/dL)	24.1 ± 8.5	22.7 ± 8.3	0.072
30-Min Delta Cortisol (mean ± SD) (µg/dL)	12.4 ± 5.7	10.6 ± 5.2	< 0.001
60-Min Delta Cortisol (mean ± SD) (µg/dL)	14.3 ± 6.8	13.3 ± 6.3	0.256
Serum Albumin (mean ± SD) (g/dL)	3.7 ± 0.7	3.9 ± 0.6	0.002
Serum Cholesterol (mean ± SD) (mg/dL)	193.4 ± 65.8	181.7 ± 54.1	0.035
Serum Creatinine (mean ± SD) (mg/dL)	0.9 ± 0.7	0.9 ± 0.9	0.758

SBP : Systolic blood pressure
DBP : Diastolic blood pressure

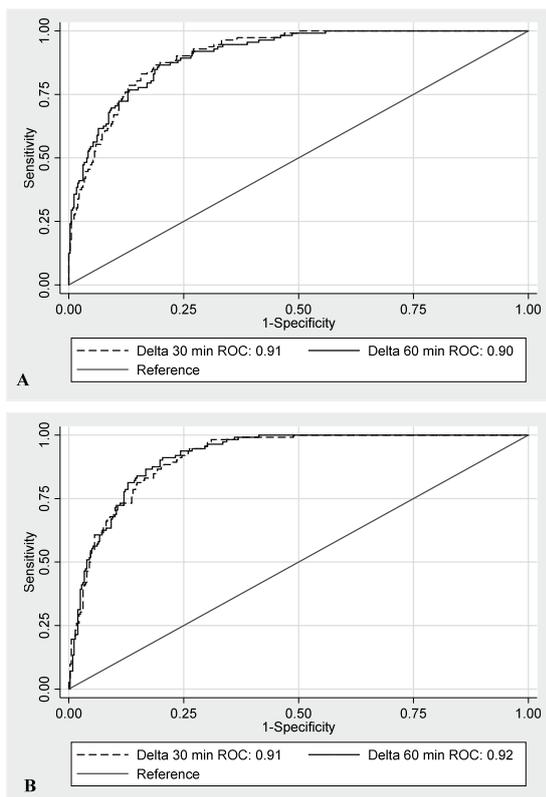
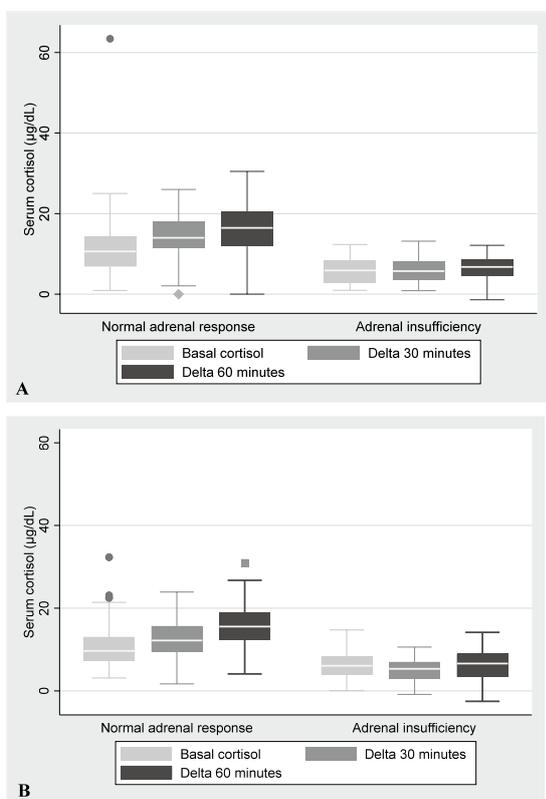


Figure 1 AuROC comparing between serum delta cortisol at 30 and 60 minutes. A) low dose ACTH stimulation test, B) high dose ACTH stimulation test



Supplementary Figure 1 Boxplot of serum basal cortisol, serum delta cortisol at 30 and 60 minutes. A) low dose ACTH stimulation test, B) high dose ACTH stimulation test

Table 2 Univariable and multivariable model for serum delta cortisol at 30 and 60 minutes to predict adrenal insufficiency

Factors	Univariable		Multivariable*	
	Odd ratio (95% CI)	P-value	Odd ratio (95% CI)	P-value
Serum delta cortisol at 30 minutes	0.98 (0.97-0.99)	< 0.001	0.98 (0.97-0.99)	< 0.001
Serum delta cortisol at 60 minutes	0.98 (0.98-0.99)	< 0.001	0.98 (0.98-0.99)	< 0.001

*Adjusted for age, sex, serum albumin, cholesterol, creatinine and ACTH dose

Supplementary Table 2 Demographic data categorized by peak serum cortisol at 30 and 60 minutes after ACTH stimulation tests

Characteristics	Peak cortisol at 30 min (n = 62)	Peak cortisol at 60 min (n = 409)	P-value
Adrenal insufficiency, N (%)	22 (35.5)	90 (22.0)	0.020
Age (Mean ± SD) (years)	51.4 ± 15.7	50.4 ± 17.7	0.641
Gender, N (%)			
- Male	22 (35.5)	204 (49.8)	
- Female	40 (64.5)	205 (50.2)	0.035
Body Weight (mean ± SD) (kg)	63.4 ± 18.7	60.2 ± 14.9	0.147
Underlying Diseases, N (%)			
- Autoimmune Diseases	6 (9.7)	51 (12.5)	0.530
- Diabetes Mellitus	8 (12.9)	59 (14.5)	0.744
- Hypertension	7 (11.3)	102 (25.0)	0.017
- Coronary Artery Disease	2 (3.2)	21 (5.1)	0.514
- Malignancy	0 (0.0)	12 (2.93)	0.172
- Other underlying diseases	46 (74.1)	268 (65.7)	0.185
Indication for Testing, N (%)			
- Exogenous Steroid Use	19 (30.6)	105 (25.7)	0.407
- Post-Pituitary Surgery	12 (19.4)	81 (19.8)	0.934
- Pituitary Tumor	4 (6.4)	11 (2.7)	0.116
- Other Pituitary Hormonal Deficiencies	18 (29.0)	121 (29.6)	0.926
- Symptoms of Adrenal Insufficiency	25 (40.3)	134 (32.8)	0.241
Baseline SBP (mean ± SD) (mmHg)	122.9 ± 19.2	120.9 ± 20.8	0.469
Baseline DBP (mean ± SD) (mmHg)	74.6 ± 15.1	73.0 ± 13.3	0.375
Serum Morning Cortisol (mean ± SD) (µg/dL)	8.7 ± 3.6	9.2 ± 3.2	0.261
Serum Basal Cortisol (mean ± SD) (µg/dL)	9.8 ± 5.8	9.7 ± 5.4	0.930
Serum Albumin (mean ± SD) (g/dL)	3.8 ± 0.6	3.9 ± 0.7	0.757
Serum Cholesterol (mean ± SD) (mg/dL)	201.3 ± 52.7	182.2 ± 58.2	0.034
Serum Creatinine (mean ± SD) (mg/dL)	0.8 ± 0.3	0.9 ± 0.8	0.197

Table 3. Comparison of AuROC between serum delta cortisol at 30 and 60 minutes

Comparison	AuROC	95%CI	P-value
LDT			
Serum delta cortisol at 30 minutes vs Serum delta cortisol at 60 minutes	0.91 vs 0.90	(0.88-0.94) vs (0.88-0.93)	0.877
HDT			
Serum delta cortisol at 30 minutes vs Serum delta cortisol at 60 minutes	0.91 vs 0.92	(0.89-0.94) vs (0.90-0.94)	0.687

LDT : Low dose ACTH stimulation test
HDT : High dose ACTH stimulation test

Proposed cut-off levels for 30-minute serum delta cortisol

As the diagnostic accuracy of delta 30- and 60-minute cortisol was similar, only delta cortisol at 30 minutes was used in the analysis for proposed cut-off levels. Data for each cut-off level categorized by LDT and HDT are depicted in Table 4.

For LDT, the lower cut-off for 30-minute delta cortisol to rule in AI which gave the highest specificity (99.3%) was $\Delta < 1.8 \mu\text{g/dL}$, while the upper cut-off to rule out AI which gave the highest sensitivity (97.4%) was $\Delta > 11.8 \mu\text{g/dL}$. Delta cortisol levels were below the cut-off in 1.6% of cases ($n = 3/186$), while they were above the cut-off in 42.5% of cases ($n = 79/186$). Similarly, for HDT the lower cut-off for 30-minute delta cortisol to rule in AI which gave the highest specificity (99.4%) was $\Delta < 1.8 \mu\text{g/dL}$, while the upper cut-off to rule out AI which had the highest sensitivity (98.6%) was $\Delta > 10.5 \mu\text{g/dL}$. Delta cortisol levels were below the cut-off in 3.2% of cases ($n = 9/285$), while levels were above the cut-off in 49.5% of cases ($n = 141/285$) in HDT. The proposed upper and lower cut-off levels resulted in only one FP and one FN, respectively. Other proposed cut-off levels for LDT and HDT also gave high sensitivity and specificity (around 90%) (Table 4). If the proposed serum delta cortisol cut-off levels were used, 58% ($n = 108/186$) of the LDT and 53% ($n = 151/285$) of the HDT patients would have been appropriately categorized.

When the above mentioned cut-off levels were used in the patients who attained peak serum cortisol at 60 minutes which is the group concerning for AI over-diagnosis, increased sensitivity and specificity were observed. With LDT using the lower cut-off level of $\Delta < 1.8 \mu\text{g/dL}$, the specificity was 100%; the upper cut-off level of $\Delta > 11.8 \mu\text{g/dL}$ also had a sensitivity of 100%. With HDT, the lower cut-off level of $\Delta < 1.8 \mu\text{g/dL}$ revealed a specificity of 99.5%, while the upper cut-off level of $\Delta > 10.5 \mu\text{g/dL}$ gave a similar sensitivity result of 98.6%. If only a 30-minute cortisol levels had been acquired during LDT, $n = 10/147$ or, in HDT,

$n = 36/262$ patients would have been incorrectly classified as having AI. If the upper and lower cut-off levels had been used with both LDT and with HDT, all patients in both groups would have been correctly identified as non-AI.

DISCUSSION

The present study highlights the important finding that serum delta cortisol at both 30 and 60 minutes after ACTH stimulation tests can predict the occurrence of AI with a high degree of diagnostic accuracy. No significant differences in diagnostic accuracy between the two groups were observed. For the health care practitioner, using only 30-minute cortisol to calculate the delta cortisol levels offers the benefits of increased convenience, being less invasiveness and requiring less time while providing high specificity and sensitivity for ruling in and ruling out AI.

Both 30- and 60-minute delta cortisol demonstrated 'excellent' overall diagnostic accuracy with AuROC of between 90-100% (11). The justification for using the "delta" value of serum cortisol after ACTH stimulation tests is a study which demonstrated that using the absolute single-time serum cortisol value, especially at 30 minutes, may increase the number of false positives for AI (8, 12). Thus using delta cortisol levels rather than absolute levels of cortisol may have high performance in AI diagnosis. The use of serum delta cortisol in AI diagnosis has been reported in multiple reports but those results have been inhomogeneous. For example, a study by Struja *et al.* stated that delta cortisol does not offer any added information compared to baseline cortisol alone. They found that delta cortisol was inferior to basal cortisol with AUCs of 0.85 (95% CI : 0.81-0.89) and 0.88 (95% CI : 0.83- 0.92), respectively (2). Delta cortisol has also been proposed for use in critically ill patients. Using serum delta cortisol after HDT in critically-ill patients with a cut-off level of $\Delta < 9 \mu\text{g/dL}$ after peak

Table 4 Proposed cut-off levels and the diagnostic accuracy for 30-minute delta cortisol levels categorized by low dose and high dose ACTH stimulation test

	Level ($\mu\text{g/dL}$)	Sensitivity (95%CI)	Specificity (95%CI)	PPV (%)	NPV (%)	LHR+	LHR-	TP (n)	FN (n)	FP (n)	TN (n)
30 minutes delta cortisol for LDT											
- Lower	$\Delta < 1.8$	5.3 (0.6-17.7)	99.3 (96.3-99.9)	66.7	80.3	7.79	0.95	2	36	1	147
	$\Delta < 2.7$	13.2 (4.4-28.1)	98.0 (94.2-99.6)	62.5	81.5	6.49	0.89	5	33	3	145
	$\Delta < 3.6$	26.3 (13.4-43.1)	97.3 (93.2-99.3)	71.4	83.7	9.74	0.76	10	28	4	144
	$\Delta < 4.5$	31.6 (17.5-48.7)	97.3 (93.2-99.3)	75.0	84.7	11.68	0.70	12	26	4	144
	$\Delta < 5.4$	47.4 (31.0-64.2)	93.2 (87.9-96.7)	64.3	87.3	7.01	0.56	18	20	10	138
- Upper	$\Delta > 11.8$	97.4 (86.2-99.9)	71.6 (63.6-78.7)	46.8	99.1	3.43	0.04	37	1	42	106
	$\Delta > 10.9$	94.7 (82.3-99.4)	77.7 (70.1-84.1)	52.2	98.3	4.25	0.07	36	2	33	115
	$\Delta > 9.9$	94.7 (82.3-99.4)	79.7 (72.3-85.9)	54.5	98.3	4.67	0.07	36	2	30	118
	$\Delta > 9.0$	89.5 (75.2-97.1)	82.4 (75.3-88.4)	56.7	96.8	5.09	0.13	34	4	26	122
30 minutes delta cortisol for HDT											
- Lower	$\Delta < 1.8$	10.8 (4.8-20.2)	99.4 (97.4-99.9)	88.9	76.1	22.81	0.90	8	66	1	210
	$\Delta < 2.7$	20.3 (11.8-31.2)	98.6 (95.9-99.7)	83.3	77.9	14.26	0.81	15	59	3	208
	$\Delta < 3.6$	27.0 (17.4-38.6)	96.7 (93.3-98.7)	74.1	79.1	8.15	0.75	20	54	7	204
	$\Delta < 4.5$	35.1 (24.4-47.1)	94.3 (90.3-97.0)	68.4	80.6	6.18	0.69	26	48	12	199
	$\Delta < 5.4$	52.7 (40.7-64.4)	93.8 (89.7-96.7)	75.0	85.0	8.55	0.50	39	35	13	198
- Upper	$\Delta > 10.5$	98.6 (92.7-99.9)	67.8 (61.0-74.0)	51.8	99.3	3.06	0.02	73	1	68	143
	$\Delta > 9.9$	91.9 (83.2-97.0)	71.1 (64.5-77.1)	52.7	96.2	3.18	0.11	68	6	61	150
	$\Delta > 9.0$	87.8 (78.2-94.3)	78.2 (72.0-83.6)	58.6	94.8	4.03	0.16	65	9	46	165

LDT : Low dose ACTH stimulation test
HDT : High dose ACTH stimulation test

serum cortisol can help diagnose critical illness-related corticosteroid-induced adrenal insufficiency conditions (13). However, this concept can be applied only in a critically ill setting.

Specific patterns in peak serum cortisol were demonstrated in the present study. Although most of the patients attained peak cortisol at 60 minutes, which is in agreement with a study by Cartaya *et al.* (8), we postulated that using only a 30-minute cortisol level could facilitate AI diagnosis as it would be less time-consuming, simpler and less expensive as only two blood samples are required (basal cortisol and 30-minutes serum cortisol). However, using only a 30-minute cortisol level raises concern regarding AI overdiagnosis in 6-14% of patients, particularly in patients who had undergone HDT. In order to reduce the incidence of false positives that could result in inappropriate treatment with corticosteroids and which, in turn, could lead to multiple long-term complications, we proposed upper and lower cut-offs for serum delta cortisol levels. Although one study reported that mean cortisol levels with LDT and HDT at 30 minutes were not different (12), other studies have reported different accuracy levels and time variabilities between LDT and HDT (14, 15). For that reason, this study used different 30-minute delta cortisol cut-off values for LDT and for HDT.

With both LDT and HDT, if the proposed 30-minute delta cortisol had been utilized all patients who formerly fell into the AI over-diagnosed group when only a 30-minute cortisol was employed would have been correctly identified as having no AI. These results suggest that our proposed cut-off levels and utilizing only 30-minute serum cortisol testing can help reduce the incidence of AI over-diagnosis. For LDT in particular, multiple studies have recommended 30-minute cortisol values as the test of choice to evaluate AI (4, 5, 16). However, data on the use of HDT and only 30-minute cortisol are still inconclusive. Only one study has reported benefits from using 30-minute cortisol in HDT to diagnose AI, while other studies have reported no diagnostic value and a very high false positive rate using this value in isolation (17-19). If our upper and lower proposed cut-off levels for 30-minute delta cortisol levels were employed with all the patients included in this study, approximately 58% of the LDT and 53% of the HDT patients would have been appropriately categorized. This would reduce by more than half the need to perform other cortisol tests or to draw specimens at 60 minutes. For those reasons, we recommend performing only a 30-minute cortisol level and calculating delta cortisol levels. If serum 30-minute cortisol levels are undiagnosed or if delta cortisol levels fall between the upper and lower cut-off levels (intermediate range), then 60-minute cortisol may be determines. This proposed measure provides markedly high diagnostic performance for AI and could be easily used in normal clinical practice.

The reason for choosing the upper cut-off level which gave the highest sensitivity to rule-out AI was that AI is a potentially lethal condition if left untreated and thus requires timely diagnosis. There is a need for a tool which provides high sensitivity to minimize the number of false negative AI diagnoses. Likewise, the lower cut-off values to rule in AI were the values which gave the highest specificity with the lowest incidence of false positives. Treating patients with glucocorticoids who may not need these medications may cause harm, e.g., physical and metabolic changes such as osteoporosis.

The present study has multiple strengths. We found that the use of delta cortisol after ACTH stimulation testing yielded excellent accuracy in AI diagnosis. In particular, 30-minute cortisol is beneficial in terms of increased convenience as well as saving time and expense. The proposed cut-off for 30-minute cortisol could make these findings more widely applicable in actual practice. As most of the peaks in serum cortisol after an ACTH stimulation test occur at 60 minutes, the likelihood of having a false

positive diagnosis of AI is high if only 30-minute cortisol level testing is performed. The cut-off levels that we have presented could diminish these diagnostic errors. Confounding variables in the multivariable analysis model were appropriately adjusted for, including serum albumin, cholesterol and creatinine levels which various studies have demonstrated to be associated with serum cortisol interference (20-22). Another strength is the large sample size with an adequate power of > 80% (23).

We acknowledge some limitations in this study. The proposed cut-off levels were based on the Roche diagnostic ECLIA assay for cortisol which has been reported as one of the three most commonly used cortisol assays in United Kingdom (24). Previous report has demonstrated that in a comparison of five modern immunoassays for serum cortisol post-ACTH stimulation tests using mass spectrometry as a reference method, the serum cortisol responses in healthy volunteers were assay- and gender-specific (25). As an example of the diversity of response, in healthy individuals, the post-ACTH cortisol levels have been reported to range from 19.6 µg/dL (542 nmol/L) in females with the Architect assay to 27.9 µg/dL (772 nmol/L) in males with the Roche Diagnostic assay (25). Thus, the results in the present study may not be directly applicable for institutes using assays other than the Roche diagnostic ECLIA assay. These results should be applied only with the subgroup with indeterminate serum 0800 h cortisol levels between 3-17.9 µg/dL as that is the only population included in the analysis. The gold standard for diagnosing AI, i.e. the insulin tolerance test, was not used in this study. The ACTH stimulation test cannot diagnose patients with partial or recent secondary AI, so the AI status of those patients cannot be determined based solely on the ACTH stimulation test. The specific time at which the ACTH stimulation test was performed was not documented. Only the period (0900 h-1300 hr) was recorded. As levels of serum basal cortisol exhibit a diurnal variation, it may have affected the delta cortisol levels in those whose ACTH stimulation tests were conducted after noon (26). In the present study, ACTH level tests were not performed in all patients who had been diagnosed with AI. The causes of AI in the present study were based solely on clinical suggestions. The population included in this study were varied in indications of testing, therefore the results may not be relevant for some subgroups although the heterogeneity of the sample does suggest a level of generalizability.

CONCLUSIONS

Serum delta cortisol, especially the 30-minute delta cortisol level from ACTH stimulation tests, can facilitate AI diagnosis. The cut-off levels proposed for 30-minute delta cortisol yield high accuracy. Using the proposed values of 30-minute delta cortisol levels offers increased simplicity for healthcare practitioners and the patients, although 60-minute serum cortisol level may be needed for almost half of the patients. Further prospective study is needed to confirm and validate these results.

CONFLICT OF INTEREST

None of the authors have any potential conflicts of interest associated with this research.

ACKNOWLEDGEMENTS

We would like to thank Ms. Laddawan Limpjarnkij for the data collection and for performing the ACTH stimulation tests

and G. Lamar Robert, PhD and Chongchit Robert, PhD for reviewing the manuscript.

REFERENCES

- Bornstein SR, Allolio B, Arlt W, Barthel A, Don-Wauchope A, Hammer GD, Husebye ES, Merke DP, Murad MH, Stratakis CA, Torpy DJ: Diagnosis and Treatment of Primary Adrenal Insufficiency: An Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab*; 101(2): 364-89, 2016.
- Struja T, Briner L, Meier A, Kutz A, Mundwiler E, Huber A, Mueller B, Bernasconi L, Schuetz P: DIAGNOSTIC ACCURACY OF BASAL CORTISOL LEVEL TO PREDICT ADRENAL INSUFFICIENCY IN COSYNTROPIN TESTING: RESULTS FROM AN OBSERVATIONAL COHORT STUDY WITH 804 PATIENTS. *Endocr Pract*; 23(8): 949-61, 2017.
- Perton FT, Mijnhout GS, Kollen BJ, Rondeel JM, Franken AA, Groeneveld PH: Validation of the 1 mug short synacthen test: an assessment of morning cortisol cut-off values and other predictors. *Neth J Med*; 75(1): 14-20, 2017.
- 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*; 8(6): 427-34, 2002.
- Dickstein G, Shechner C, Nicholson WE, Rosner I, Sherrill Z, Adawi F, Lahav M: Adrenocorticotropin stimulation test: effects of basal cortisol level, time of day, and suggested new sensitive low dose test. *J Clin Endocrinol Metab*; 72(4): 773-8, 1991.
- Chitale A, Musonda P, McGregor AM, Dhatariya KK: Determining the utility of the 60 min cortisol measurement in the short synacthen test. *Clin Endocrinol (Oxf)*; 79(1): 14-9, 2013.
- 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*; 83(5): 1558-62, 1998.
- Cartaya J, Misra M: The low-dose ACTH stimulation test: is 30 minutes long enough? *Endocr Pract*; 21(5): 508-13, 2015.
- Manosroi W, Phimphilai M, Khorana J, Atthakomol P: Diagnostic performance of basal cortisol level at 0900-1300 h in adrenal insufficiency. *PLoS One*; 14(11): e0225255, 2019.
- Anantharaman R, Menezes G, Yusuf R, Ganapathi B, Ayyar SV, Srinivasan R: The 1 mug cosyntropin test in normal individuals: A reappraisal. *Indian J Endocrinol Metab*; 17(4): 693-6, 2013.
- Zou KH, O'Malley AJ, Mauri L: Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*; 115(5): 654-7, 2007.
- Nye EJ, Grice JE, Hockings GI, Strakosch CR, Crosbie GV, Walters MM, Jackson RV: Comparison of adrenocorticotropin (ACTH) stimulation tests and insulin hypoglycemia in normal humans: low dose, standard high dose, and 8-hour ACTH-(1-24) infusion tests. *J Clin Endocrinol Metab*; 84(10): 3648-55, 1999.
- Annane D, Pastores SM, Rochwerger B, Arlt W, Balk RA, Beishuizen A, Briegel J, Carcillo J, Christ-Crain M, Cooper MS, Marik PE, Umberto Meduri G, Olsen KM, Rodgers S, Russell JA, Van den Berghe G: Guidelines for the diagnosis and management of critical illness-related corticosteroid insufficiency (CIRCI) in critically ill patients (Part I): Society of Critical Care Medicine (SCCM) and European Society of Intensive Care Medicine (ESICM) 2017. *Intensive Care Med*; 43(12): 1751-63, 2017.
- 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*; 162(1): 109-13, 2010.
- Tordjman K, Jaffe A, Grazas N, Apter C, Stern N: The role of the low dose (1 microgram) adrenocorticotropin test in the evaluation of patients with pituitary diseases. *J Clin Endocrinol Metab*; 80(4): 1301-5, 1995.
- Rasmuson S, Olsson T, Hagg E: A low dose ACTH test to assess the function of the hypothalamic-pituitary-adrenal axis. *Clin Endocrinol (Oxf)*; 44(2): 151-6, 1996.
- Zueger T, Jordi M, Laimer M, Stettler C: Utility of 30 and 60 minute cortisol samples after high-dose synthetic ACTH-1-24 injection in the diagnosis of adrenal insufficiency. *Swiss Med Wkly*; 144: w13987, 2014.
- Munro V, Elnenaei M, Doucette S, Clarke DB, Imran SA: The effect of time of day testing and utility of 30 and 60minute cortisol values in the 250 mcg ACTH stimulation test. *Clin Biochem*; 54: 37-41, 2018.
- 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*; 47(2): 272-4, 1978.
- Bacon GE, Kenny FM, Murdaugh HV, Richards C: Prolonged serum half-life of cortisol in renal failure. *Johns Hopkins Med J*; 132(2): 127-31, 1973.
- Raff H, Trivedi H: Circadian rhythm of salivary cortisol, plasma cortisol, and plasma ACTH in end-stage renal disease. *Endocr Connect*; 2(1): 23-31, 2013.
- Verbeeten KC, Ahmet AH: The role of corticosteroid-binding globulin in the evaluation of adrenal insufficiency. *J Pediatr Endocrinol Metab*; 31(2): 107-15, 2018.
- Bujang MA, Adnan TH: Requirements for Minimum Sample Size for Sensitivity and Specificity Analysis. *Journal of clinical and diagnostic research: JCDR*; 10(10): YE01-YE6, 2016.
- Wallace I, Cunningham S, Lindsay J: The diagnosis and investigation of adrenal insufficiency in adults. *Ann Clin Biochem*; 46(Pt 5): 351-67, 2009.
- El-Farhan N, Pickett A, Ducrocq D, Bailey C, Mitchem K, Morgan N, Armston A, Jones L, Evans C, Rees DA: Method-specific serum cortisol responses to the adrenocorticotropin test: comparison of gas chromatography-mass spectrometry and five automated immunoassays. *Clin Endocrinol (Oxf)*; 78(5): 673-80, 2013.
- Kline GA, Buse J, Krause RD: Clinical implications for biochemical diagnostic thresholds of adrenal sufficiency using a highly specific cortisol immunoassay. *Clin Biochem*; 50(9): 475-80, 2017.