OPEN

Diagnostic Accuracy of Perioperative Measurement of Basal Anterior Pituitary and Target Gland Hormones in Predicting Adrenal Insufficiency After Pituitary Surgery

Vatroslav Cerina, MD, PhD, Ivan Kruljac, MD, Jelena Marinkovic Radosevic, MD, Lora Stanka Kirigin, MD, Darko Stipic, MD, Hrvoje Ivan Pecina, MD, PhD, and Milan Vrkljan, MD, PhD

Abstract: The insulin tolerance test (ITT) is the gold standard for diagnosing adrenal insufficiency (AI) after pituitary surgery. The ITT is unpleasant for patients, requires close medical supervision and is contraindicated in several comorbidities. The aim of this study was to analyze whether tumor size, remission rate, preoperative, and early postoperative baseline hormone concentrations could serve as predictors of AI in order to increase the diagnostic accuracy of morning serum cortisol.

This prospective study enrolled 70 consecutive patients with newly diagnosed pituitary adenomas. Thirty-seven patients had nonfunctioning pituitary adenomas (NPA), 28 had prolactinomas and 5 had somatotropinomas. Thyroxin (T4), thyrotropin (TSH), prolactin, follicle-stimulating hormone (FSH), luteinizing hormone (LH), testosterone, and insulin-like growth factor 1 (IGF-I) were measured preoperatively and on the sixth postoperative day. Serum morning cortisol was measured on the third postoperative day (CORT3) as well as the sixth postoperative day (CORT6). Tumor mass was measured preoperatively and remission was assessed 3 months after surgery. An ITT was performed 3 to 6 months postoperatively.

Remission was achieved in 48% of patients and AI occurred in 51%. Remission rates and tumor type were not associated with AI. CORT3 had the best predictive value for AI (area under the curve (AUC) 0.868, sensitivity 82.4%, specificity 83.3%). Tumor size, preoperative T4, postoperative T4, and TSH were also associated with AI in a multivariate regression model. A combination of all preoperative and postoperative variables (excluding serum cortisol) had a sensitivity of 75.0% and specificity of 77.8%. The predictive power of CORT3 substantially improved by adding those variables into the model (AUC 0.921, sensitivity 94.1%, specificity 78.3%, PPV 81.9%, NPV of 92.7%). In a subgroup analysis that included only female patients with NPA, LH had exactly the same predictive value as CORT3. The addition of baseline LH to CORT3, increased sensitivity to 100.0%, specificity to 88.9%, PPV to 90.4%, and NPV to 100.0%.

Besides CORT3, tumor size, thyroid hormones, and gonadotropins can serve as predictors of AI. LH in postmenopausal female patients with NPA has similar diagnostic accuracy as CORT3. Further studies are needed in order to validate the scoring system proposed by this study.

Abbreviations: AI = adrenal insufficiency, AUC = area under the curve, CORT3 = serum morning cortisol on the third postoperative day, CORT6 = serum morning cortisol on the sixth postoperative day, FSH = follicle-stimulating hormone, HPA = hypothalamus−pituitary−adrenal axis, IGF-I = insulin-like growth factor 1, ITT = insulin tolerance test, LH = luteinizing hormone, MRI = magnetic resonance imaging, NPA = nonfunctional pituitary adenoma, NPV = negative predictive value, PPV = positive predictive value, ROC = receiver operating characteristic, T3 = thyronin, T4 = thyroxin, TSH = thyrotropin.

INTRODUCTION

Pituitary tumors and pituitary surgery are the most common causes of adrenal insufficiency (AI). Approximately 44% to 52% of patients will have AI after pituitary surgery.[1,2] AI requires hydrocortisone replacement therapy in order to avoid adrenal crisis. The insulin tolerance test (ITT) is the gold standard test for detecting AI[3]; however, there are several disadvantages when using the ITT to assess the hypothalamic−pituitary−adrenal (HPA) axis. It is unpleasant for patients and close medical supervision is required during testing. In addition, it is contraindicated in patients with ischemic heart disease, arrhythmias, or epilepsy, and can only be used in older patients with caution.[3] Although rare, dynamic pituitary testing in patients with large adenomas can precipitate pituitary apoplexy.[4] Hence, it may be inconvenient and dangerous to perform an ITT in the early postoperative period. Low dose and standard dose synthetic ACTH stimulation tests are safe but unreliable and should not be used in the early postoperative period.[5,6]

Randomized control trials evaluating the appropriate postoperative assessment and ideal replacement therapy have not been conducted. Few studies have assessed the use of serum morning cortisol in the early postoperative period as a predictor of AI in order to select patients who require replacement therapy.[1,7] Unfortunately, these studies failed to accurately detect AI in the largest group of patients that had serum cortisol between 100 and 450 nmol/L. Current recommendations based on small retrospective studies suggest that the early postoperative assessment should include daily clinical examination as well as 0800 hours plasma cortisol levels.[8] Cortisol levels over 450 nmol/L reflect normal HPA function, and levels less than 100 nmol/L are consistent with ACTH deficiency. Patients with

cortisol levels between 100 and 250 nmol/L may be ACTH deficient and should receive morning hydrocortisone replacement until definitive HPA axis testing. Patients with cortisol levels between 250 and 450 nmol/L are unlikely to be ACTH deficient but should receive additional steroids in times of stress until a definitive test is performed.[6,8]

Corticotropic, thyrotropic, gonadotropic, somatotropic, and lactotropic cells are scattered within the anterior pituitary gland. Evidence suggests that injury to the anterior pituitary rarely leads to isolated corticotropin deficiency.[9] Deficiencies of other hormonal axes and local tumor compressive symptoms may accompany, precede, or follow AI.[10] Thus, other pituitary hormones and hormones of their target glands might also predict the onset of AI. The aim of our study was to analyze whether tumor size, remission rates, preoperative, and early postoperative baseline hormone concentrations could serve as predictors of AI in order to increase the diagnostic accuracy of morning serum cortisol.

PATIENTS AND METHODS

Study Protocol

This prospective observational study enrolled 70 consecutive patients with newly diagnosed pituitary adenomas. Thirty-seven patients had nonfunctional pituitary adenomas (NPA), 28 had prolactinomas and 5 patients had somatotropinomas. Microprolactinomas were diagnosed in patients with adenomas <10 mm and serum prolactin (PRL) levels >70 mg/L, and macroprolactinomas in patients with adenomas >10 mm and serum PRL >100 mg/L.[11] Somatotropinomas were diagnosed in patients with increased age-adjusted IGF-I concentrations and GH >0.5 ng/mL during the oral glucose tolerance test (OGTT).[12] Cushing disease was diagnosed in patients with increased morning serum cortisol, increased urinary-free cortisol and morning cortisol >50 nmol/L after an overnight 1 mg dexamethasone suppression test, along with normal or elevated serum concentrations of ACTH.[13] Patients who met the criteria for Cushing disease were excluded from the study. Patients that did not meet the previously mentioned criteria were diagnosed with NPA.

All patients underwent magnetic resonance imaging (MRI) of the sellar region using the standard protocol for pituitary imaging. This included sagittal and coronal T1-weighted images and T1 postcontrast images, as well as T2-weighted images in the coronal plane. The same experienced neuroradiologist (HIP) reviewed all MRI images in order to reduce observer bias. The neuroradiologist was blinded to the clinical information during MRI evaluation. Two days before surgery, hormone measurements were obtained at 8 AM following an overnight fast. The hormonal evaluation included measurements of serum total T4, TSH, PRL, FSH, LH, testosterone in men, IGF-I, morning serum cortisol, ACTH, and 24-hour urinary-free cortisol. The same experienced neurosurgeon (VC) operated the patients using the endoscopic transsphenoidal approach.[14] On the day of surgery, all patients received 100 mg of hydrocortisone intravenously. Oral hydrocortisone replacement therapy was continued in all patients at a dose of 20 mg in the morning and 10 mg at 5:00 PM. On the sixth postoperative day, serum total T4, TSH, PRL, FSH, LH, testosterone in men, and IGF-I was measured. Serum morning (8 AM) cortisol was measured on the third postoperative day (CORT3) as well as the sixth postoperative day (CORT6), before patients took their morning dose of hydrocortisone.

An ITT was performed in all patients 6 to 9 months postoperatively. Hydrocortisone replacement therapy was withheld 36 hours before testing and the ITT was performed according to standard protocol. Initial baseline cortisol and glucose levels were measured though an indwelling cannula after 30 minutes. Insulin was administered at a dose of 0.1 to 0.2 U/kg. Glucose, cortisol and insulin levels were measured 10, 20, 30, 60, and 120 minutes later. Patients with peak cortisol levels of ≥550 nmol/L, along with plasma glucose ≤2.2 mmol/L were considered to have normal adrenal function.

Remission was assessed on the third postoperative month and was defined as the following: normalization of serum PRL 3 months after surgery in patients with prolactinomas; normalization of IGF-I along with GH <0.5 ng/mL during the OGTT in patients with somatotropinomas; absence of residual tumor tissue on MRI in patients with NPA.

The ethics committee at the University Hospital Center Sestre milosrdnice, "Etičko povjerenstvo" approved the study and all patients gave written informed consent.

Hormonal Assays

PRL was measured using the flow injection analysis method—reference range for men was 2.0 to 20.0 μg/L and for women 2.0 to 30.0 μg/L. Electrochemiluminescence immunoassay was used for measurement of serum cortisol, growth hormone (GH), insulin-like growth factor-I (IGF-I), follicle-stimulating hormone (FSH), and luteinizing hormone (LH). The reference ranges for morning cortisol, GH, IGF-I were 138 to 800 nmol/L, 0.0 to 0.5 ng/mL, and 115 to 420 ng/mL, respectively. The reference range for FSH in the follicular phase was 3.5 to 12.5 IU/L for women in reproductive age, 25.8 to 134.8 IU/L for menopausal women, and 1.5 to 12.4 IU/L for men. The reference range for LH in the follicular phase was 2.4 to 12.6 IU/L, 7.7 to 58.5 IU/L for menopausal women and 1.7 to 8.6 IU/L for men. The radioimmunoassay method was used to determine the 24-hour urinary-free cortisol and testosterone. The reference range for 24-hour urinary-free cortisol was 72.5 to 372.0 nmol/24 hours. The reference range for testosterone in men aged 20 to 50 years was 8.6 to 29.0 nmol/L and in men older than 50 years 6.7 to 25.7 nmol/L. Chemiluminescence immunoassays were used for thyroid-stimulating hormone (TSH), T4, and adrenocorticotropic hormone (ACTH) measurement. Reference ranges were the following: for TSH 0.4 to 4.0 mIU/L, for T4 60 to 165 nmol/L, and for ACTH 2.0 to 13.3 pmol/L.

Statistical Analyses

Patient characteristics were assessed using descriptive statistics and presented as median values with interquartile ranges. Independent continuous variables were compared using Mann–Whitney test and categorical variables were compared using Fisher exact test. Longitudinal values were reported as median and interquartile range. The association between tumor type, age, gender, and remission and the development of AI was analyzed using binary and multinomial logistic regression when appropriate. It was also used to adjust for confounding factors. A linear mixed effects model was used to determine which hormones were different in patients with and without AI. Since the majority of hormones did not follow normal distribution, they were all transformed by logarithm to base 10 before the liner mixed effects model analysis. A linear mixed effects model with AR-heterogeneous covariance structure was used to compare the difference between patients with and without AI. The

subject-specific intercept was considered a random effect, and time, group, and time by group interaction were considered fixed effects. Tumor size was also considered a fixed effect covariate, adjusting for baseline individual differences. Mean differences reported within the text present the difference of log-transformed values for each variable. Mean differences are results of linear mixed effects model analysis and compare the magnitude of difference between the variables. Concentrations of hormones and the true differences in hormone concentrations are presented in Table 1. However, these differences are only informative and we did not report *P* values for changes between preoperative and postoperative concentrations. Variables that were independently associated with AI were included in receiver operating characteristic (ROC) analysis. ROC analysis was performed on raw data, but only for parameters that showed significant differences on linear mixed effect model analysis. ROC analysis aimed to establish cut-off hormone concentrations and their sensitivity, specificity, positive, and negative predictive values. Scoring system was developed based on z-scores obtained from ROC analysis. The initial power analysis was performed using the area under the curve (AUC) values obtained from previous studies. When considering type 1 error (α) to be 5% and type 2 error (β) to be 20%, the minimum sample size was 20 (10 patients with AI and 10 patients without AI). However, a larger sample size was needed, since we aimed to analyze several other parameters, which have not been analyzed previously. Our sample size of 70 patients with patients divided in a 1:1 ratio based on the presence of AI, had a power of 99% for 3rd day cortisol. Power was sufficiently high for all associations with AUC 0.689 or greater. IBM SPSS Statistics for Windows, Version 20.0 (IBM Corp., Armonk, NY) was used to perform all analyses. *P*-value < 0.05 was considered significant.

RESULTS

Anthropometric Characteristics

The study was performed from the year 2012 to 2013 and included 42 women with a mean age of 38 years (range 30.0–52.0) and 28 men with a mean age of 54 years (range 42.5–61.0). Male patients were significantly older ($P < 0.001$), but neither gender nor age was associated with AI. Patients had a median tumor size of 20.0 mm (range 10.0–27.5). Patients with AI had significantly larger tumors (23 mm vs 17 mm, $P = 0.004$). The association between tumor size and AI remained significant in a logistic regression model adjusted for tumor type (OR 1.070, 95% CI: 1.013–1.129, $P = 0.015$). Remission was achieved in 48% of patients and AI occurred in 51% of patients. There was no statistically significant difference in remission between certain types of tumors nor did AI depend on tumor type (Table 1).

TABLE 1. Characteristics of the Study Population Divided Based on the Presence of Adrenal Insufficiency

	Adrenal Insufficiency (N = 36)	Without Adrenal Insufficiency (N = 34)	*P*
Age (y)	48.0 (33.5–58.5)	47.0 (35.2–60.7)	0.796
Tumor size (mm)	23.0 (16.0–30.0)	17.0 (9.0–22.2)	0.004
Remission, n (%)	40.0 (14)	57.6 (19)	0.231
Gender (male) % (n)	47.2 (17)	32.4 (11)	0.231
Tumor type			0.644
Somatotropinoma % (n)	5.6 (2)	8.8 (3)	
Prolactinoma % (n)	36.1 (13)	44.1 (15)	
Nonfunctional adenoma % (n)	58.3 (21)	47.1 (16)	
PRL (μg/L) baseline	15.0 (11.8–25.6)	9.1 (7.5–26.2)	0.655
6th day	9.2 (3.3–15.3)	4.7 (3.6–9.2)	0.258
T4 (nmol/L) baseline	84.8 (68.5–97.6)	97.5 (84.6–119.0)	0.008
6th day	95.8 (80.7–111.2)	118.0 (93.4–130.7)	0.022
TSH (mIU/L) baseline	1.7 (1.1–2.8)	2.1 (1.2–3.1)	0.342
6th day	0.6 (0.3–1.2)	1.2 (0.6–1.8)	0.013
IGF-I (ng/ml) baseline	109.0 (64.6–152.0)	143.5 (114.0–190.0)	0.039
6th day	119.0 (74.3–175.0)	180.0 (139.7–243.0)	0.017
Testosterone (nmol/L) baseline	4.3 (0.4–9.2)	8.7 (5.9–13.5)	0.059
6th day	3.8 (0.6–10.1)	7.0 (5.1–11.8)	0.200
FSH (IU/L) baseline	4.6 (1.7–7.4)	8.4 (3.4–35.9)	0.016
6th day	4.0 (2.3–6.8)	5.2 (3.1–31.9)	0.052
LH (IU/L) baseline	1.6 (0.6–3.8)	5.5 (2.4–23.0)	0.001
6th day	1.8 (1.0–4.0)	7.1 (1.7–17.5)	0.003
Cortisol (nmol/L)			
Baseline	360.5 (150.0–579.5)	476.0 (348.0–598.5)	0.062
CORT3	248.0 (115.0–340.7)	491.0 (364.5–547.0)	<0.001
CORT6	256.0 (174.5–420.2)	461.0 (397.7–546.7)	<0.001
Urinary-free cortisol (nmol/24h)	92 (41–167)	145 (98–239)	0.041

Continues variables are presented as median with interquartile ranges (25–75%).
CORT3 = morning cortisol on the 3rd postoperative day, CORT6 = morning cortisol on the 6th postoperative day, FSH = follicle-stimulating, IGF-I = insulin-like growth factor 1, LH = luteinizing hormone, PRL = prolactin, T4 = thyroxin, TSH = thyrotropin.

The Association Between Laboratory Parameters and Development of AI

The gold standard test (ITT) was performed 6 to 9 months postoperatively. AI was diagnosed in 36 (51%) patients, while 34 (49%) patients did not have AI. Patients with AI had significantly lower preoperative T4, IGF-I, FSH, LH, 24-hour urinary-free cortisol and lower postoperative T4, TSH, IGF-I, LH, CORT3, and CORT6. There was no difference in other parameters between the study groups (Table 1). In a linear mixed model adjusted for age and tumor size, patients with AI had significantly lower cortisol levels (mean difference −0.238, 95% CI: −0.355 to −0.121, $P < 0.001$) (Figure 1A), T4 (mean difference −0.104, 95% CI: −0.167 to −0.041, $P = 0.002$) (Figure 1B), and TSH (mean difference −0.257, 95% CI: −0.472 to −0.042, $P = 0.02$) (Figure 1C). The association between baseline 24-hour urinary-free cortisol and AI diminished after adjustment for tumor size.

T4 levels significantly increased after the surgery (mean difference 0.032, 95% CI: 0.001–0.063, $P = 0.046$), whereas TSH decreased (mean difference −0.325, 95% CI: −0.481 to −0.168, $P < 0.001$). After additional adjustment for tumor type, serum PRL decreased in NPA (mean difference −0.446, 95% CI: −0.591 to −0.300, $P < 0.001$), but substantially more in prolactinomas (mean difference −0.768, 95% CI: −1.020 to −0.515, $P < 0.001$). There were no differences in serum PRL levels in patients with and without AI. Serum IGF-I increased after surgery (mean difference 0.078, 95% CI: 0.014–0.141, $P = 0.018$), but it was not correlated with AI. After adjustment for tumor size, there were no differences between study groups.

Overall, serum FSH and LH levels were lower in patients with AI. Female patients with prolactinomas had significantly lower baseline FSH (6.3 vs 22.3, $P = 0.018$) and lower postoperative FSH (4.9 vs 22.9, $P = 0.004$) and LH (4.8 vs 12.2, $P = 0.027$). Therefore, gonadotropins were analyzed separately in patients with prolactinomas. In patients with prolactinomas, there was no difference in FSH, LH, and testosterone levels in those with and without AI. However, female patients with NPA and AI had significantly lower serum LH (mean difference −1.065, 95% CI: −1.672 to −0.458, $P = 0.002$) (Figure 1D) and FSH (mean difference −0.863, 95% CI: −1.207 to −0.520, $P < 0.001$) levels.

ROC Analysis and Scoring System

The sensitivity, specificity, positive, and negative predictive value for each parameter are presented in Table 2. CORT3 had the best predictive value for AI (AUC 0.868, 95% CI: 0.727–0.954). However, in a subgroup analysis of female patients with NPA, both preoperative and postoperative LH had exactly the same predictive value as CORT3 (Figure 2).

In order to make our results more clinically applicable, we developed a scoring system based on z scores derived from ROC analysis. Each variable was labeled as follows: tumor
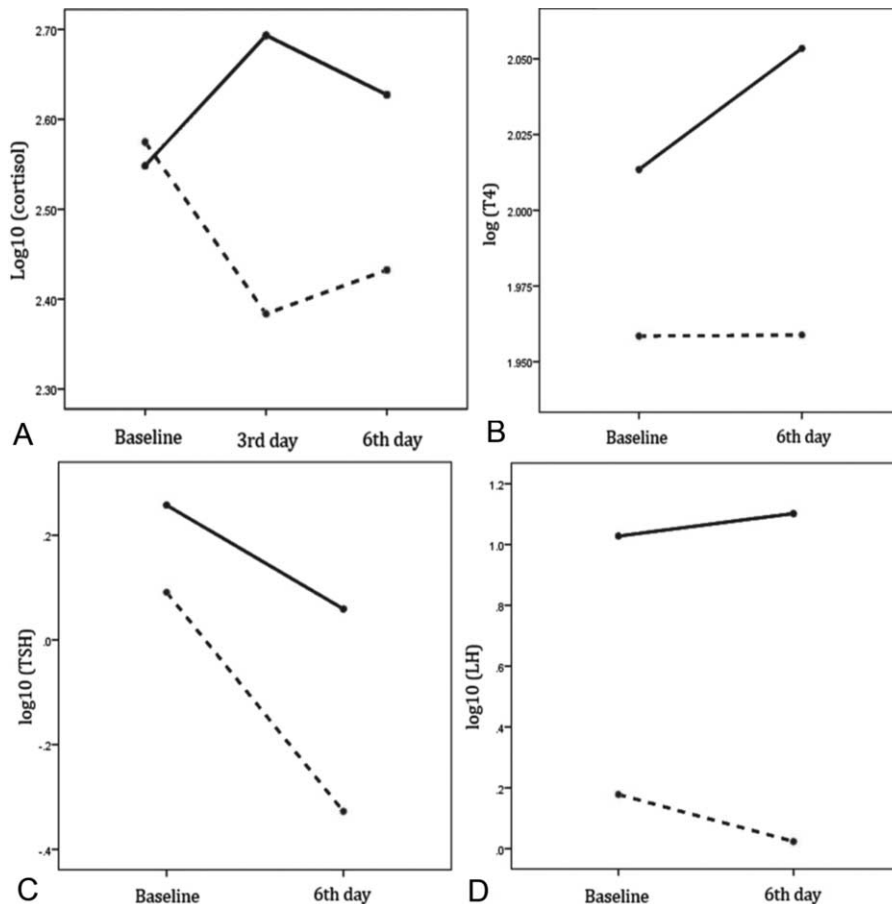

FIGURE 1. Linear mixed model analysis adjusted for tumor size showing the changes in serum cortisol (A), T4 (B), TSH (C), and LH (D) between patients with adrenal insufficiency (interrupted line) and without adrenal insufficiency.

TABLE 2. Results of ROC Analysis Showing the Sensitivity, Specificity, Positive, and Negative Predicitive Values for Each Cut-Off Value of the Variable

	AUC	95% CI for AUC	Cut-Off Value	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Tumor size	0.694	0.570–0.801	<9	91.9	30.0	57.7	78.1
			>19	67.6	63.3	65.7	65.2
			> 29	35.1	93.3	84.5	58.0
T4 baseline	0.665	0.538–0.776	≤69.9	27.8	90.0	74.3	54.5
			≤93.7	69.4	60.0	64.4	65.3
			≤123	94.4	20.0	55.1	77.4
T4 6th day	0.662	0.529–0.778	≤81	26.7	93.6	81.3	55.1
			≤114	83.3	54.8	65.7	75.9
			≤134	90.0	19.4	52.0	56.5
TSH 6th day	0.670	0.538–0.785	≤0.43	36.7	90.3	79.7	57.8
			≤1.24	80.0	48.4	61.7	69.9
			≤1.96	93.3	22.6	56.6	78.6
CORT3	0.868	0.727–0.954	≤248	52.9	95.8	92.8	65.9
			≤343	82.4	83.3	83.7	82.0
			≤464	100.0	58.3	71.4	100.0
CORT6	0.780	0.642–0.884	≤320	62.5	94.7	92.5	70.8
			≤515	90.6	36.8	59.9	79.0
LH baseline[*]	0.868	0.618–0.981	≤3.4	77.8	87.5	86.6	79.1
			≤5.2	100.0	75.0	80.6	100.0
LH 6th day[*]	0.854	0.645–0.965	≤5.3	80.0	92.3	91.5	81.6
			≤14.3	90.0	69.2	75.3	86.9
CORT6+ all baseline and postoperative	0.867	0.743–0.946	>2.5	96.8	20.0	55.7	85.7
			>7.9	80.0	90.0	88.1	74.9
			>11	35.5	100.0	100.0	59.8
CORT3+ all baseline and postoperative	0.921	0.790–0.982	>5.3	100.0	56.5	70.5	100.0
			>7.9	94.1	78.3	81.9	92.7
			>11	76.5	95.7	94.9	79.6
Baseline and postoperative	0.788	0.657–0.887	0	100.0	22.2	57.2	100.0
			>5.3	75.0	77.8	77.9	74.9
			>11	0.0	100.0	100.0	NA
Baseline parameters	0.728	0.591–0.839	≥0	100.0	0.0	NA	100.0
			>3.1	53.6	85.2	79.0	63.8
			>5.9	0.0	100.0	100.0	NA
CORT3+ baseline LH[*]	0.944	0.669–1.000	>11	100.0	88.9	90.4	100.0
Baseline and postoperative + baseline LH[*]	0.889	0.596–0.993	>5.9	100.0	77.8	82.4	100.0
			>12.9	50.0	100.0	100.0	65.8

AUC = area under the curve, CI = confidence interval, CORT3 = morning cortisol on the 3rd postoperative day, CORT6 = morning cortisol on the 6th postoperative day, LH = luteinizing hormone, NA = not analysed, NPV = negative predictive value, PPV = positive predictive value, T4 = thyroxin, TSH = thyrotropin.
[*] Females with nonfunctional adenomas.

diameter >19 mm as 3.1, baseline T4 < 93.7 nmol/L as 2.8, CORT3 < 343 nmol/L as 6.2, CORT6 < 320 nmol/L as 4.2, postoperative T4 < 114 nmol/L as 2.5, postoperative TSH < 0.93 mIU/L as 2.6 and baseline LH in females with NF adenoma < 5.2 IU/L as 7.0. A score calculated based on the combination of all preoperative and postoperative variables excluding serum cortisol, yielded a similar predictive value as CORT3 (AUC 0.788, 95% CI: 0.657–0.887). Moreover, the predictive power of CORT3 substantially improved by adding other variables into the model (AUC 0.921, 95% CI: 0.790–0.982). A score of 7.9 had a sensitivity of 94.1%, specificity of 78.3%, PPV of 81.9%, and NPV of 92.7% (Figure 2). In female patients with NPA, when baseline LH was added to the other variables the AUC increased to 0.944 (95% CI: 0.669–1.000).

A score of 11 had a sensitivity of 100.0%, specificity of 88.9%, PPV of 90.4%, and NPV of 100.0%.

DISCUSSION

Our study is the first to demonstrate that tumor size, thyroid hormones, and gonadotropins can be used to predict the development of AI after pituitary surgery in postmenopausal women with NPA. Moreover, a combination of these parameters and LH levels alone, have a similar diagnostic value as CORT3.

Two previous studies observed an association between postoperative serum morning cortisol and the development of AI. In a small retrospective study by Jayasena et al a cut-off
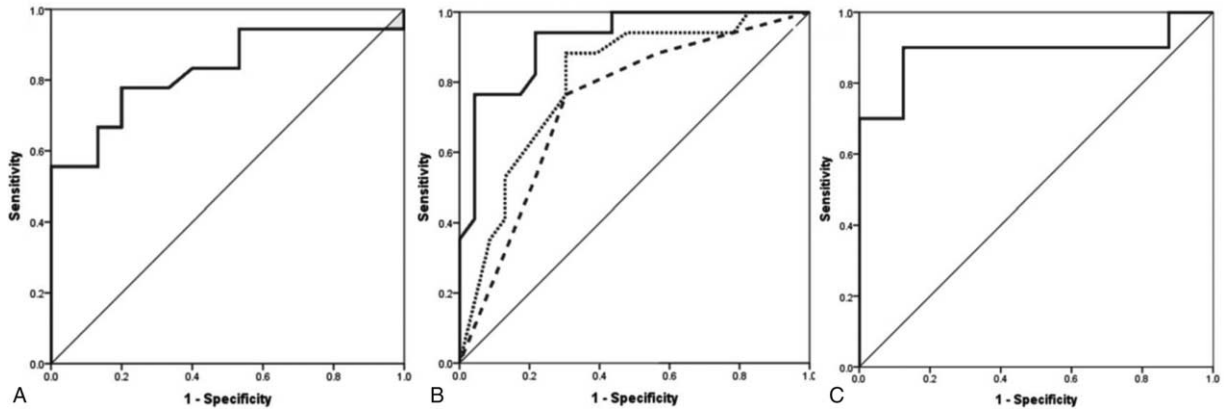
FIGURE 2. ROC curves showing the sensitivity and specificity of CORT3 (A); tumor size and preoperative T4 (interrupted line), tumor size, preoperative T4 and postoperative T4 and TSH (dotted line) and the combination of CORT3 with all the latter variables (full line) (B); preoperative LH in female patients with nonfunctioning adenomas (C).

cortisol concentration of 300 nmol/L on the 5th postoperative day had a 86.9% diagnostic sensitivity and a 69.2% diagnostic specificity7. In our study, CORT3 < 343 nnol/L had an 82.4% sensitivity and an 83.3% specificity, while CORT6 < 320 nmol/L had a 62.5% sensitivity and a 94.7% specificity. Our study is the first to demonstrate a difference between cut-off cortisol levels and time of cortisol measurement, suggesting that cortisol measured within the first 3 days after surgery might be a better predictor of AI. In a study by Karaca et al, morning serum cortisol was measured on the 2nd, 3rd, 4th, 5th, and 6th postoperative day. The authors reported only lower cut-off values with 100% specificity, but did not report cut-off values with the best sensitivity and specificity (based on Youden index) nor did they report AUC values. However, morning cortisol cut-off concentrations with 100% specificity ranged from 193 nmol/L (second and third postoperative day) to 83 nmol/L (sixth postoperative day). The various cut-off levels used on different days further demonstrates the importance of timing when measuring cortisol.[1] All patients included in our study were prescribed with hydrocortisone replacement therapy after surgery and the potential influence of glucocorticoid replacement therapy on morning cortisol cut-off values may be questioned. Because glucocorticoid replacement therapy increases serum cortisol levels, higher cortisol cut-off values might be expected in these patients. In a study by Jayasena et al all patients received glucocorticoid replacement therapy, while in a study by Karaca et al only patients with preoperatively diagnosed AI with ITT received glucocorticoid replacement therapy. Surprisingly, morning cortisol cut-off values with 100% specificity were higher in Karaca study (165 vs 111 nmol/L).

The cut-off cortisol values with 100% specificity in our study are comparable with those in Karaca study, which suggests that cut-off values are not influenced by glucocorticoid replacement therapy.

Our study disclosed several new insights on pituitary function in the early postoperative period. Although the correlation between tumor size and AI may be obvious, our study is first to determine the true prognostic value of tumor size in prediction of AI. Remission implies more radical tumor resection and potentially more pronounced pituitary injury. However, there was no correlation between remission rates and AI. The main goal of our study was to determine whether other pituitary and target gland hormones could predict AI.

Preoperative T4, LH, FSH and postoperative T4, TSH, LH, and FSH were associated with AI after adjustment for tumor size. This finding is interesting and could substantially change the evaluation of the HPA axis. The HPA axis is a critical stress-responsive system with prominent circadian rhythmicity, with highest cortisol levels at 8:00 AM and lowest levels during the night.[15] During severe stress, a 10-fold increase in cortisol may occur.[16] This is why a single cortisol measurement is not a reliable indicator of the HPA axis. GH has similar physiological properties but IGF-I has a more favorable half-life (12–15 hours).[17] However, dynamic tests are required for adequate evaluation of somatotropic function as well. On the other hand, thyroid hormones and sex hormones in males and postmenopausal women do not exhibit any circadian rhythmicity.[18,19] Hence, basal endocrinological testing is usually adequate for proper evaluation of thyroid and gonadal function. Since corticotropic, thyrotropic, gonadotropic, somatotropic, and lactotropic cells are scattered within the anterior pituitary gland, it may be justifiable to use thyrotropic and gonadotropic hormones to indirectly evaluate corticotropic and somatotropic function. Our study is the first to report that thyroid hormones and gonadotropins can be used as indicators of HPA axis function. A combination of tumor size, preoperative T4, postoperative T4, and TSH had a sensitivity of 75.0% and specificity of 77.8%. Similar findings were observed using only CORT3. Moreover, the addition of these parameters to CORT3, increased sensitivity to 94.1% and specificity to 78.3%. Interestingly, preoperative LH in postmenopausal women with NPA had a sensitivity of 100.0%, specificity of 75.0% and an AUC identical to CORT3. In premenopausal female patients LH was not associated with AI, possibly due to different stages of the menstrual cycle in which LH was measured and the additional suppressive effect of hyperprolactinemia. In male patients, this association could not be elucidated due to the relatively small sample size in the subgroup analyses.

Finally, we wish to comment on changes in T4, TSH, IGF-I, and PRL levels. Previous studies have documented a drop in serum PRL levels both in patients with prolactinomas and NPA immediately after surgery.[20] In patients with prolactinomas, this effect is explained by the reduction of PRL-secreting tumor tissue. In patients with NPA, the drop in PRL follows restoration of hypothalamic control over pituitary hormone secretion which leads to increased dopamine flow to the anterior pituitary;

therefore, changes in serum PRL in patients with NPA may serve as an indicator of overall pituitary function. Although a significant decrease in PRL levels occurred in patients with prolactinomas and NPAs in our study, we found no association between serum PRL and AI. The immediate postoperative increase in IGF-I we observed implies an overall improved pituitary function. The dynamics of thyroid hormones in our study remain elusive. The increase in total T4 and decrease in TSH could be explained by the interference of anesthetics, which could affect thyroid hormone binding proteins or peripheral conversion of T4 to T3. Unfortunately, free thyroid hormone levels were not determined and these questions need to be answered in future studies.

A limitation of our study was the small study population; therefore, we could not divide our population into 2 groups. Ideally, one group should have been used to develop the scoring system, and the other group to test the power of the proposed scoring system in a separate series of patients. Thus, future studies are needed to validate our scoring system in a different patient population.

In conclusion, our study has confirmed the role of basal cortisol measurement in the early postoperative period in predicting AI. CORT3 is a better predictor of AI than CORT6. Preoperative tumor size, perioperative thyroid hormones, and gonadotropins can serve as predictors of AI. Preoperative LH in postmenopausal female patients with NPA has similar diagnostic accuracy as CORT3.

REFERENCES

1. Karaca Z, Tanriverdi F, Atmaca H, et al. Can basal cortisol measurement be an alternative to the insulin tolerance test in the assessment of the hypothalamic-pituitary-adrenal axis before and after pituitary surgery? *Eur J Endocrinol*. 2010;163:377–382.

2. Berg C, Meinel T, Lahner H, et al. Recovery of pituitary function in the late-postoperative phase after pituitary surgery: results of dynamic testing in patients with pituitary disease by insulin tolerance test 3 and 12 months after surgery. *Eur J Endocrinol*. 2010;162:853–859.

3. Fish HR, Chernow B, O'Brian JT. Endocrine and neurophysiologic responses of the pituitary to insulin-induced hypoglycemia: a review. *Metabolism*. 1986;35:763–780.

4. Yoshino A, Katayama Y, Watanabe T, et al. Apoplexy accompanying pituitary adenoma as a complication of preoperative anterior pituitary function tests. *Acta Neurochir (Wien)*. 2007;6:557–565.

5. Klose M, Lange M, Kosteljanetz M, et al. Adrenocortical insufficiency after pituitary surgery: an audit of the reliability of the conventional short synacthen test. *Clin Endocrinol (Oxf)*. 2005;63:499–505.

6. Courtney CH, McAllister AS, McCance DR, et al. Comparison of one week 0900 h serum cortisol, low and standard dose synacthen tests with a 4 to 6 week insulin hypoglycaemia test after pituitary surgery in assessing HPA axis. *Clin Endocrinol (Oxf)*. 2000;53:431–436.

7. Jayasena CN, Gadhvi KA, Gohel B, et al. Day 5 morning serum cortisol predicts hypothalamic–pituitary–adrenal function after transsphenoidal surgery for pituitary tumors. *Clin Chem*. 2009;55:972–977.

8. Inder WJ, Hunt PJ. Glucocorticoid replacement in pituitary surgery: guidelines for perioperative assessment and management. *J Clin Endocrinol Metab*. 2002;87:2745–2750.

9. Vance ML. Hypopituitarism. *N Engl J Med*. 1994;330:1651–1662.

10. Oelkers W. Adrenal insufficiency. *N Engl J Med*. 1996;335:1206–1212.

11. Casanueva FF, Molitch ME, Schlechte JA, et al. Guidelines of the Pituitary Society for the diagnosis and management of prolactinomas. *Clin Endocrinol (Oxf)*. 2006;65:265–273.

12. Melmed S, Colao A, Barkan A, et al. Guidelines for acromegaly management: an update. *J Clin Endocrinol Metab*. 2009;94:1509–1517.

13. Nieman LK, Biller BM, Findling JW, et al. The diagnosis of Cushing's syndrome: an Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab*. 2008;93:1526–1540.

14. Marić A, Kruljac I, Čerina V, et al. Endocrinological outcomes of pure endoscopic transsphenoidal surgery: a Croatian Referral Pituitary Center experience. *Croat Med J*. 2012;53:224–233.

15. Veldhuis JD, Iranmanesh A, Johnson ML, et al. Amplitude, but not frequency, modulation of adrenocorticotropin secretory bursts gives rise to the nyctohemeral rhythm of the corticotropic axis in man. *J Clin Endocrinol Metab*. 1990;71:452–463.

16. Taylor AL, Fishman LM. Corticotropin-releasing hormone. *N Engl J Med*. 1988;319:213–222.

17. LeRoith D, Buler A. What is the role of circulating IGF-I? *Trends Endocrinol Metab*. 2001;12:48–52.

18. Izumi M, Larsen PR. Triiodothyronine, thyroxine, and iodine in purified thyroglobulin from patients with Graves' disease. *J Clin Invest*. 1977;59:1105–1112.

19. Bhasin S, Fisher CE, Sverdloff RS. Follicle-stimulating hormone and luteinizing hormone. In: Melmed S, ed. *The Pituitary*. 2nd ed Malden, MA: Blackwell Science; 2002:216–278.

20. Arafah BM, Nekl KE, Gold RS, et al. Dynamics of prolactin secretion in patients with hypopituitarism and pituitary macroadenomas. *J Clin Endocrinol Metab*. 1995;80:3507–3512.