CLINICAL STUDY

Accuracy of several parameters of hypothalamic–pituitary–adrenal axis activity in predicting before surgery the metabolic effects of the removal of an adrenal incidentaloma

Cristina Eller-Vainicher1, Valentina Morelli1, Antonio Stefano Salcuni2, Claudia Battista2, Massimo Torlontano2, Francesca Coletti3, Laura Iorio4, Elisa Cairoli1, Paolo Beck-Pecoz3, Maura Arosio1,3, Bruno Ambrosi1, Alfredo Scillitani2 and Iacopo Chiodini1

1Endocrinology and Diabetology Unit, Department of Medical Sciences, University of Milan, Fondazione IRCCS Cà Granda – Ospedale Maggiore Policlinico, Pad. Granelli, Via F. Sforza 35, 20122 Milan, Italy, 2Endocrinology Unit, ‘Casa Sollievo della Sofferenza’ IRCCS, San Giovanni Rotondo, 71013 Foggia, Italy, 3Endocrinology Unit, Ospedale ‘San Giuseppe Multimedica’, 20123 Milan, Italy and 4Endocrinology Unit, Department of Medical and Surgical Sciences, IRCCS Policlinico San Donato Institute, University of Milan, San Donato Milanese, 20097 Milan, Italy

(Correspondence should be addressed to I Chiodini; Email: iacopo.chiodini@email.it)

Abstract

Context: It is unknown whether the metabolic effects of the removal of an adrenal incidentaloma (AI) can be predicted by the assessment of cortisol hypersecretion before surgery.

Objective: To evaluate the accuracy of several criteria of hypothalamic–pituitary–adrenal axis activity in predicting the metabolic outcome after adrenalectomy.

Design: Retrospective longitudinal study.

Patients: In 55 surgically treated AI patients (Group 1) before surgery and in 53 nontreated AI patients (Group 2) at the baseline, urinary free cortisol (UFC), cortisol after 1 mg overnight dexamethasone-suppression test (1 mg-DST), ACTH, and midnight serum cortisol (MSC) were measured. In Groups 1 and 2, metabolic parameters were evaluated before and 29.6±13.8 months after surgery and at the baseline and after 35.2±10.9 months respectively.

Main outcome measures: The improvement/worsening of weight, blood pressure, glucose, and cholesterol levels (endpoints) was defined by the presence of a >5% weight decrease/increase and following the European Society of Cardiology or the ATP III criteria respectively. The accuracy of UFC, 1 mg-DST, ACTH, and MSC, singularly taken or in combination, in predicting the improvement/worsening of ≥2 endpoints was calculated.

Results: The presence of ≥2 among UFC >70 µg/24 h (193 nmol/l), ACTH <10 pg/ml (2.2 pmol/l), 1 mg-DST >3.0 µg/dl (83 nmol/l) (UFC–ACTH–DST criterion) had the best accuracy in predicting the endpoints’ improvement (sensitivity (SN) 65.2%, specificity (SP) 68.8%) after surgery. In the nontreated AI patients, this criterion predicted the worsening of ≥2 endpoints (SN 55.6%, SP 82.9%).

Conclusions: The UFC–ACTH–DST criterion seems to be the best for predicting the metabolic outcome in surgically treated AI patients.

European Journal of Endocrinology 163 925–935

Introduction

Incidentally discovered adrenal masses (adrenal incidentalomas (AIs)) are found with increasing frequency due to the widespread use of imaging techniques of the abdomen and their prevalence is estimated to be around 7% in the general population (1–4). Although, generally benign, AI may be associated with a subtle cortisol overproduction. When the typical signs or symptoms of overt hypercortisolism (Cushing’s syndrome (CS)) are absent, then this condition is called ‘subclinical hypercortisolism’ (SH) (5–9).

In the past years, several studies have shown that SH may be associated with some clinical consequences typical of CS, particularly regarding bone and features of the metabolic syndrome (10–18). More recently, some authors showed that the surgical removal of the adrenal mass causing SH leads to the improvement of some complications of the metabolic syndrome, thus convincingly demonstrating that this ‘subclinical’ cortisol excess has, in fact, clinical consequences (18–25). Nevertheless, some of these studies also reported that patients operated on for the size of AI and without a diagnosis of SH before surgery experienced the improvement of some of these metabolic complications after the surgery (18, 19, 21, 25), thus suggesting, that the diagnosis of SH had not been sensitive enough in predicting the outcome of the surgical approach.
Regarding the SH diagnosis, it must be considered that widely accepted guidelines are not available (26), and the criteria used are extremely variable among the different studies (27, 28). Up to now, no study evaluated the accuracy of the different parameters of SH in predicting the improvement of the SH consequences after recovery. This lack of knowledge is of utmost clinical importance since the finding of a parameter or a combination of parameters, which can predict the outcome of the recovery from SH, may help to identify subjects who can better take advantage from surgery. Therefore, the aim of this study is to evaluate the accuracy of several hypothalamic–pituitary–adrenal axis (HPA) activity parameters in predicting before surgery the improvement of some features of the metabolic syndrome after the surgical removal of an AI.

Subjects and methods

Subjects

The study was performed in four referral Italian Endocrinology Units: ‘Fondazione Ospedale Maggiore Policlinico, Mangiagalli e Regina Elena’ IRCCS in Milan, ‘Casa Sollievo della Sofferenza’ IRCCS in San Giovanni Rotondo, ‘San Giuseppe Multimedica’ Hospital, and IRCCS Policlinico San Donato, in Milan. From December 2002 to December 2007, we studied 55 consecutive patients who underwent the surgical excision of an AI (treated subjects, Group 1) and 53 consecutive AI patients who were followed up with a conservative approach (nontreated subjects, Group 2). Patients from both the groups were regularly followed up in our institutions. The diagnosis of AI was based on the detection of a unilateral adrenal mass (size > 1 cm) by cross-sectional imaging methods of the abdomen, in the course of diagnostic testing or treatment for other clinical conditions not related to adrenal disorders.

For both the groups the exclusion criteria were i) past or current history of hypogonadism (in men, testosterone levels < 300 ng/dl (10.4 nmol/l), and in premenopausal women, less than six menstrual cycles/year) and of diseases known to affect glucose metabolism other than type 2 diabetes (i.e. thyrotoxicosis, bowel diseases, chronic renal failure, chronic hepatic disease, alcoholism, eating disorders, rheumatologic, or hematological diseases); ii) administration of drugs influencing cortisol and dexamethasone metabolism or cortisol secretion (29, 30); and iii) signs or symptoms specific for cortisol excess (easy bruising, facial plethora, striae rubrae, and proximal muscle weakness), which best discriminate CS (29).

No subject had evidence of metastatic diseases. Computed tomography revealed that all adrenal masses were homogeneous, hypodense, and with well-shaped features, consistent with the diagnosis of an adenocortical adenoma. In all patients, the diagnosis of pheochromocytoma and aldosteronoma was excluded by appropriate hormonal determinations (24 h urinary catecholamines and/or metanephrines and upright plasma renin activity and aldosterone).

HPA axis function was studied before the surgery in treated and nontreated patients at the baseline and at the end of follow-up by measuring at least once urinary free cortisol (UFC) levels (normal values 10–70 μg/24 h, 28–193 nmol/l), serum cortisol levels after 1 mg overnight dexamethasone-suppression test (1 mg-DST), morning (0800 h) ACTH levels (normal values 10–55 pg/ml, 2–12 pmol/l), and midnight serum cortisol (MSC, normal values < 5.4 μg/dl, 149 nmol/l). When more than one determination of each parameter of HPA axis secretion was available at the baseline, then the mean value was considered.

Patients from Group 1 underwent surgery. In 41 patients (32 females, 9 males), the surgical approach was indicated mainly on the basis of the increasing dimensions (> 1 cm increase during the 12-month follow-up) and/or to the size larger than 4 cm at the diagnosis. In 14 subjects (11 females, 3 males), the surgical option was suggested on the basis of the presence of at least three of the following alterations of HPA axis activity: UFC levels > 70 μg/24 h (193 nmol/24 h), 1 mg-DST > 3.0 μg/dl (83 nmol/l), ACTH levels < 10 pg/ml (2.2 pmol/l), and MSC > 5.4 μg/dl (149 nmol/l). In our protocol, we used a 1 mg-DST cut-off of 3.0 μg/dl (83 nmol/l) rather than 5 μg/dl (138 nmol/l) as recommended by the NIH (31) in order to increase the test sensitivity (SN). Laparoscopic or laparotomic (open) adrenalectomy was performed depending on the size of the adrenal adenoma and the clinical characteristics of subjects. No patient had surgical complications. In all patients the histological findings were consistent with adrenal adenoma. After adrenalectomy, a precautionary steroid therapy with hydrocortisone 100 mg i.v., during the surgery, and with cortisone acetate per os (at weight-related doses ranging between 25 and 37.5 mg/day in three subdivided doses during the day), immediately after the operation, was administered. The commonly used cortisone acetate dose was 25 mg/day, while higher doses of 31.3 and 37.5 mg/day were used in four and two obese patients respectively. In all patients, cortisol secretion was re-evaluated, after 2 months, by ACTH stimulation test. If inconclusive results were obtained, an insulin tolerance test was performed. In patients with persistent adrenal insufficiency, steroid substitutive therapy was continued and the HPA axis function was re-assessed every 6 months. The duration of the steroid substitutive therapy was comparable between patients operated on for the size of the adenoma and for the presence of the alterations of HPA axis activity (13.9 ± 22.6 vs 11.6 ± 16.3 months, P = 0.731). After the discontinuation of the steroid substitutive therapy, the HPA axis secretion was restored in all subjects treated for the alterations of HPA axis activity (n = 14), with a significant increase in ACTH (before 8.1 ± 3.2 pg/ml,
after 12.3 ± 4.1 pg/ml, P = 0.006) and decrease in UFC (before 80.1 ± 37.2 μg/24 h, after 47.9 ± 17.4 μg/24 h, P = 0.007), 1 mg-DST (before 3.9 ± 1.1 μg/dl, after 1.3 ± 0.5 μg/dl, P < 0.0001), and MSC (before 7.01 ± 4.0 μg/dl, after 2.8 ± 0.9 μg/dl, P = 0.001) levels.

Post operation, at the end of the follow-up period, in patients operated on for the size of the adenoma (n = 41), the biochemical testing showed a significant increase in ACTH values (before 10.4 ± 5.6 pg/ml, after 13.2 ± 3.8, P = 0.009), and decrease in 1 mg-DST levels (before 2.4 ± 1.1 μg/dl, after 1.4 ± 0.4 μg/dl, P < 0.0001), while UFC (before 40.8 ± 19.2 μg/24 h, after 39.8 ± 18.3 μg/24 h, P = NS) and MSC (before 3.4 ± 1.4 μg/dl, after 3.6 ± 2 μg/dl, P = NS) levels did not show a significant change with respect to the preoperative period.

Group 2 included patients for whom surgery was not indicated or was refused. In these patients, the HPA axis activity was re-evaluated at 9-months interval during the follow-up period, and in all cortisol secretion did not change. No patient was excluded or switched groups during the follow-up period. Therefore, in no patient the clinical management needed to be modified.

Patients gave informed consent to participate and the study was conducted in accordance with Helsinki Declaration II and was approved by local ethical committees.

Methods

Serum and urinary samples were collected and stored at −20 °C until assayed. In all patients, serum ACTH levels (mean of three determinations at 20 min intervals) were measured by IRMA (BRAHMS Diagnostica GmbH, Berlin, Germany), and serum cortisol and UFC levels (after dichloromethane extraction) were determined immunofluorimetrically by TDX-FLX Abbott GmbH Diagnostika kits. The intra- and inter-assay coefficients of variation were < 15% for ACTH and < 10% for all other assays.

In all patients, clinical examination was performed at the baseline and at 9-months interval during the follow-up period, measuring weight, height, waist circumference, and recording the presence of obesity, hypertension, dyslipidemia, and type 2 diabetes mellitus. Subjects with body mass index (BMI) > 30 kg/m² were considered obese. Subjects with systolic blood pressure of ≥ 130 mmHg and/or diastolic blood pressure ≥ 85 mmHg and/or on antihypertensive treatment were defined as hypertensive. Diabetes mellitus was diagnosed using the WHO criteria (32) and patients were also considered diabetic if any hypoglycemic drug was given. Patients underwent oral glucose tolerance test (OGTT) unless the diagnosis of diabetes was manifested on the basis of fasting glucose levels. Dyslipidemia was defined as serum triglycerides levels of ≥ 150 mg/dl (≥ 1.7 mmol/l), or high-density lipoprotein (HDL) cholesterol levels of < 40 mg/dl (< 1.0 mmol/l) in men and < 50 mg/dl (< 1.3 mmol/l) in women (33). Patients were also considered dyslipidemic if any antidyslipidemic treatment was given.

The improvement or worsening of body weight was defined by a > 5% decrease or increase in body weight (34) respectively between the baseline and the end of the follow-up period. The improvement or worsening of arterial blood pressure was defined if during the follow-up period the nonhypertensive patients passed from a prehypertension category to another or the hypertensive patients from a hypertension grade to another, following the guidelines for the Management of Arterial Hypertension of the European Society of Cardiology (35). Fasting glucose and cholesterol (HDL and/or low-density lipoprotein, LDL) levels were considered improved or worsened if they passed from a category to another in agreement with the ATP III criteria (32). The improvement or worsening of body weight, arterial blood pressure, fasting glucose, and cholesterol levels (HDL and or LDL) was established at the last follow-up.

Treated patients were considered ‘improved’ or ‘not improved’ and nontreated patients ‘worsened’ or ‘not worsened’ if they experienced the improvement or worsening of at least two of the following parameters: body weight, blood pressure, fasting glucose, and cholesterol levels during the follow-up. We decided to use the LDL levels as endpoint since they are considered the best parameter to evaluate the efficacy of the hypolipidemic therapy (33).

Statistical analysis

Statistical analysis was performed by SPSS version 12.0 statistical package (SPSS Inc., Chicago, IL, USA). Results are expressed as mean ± S.D. along with the range for normally distributed continuous variables or as absolute frequencies and percentages for categorical variables.

Comparison of continuous variables was performed using t-test or Mann–Whitney U test for not normally distributed variables. Categorical variables were compared by χ² test or Fisher’s exact test, as appropriate.

Body weight, fasting glucose, blood pressure, and LDL cholesterol levels at the end of follow-up were used as endpoints. Receiver operator characteristics (ROC) analysis was performed to assess the best threshold values of 1 mg-DST and MSC for detecting treated and nontreated patients with improvement and worsening of at least two endpoints respectively.

The association between the number of improved endpoints and the parameters of the HPA axis activity was tested by Pearson or Spearman correlation as appropriate.

We calculated by χ² test or Fisher’s exact test, as appropriate, the SN, specificity (SP), positive predictive value (PPV), negative predictive value (NPV), and accuracy in predicting in treated patients the improvement after surgery of at least two endpoints for the
following parameters taken alone or in all possible combinations:

1 mg-DST: > 1.8 (50 nmol/l), > 2.0 (55 nmol/l) or > 3.0 (83 nmol/l) or > 5.0 μg/dl (138 nmol/l);
MSC: > 4.0 (110 nmol/l) or > 5.4 (149 nmol/l) or > 7.5 μg/dl (210 nmol/l);
ACTH: < 10.0 pg/ml (2.2 pmol/l);
UFC: > 70.0 μg/dl (193 nmol/l).

These cut-offs were chosen on the basis of literature data (25–28, 33) or the ROC curve (i.e. MSC > 4.0 μg/dl, 110 nmol/l and > 2.0 μg/dl, 55 nmol/l). The parameters or combinations of parameters that showed SN or SP or overall accuracy below 65.5% were excluded.

The same analyses were performed looking for the accuracy of the above mentioned criteria in predicting the improvement of at least one endpoint. The results are reported in the Supplementary Table (see section on supplementary data given at the end of this article).

Subsequently, a logistic regression analysis was performed to assess the odds ratio of the statistically significant parameters or combination of parameters of the HPA axis activity (independent variables) in predicting in the treated subjects the improvement after surgery of ≥ 2 endpoints (dependent variable), after adjustment for duration of the follow-up, age, size of the adenoma, and the presence of at least two of the following parameters: obesity, hypertension, type 2 diabetes, and dyslipidemia at the baseline (independent variables as covariates).

Subsequently, the parameters or combination of parameters that were found to be statistically associated with the improvement of ≥ 2 endpoints after adjustment for the above mentioned covariates were tested in the nontreated AI patients. The SN, SP, and accuracy of these HPA axis activity criteria for predicting the worsening of ≥ 2 endpoints were calculated by χ² test or Fisher’s exact test as appropriate. Finally, the logistic regression analysis was performed to evaluate if these criteria of the HPA axis activity (independent variables) were predictive of the worsening of ≥ 2 endpoints (dependent variable) even after adjustment for duration of the follow-up, age, and presence of ≥ 2 out of obesity, hypertension, type 2 diabetes, and dyslipidemia at the baseline (independent variables as covariates).

### Results

The ROC analysis, performed to assess the best threshold values of 1 mg-DST and MSC for detecting treated patients with improvement of at least two endpoints, showed that the cut-off with the best compromise between SN and SP was set at 2.0 μg/dl (55 nmol/l) for 1 mg-DST and at 4.0 μg/dl (110 nmol/l) for MSC. These cut-offs for 1 mg-DST and MSC were substantially confirmed (2.1 μg/dl, 58 nmol/l for 1 mg-DST and 3.9 μg/dl, 107 nmol/l for MSC) when the ROC analysis was performed for detecting nontreated patients with worsening of at least two endpoints. Therefore, we have considered these cut-offs as adjunctive parameters beside those already proposed and used in clinical practice so far (31, 36).

The clinical characteristics of all, ‘improved,’ and ‘not improved’ treated AI subjects (Group 1) are reported in Table 1. The cortisol levels after 1 mg-DST and the prevalence at the baseline of hypertension and dyslipidemia were higher in ‘improved’ than in ‘not improved’ patients, while ACTH, MSC, and UFC levels, and the prevalence of obesity and diabetes were comparable between the two groups (Table 1).

Table 2 reports the clinical characteristics of all, ‘worsened,’ and ‘not worsened’ untreated AI patients (Group 2). The ‘worsened’ patients at the baseline had

<table>
<thead>
<tr>
<th>Table 1 Clinical data at the baseline of all treated and of ‘improved’ and ‘not improved’ AI patients. Data are expressed as mean ± s.d. with range in parenthesis or as absolute numbers with percentage in parenthesis.</th>
<th>All (n = 55)</th>
<th>Improved (n = 23)</th>
<th>Not improved (n = 32)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>56.0 ± 11.2 (24–75)</td>
<td>56.1 ± 11.5 (33–74)</td>
<td>55.9 ± 11.2 (24–75)</td>
<td>0.962</td>
</tr>
<tr>
<td>Follow-up (months)</td>
<td>29.6 ± 13.8 (18–54)</td>
<td>29.8 ± 14.9 (18–54)</td>
<td>29.4 ± 13.2 (18–54)</td>
<td>0.915</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>28.1 ± 5.6 (19–40)</td>
<td>29.1 ± 6.7 (21–40)</td>
<td>27.4 ± 4.7 (19–39)</td>
<td>0.266</td>
</tr>
<tr>
<td>Waist circumference (cm)</td>
<td>92.1 ± 12.9 (64–120)</td>
<td>95.2 ± 15.4 (64–120)</td>
<td>89.8 ± 10.5 (70–118)</td>
<td>0.132</td>
</tr>
<tr>
<td>Diameter of the tumor (cm)</td>
<td>3.4 ± 1.2 (1.2–9.0)</td>
<td>3.4 ± 0.8 (2.0–5.0)</td>
<td>3.4 ± 1.4 (1.2–9.0)</td>
<td>0.853</td>
</tr>
<tr>
<td>Obese patients, n (%)</td>
<td>20 (36.4)</td>
<td>9 (39.1)</td>
<td>11 (34.4)</td>
<td>0.718</td>
</tr>
<tr>
<td>Hypertensive patients, n (%)</td>
<td>30 (54.5)</td>
<td>16 (69.6)</td>
<td>14 (43.8)</td>
<td>0.050</td>
</tr>
<tr>
<td>Diabetic patients, n (%)</td>
<td>11 (20)</td>
<td>7 (30.4)</td>
<td>4 (12.5)</td>
<td>0.170</td>
</tr>
<tr>
<td>Dyslipidemic patients, n (%)</td>
<td>21 (38.2)</td>
<td>14 (60.9)</td>
<td>7 (21.9)</td>
<td>0.003</td>
</tr>
<tr>
<td>ACTH (pg/ml)</td>
<td>9.8 ± 5.1 (1.0–26.9)</td>
<td>9.1 ± 3.8 (4.4–18.4)</td>
<td>10.3 ± 5.9 (1.0–26.9)</td>
<td>0.387</td>
</tr>
<tr>
<td>1 mg-DST (μg/dl)</td>
<td>2.9 ± 1.3 (0.9–5.4)</td>
<td>3.2 ± 1.2 (1.4–5.4)</td>
<td>2.4 ± 1.2 (0.9–5.1)</td>
<td>0.013</td>
</tr>
<tr>
<td>UFC (μg/24 h)</td>
<td>52.4 ± 31.3 (10.0–135.5)</td>
<td>49.6 ± 28.6 (10.4–107.6)</td>
<td>51.7 ± 31.5 (10.0–135.5)</td>
<td>0.806</td>
</tr>
<tr>
<td>MSC (μg/dl)</td>
<td>4.9 ± 3.2 (1.0–13.5)</td>
<td>4.2 ± 2.0 (1.0–8.4)</td>
<td>4.5 ± 3.3 (1.0–13.5)</td>
<td>0.719</td>
</tr>
</tbody>
</table>

The P levels are referred to the comparison between improved and not improved patients. AI, adrenal incidentaloma; ‘improved’/‘not improved’ patients who did or did not experience the improvement of ≥ 2 parameters out of body weight, blood pressure, fasting glucose, low-density lipoprotein levels; ACTH (SI conversion factor × 0.22); 1 mg-DST, cortisol after 1 mg overnight dexamethasone-suppression test (SI conversion factor × 27.56); UFC, urinary free cortisol (SI conversion factor × 27.56); MSC, midnight serum cortisol (SI conversion factor × 27.56).
higher BMI, prevalence of obesity, 1 mg-DST, and UFC levels than ‘not worsened’ patients.

In the treated subjects, among the parameters of the HPA axis activity, only 1 mg-DST was significantly correlated with the number of improved end points (r = 0.37, P = 0.005). The SN, SP, PPV, NPV, and accuracy of the criteria for diagnosing SH, which turned out to be accurate in predicting the improvement of at least two among body weight, fasting glucose, blood pressure, and LDL levels in treated patients, are reported in Table 3. The criteria showing an overall accuracy, SN, and SP > 65% were the following: two parameters out of DST ≥ 3.0 μg/dl (≥ 83 nmol/l), UFC ≥ 70.0 μg/24 h (≥ 193 nmol/24 h), ACTH ≤ 10.0 pg/ml (≤ 22 pmol/l) (Criterion I); two parameters out of 1 mg-DST ≥ 2.0 μg/dl (≥ 55 nmol/l), MSC ≥ 4.0 μg/dl (≥ 110 nmol/ml) (Criterion II); 1 mg-DST ≥ 2.0 μg/dl (≥ 55 nmol/l) plus MSC ≥ 4.0 μg/dl (≥ 110 nmol/l) (Criterion III).

Although excluded from the subsequent analyses due to low SN, Criterion V (1 mg-DST ≥ 5.0 μg/dl, 138 nmol/l) turned out to have the best SP (96.9%) and PPV (83.3%), as only one subject did not improve after surgery among the six patients with 1 mg-DST ≥ 5 μg/dl. Similarly, although excluded from the subsequent analyses due to low SP, Criterion IV (1 mg-DST ≥ 2.0 μg/dl, 55 nmol/l, cut-off obtained by the ROC analysis, see section Methods) showed the best SN (91.3%) and NPV (90.0%): among 20 patients with 1 mg-DST < 2.0 μg/dl only 2 subjects improved after surgery (Table 3).

Criteria I, II, and III remained statistically significant after adjusting for potential confounders, such as age, BMI, duration of the follow-up period, diameter of the adrenal mass, and the presence at the baseline of ≥ 2 of the considered endpoints (Table 4).

As in treated patients, 1 mg-DST < 2.0 μg/dl (55 nmol/l) and ≥ 5.0 μg/dl (138 nmol/l) showed a good NPV and PPV for excluding and confirming the SH respectively (Table 3), we analyzed the data of the 29 patients (out of 55 subjects) with 1 mg-DST ≥ 2.0 μg/dl (55 nmol/l) and < 5.0 μg/dl (138 nmol/l). In this subgroup, no biochemical criterion showed an acceptable accuracy (i.e. > 65%). However, the criterion characterized by the presence of MSC ≥ 4.0 μg/dl
(110 nmol/l) or ACTH ≤ 10.0 pg/ml (2.2 pmol/l) showed the best PPV (93.3%, accuracy 53.8%) in predicting the improvement of at least two endpoints in this subgroup. Therefore, in order to identify a criterion able to predict the improvement in this subgroup with a > 65% accuracy, we investigated the diagnostic accuracy of the different criteria of SH combined with the presence at the baseline of at least two metabolic consequences (i.e. obesity, type 2 diabetes, hypertension, and dyslipidemia) of SH in predicting the metabolic outcomes (Fig. 1). The presence of MSC ≥ 4.0 µg/dl (≥ 110 nmol/l) or ACTH ≤ 10.0 pg/ml (≤ 2.2 pmol/l) combined with the presence of at least two metabolic complications at the baseline (biochemical–clinical criterion) showed the best accuracy (75.9%) in predicting the improvement of ≥ 2 endpoints after surgery. Indeed, 11 out of the 13 subjects showing this biochemical–clinical criterion experienced the improvement of at least two endpoints after surgery (PPV 84.6%). On the other hand, among the 16 subjects without this biochemical–clinical criterion, 11 did not improve after surgery (NPV 84.6%). Using the protocol illustrated in Fig. 1, in 45 out of the 55 (81.2%) treated AI patients the improvement after surgery of ≥ 2 endpoints was correctly predicted before surgery.

Since the reasons for operation (altered parameters of the HPA axis activity or the size of the adenomas)
Table 5 Baseline clinical data of AI patients operated on for the dimensions of the adenoma or for high likelihood of subclinical hypercortisolism before surgery. Data are expressed as mean±S.D. with range in parenthesis or as absolute numbers with percentage in parenthesis.

<table>
<thead>
<tr>
<th></th>
<th>All patients (n=55)</th>
<th>Operated on for dimensions (n=41)</th>
<th>Operated on for high likelihood of SH (n=14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (F/M)</td>
<td>43/12</td>
<td>32/9</td>
<td>11/3</td>
</tr>
<tr>
<td>Age (years)</td>
<td>56.0±11.2 (24–75)</td>
<td>56.8±10.9 (24–75)</td>
<td>53.5±12.3 (33–70)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>28.1±5.6 (19–40)</td>
<td>27.4±5.2 (19–40)</td>
<td>30.2±6.5 (22.3–39.1)</td>
</tr>
<tr>
<td>Waist circumference (cm)</td>
<td>92.1±12.9 (64–120)</td>
<td>91.0±12.8 (64.0–120.0)</td>
<td>95.4±13.8 (77.0–118.0)</td>
</tr>
<tr>
<td>Diameter of the tumor (cm)</td>
<td>3.4±1.2 (1.2–9.0)</td>
<td>3.7±1.2* (2.5–9.0)</td>
<td>2.6±0.6 (1.2–3.4)</td>
</tr>
<tr>
<td>Follow-up (months)</td>
<td>29.6±13.8 (18–54)</td>
<td>30.4±13.9 (18–54)</td>
<td>26.9±13.7 (18–54)</td>
</tr>
<tr>
<td>Obese patients, n (%)</td>
<td>20 (36.4)</td>
<td>13 (31.7)</td>
<td>7 (50.0)</td>
</tr>
<tr>
<td>Hypertensive patients, n (%)</td>
<td>30 (54.5)</td>
<td>24 (58.5)</td>
<td>6 (42.9)</td>
</tr>
<tr>
<td>Diabetic patients, n (%)</td>
<td>11 (20)</td>
<td>6 (14.6)</td>
<td>5 (35.7)</td>
</tr>
<tr>
<td>Dyslipidemic patients, n (%)</td>
<td>21 (38.2)</td>
<td>16 (39.0)</td>
<td>5 (35.7)</td>
</tr>
<tr>
<td>ACTH (pg/ml)</td>
<td>9.8±5.1 (1.0–26.9)</td>
<td>10.4±5.6 (2.0–26.9)</td>
<td>8.1±3.2 (5.0–14.8)</td>
</tr>
<tr>
<td>1 mg-DST (µg/dl)</td>
<td>2.8±1.3 (0.9–5.4)</td>
<td>2.4±1.1† (0.9–5.1)</td>
<td>3.9±1.1 (1.8–5.4)</td>
</tr>
<tr>
<td>UFC (µg/24 h)</td>
<td>52.4±31.3 (10.0–135.5)</td>
<td>40.8±19.2† (10.0–96.0)</td>
<td>80.1±37.2 (18.8–135.5)</td>
</tr>
<tr>
<td>MSC (µg/dl)</td>
<td>4.9±3.2 (1.0–13.5)</td>
<td>3.4±1.4† (1.0–12.0)</td>
<td>7.0±4.0 (1.2–13.5)</td>
</tr>
</tbody>
</table>

SH, subclinical hypercortisolism; AI, adrenal incidentaloma; ACTH (SI conversion factor×0.22); 1 mg-DST, cortisol after 1 mg overnight dexamethasone-suppression test (SI conversion factor×27.56); UFC, urinary free cortisol (SI conversion factor×2.756); MSC, midnight serum cortisol (SI conversion factor×27.56). Patients were considered with high likelihood of SH if showing at least three out of UFC levels >70 µg/24 h (193 nmol/24 h), 1 mg-DST >3.0 µg/dl (83 nmol/l), ACTH levels <10 pg/ml (2.2 pmol/l), MSC >5.4 µg/dl (149 nmol/l). ACTH, 1 mg-DST, UFC, and MSC were determined at the baseline. *P=0.003, †P<0.0001, and ‡P=0.004 versus patients operated on for high likelihood of SH.

Discussion

This study was designed to evaluate in AI patients before surgery the accuracy of several combined parameters of the HPA axis activity in predicting the modifications of some metabolic indexes (i.e. body weight, blood pressure, fasting glucose, and LDL cholesterol levels), taken as endpoints, after the excision of the adrenal adenoma. We found that the presence before surgery of Criterion I (two out of 1 mg-DST >3.0 µg/dl, UFC >70.0 µg/24 h, and ACTH <10.0 pg/ml; UFC–ACTH–DST criterion) or Criterion III (1 mg-DST >2.0 µg/dl plus MSC >4.0 µg/dl) seemed able to predict the improvement of ≥2 endpoints after the excision of the adrenal adenoma, regardless of age, size of the adenoma, and the presence of obesity, hypertension, diabetes mellitus, and dyslipidemia at the baseline. The UFC–ACTH–DST criterion seemed also to be predictive of the worsening of ≥2 endpoints in nontreated AI patients. Finally, the present data suggest that the improvement of ≥2 endpoints after surgery is unlikely in patients showing before surgery 1 mg-DST <2.0 µg/dl (<55 nmol/l), whereas it is probable in patients showing before surgery 1 mg-DST ≥5.0 µg/dl (≥138 nmol/l). In patients with 1 mg-DST between these cut-offs, the presence of ACTH ≤10.0 pg/ml (2.2 pmol/l) or MSC >4.0 µg/dl (110 nmol/l) combined with the presence of ≥2 of the following: obesity, hypertension, type 2 diabetes, and dyslipidemia at the baseline seems the best criterion for identifying subjects that may benefit from surgery.

Recently, several studies suggested that the surgical approach in AI patients with SH may lead to recovery and/or improvement of some features of the metabolic syndrome, particularly, type 2 diabetes, hypertension.
and obesity (18–25). However, due to the wide differences in the criteria used for defining SH among the different studies, it is not known whether the beneficial effect of the recovery from SH on these endpoints can be predicted by the hormonal assessment before surgery.

On the other hand, it is known that the diagnosis of SH is debated and several criteria have been proposed (26, 27). Recently, a cross-sectional study specifically designed to validate the SH criteria on clinical ground suggested that the UFC–ACTH–DST criterion is the best in predicting the presence of chronic manifestations of subtle cortisol excess (28). This study confirms that previous finding, since the UFC–ACTH–DST criterion was shown to be the best one in predicting the improvement of ≥2 metabolic endpoints in surgically treated patients. It is important to note that this association is present regardless of the presence of the metabolic complications before surgery. It must be considered, however, that these metabolic consequences are not specific of the condition of SH. For this reason, even if the improvement of at least one co-morbidity may be enough to consider surgery worthwhile, we decided to consider the improvement of at least one co-morbidity as the hallmark of the presence of SH before surgery, in order to strengthen the diagnostic accuracy of the HPA axis parameters for suspecting SH. Indeed, as shown in the Supplementary Table using the improvement of at least one co-morbidity as the reference endpoint for suspecting SH, the NPVs of the biochemical criteria for SH were rarely acceptable. For these reasons and on the basis of the present findings, the diagnosis of SH probably should not rely only on the improvement of already present morbidities, since even in patients with clinical CS the metabolic alterations are not uniformly present. Moreover, it must also be noted that not all possible consequences of SH have been considered in this study (i.e. bone involvement) and not all parameters of metabolic consequences have been evaluated (i.e. insulin resistance and HbAlc). Therefore, a more comprehensive study with a similar design could give more conclusive data in order to obtain a better definition of SH.

The strength of this study is that it further validates the most accurate SH criteria, testing them on a sample of nontreated AI patients, in order to look for the accuracy of these criteria in predicting the worsening of the metabolic endpoints in nontreated AI patients. The finding that the UFC–ACTH–DST criterion is also able to predict the worsening of two metabolic endpoints in nontreated patients further reinforces the idea that this combination of parameters may be considered the best way to suspect SH. Moreover, it is important to note that both the presence of this criterion and the presence of complications before surgery are independent predictors of the amelioration of the metabolic parameters after adrenalectomy (Table 4). On the other hand, looking into those patients who were not operated (Table 6), the worsening of the metabolic parameters is associated only with the presence of this criterion but not with the presence of the metabolic complications. Taken together, these data may suggest that diagnosing SH may have a definite impact on the decision for surgery. Unfortunately, an age-, BMI-, and gender-matched

### Table 6: Odds ratio of the most accurate diagnostic criteria of subclinical hypercortisolism for predicting, in followed up patients, the worsening of at least two among body weight, blood pressure, fasting glucose, and low-density lipoprotein levels, adjusted for confounding factors.

<table>
<thead>
<tr>
<th>Criteria and confounders</th>
<th>OR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: 2 out of DST &gt; 3.0 μg/dl, UFC &gt; 70.0 μg/24 h, ACTH &lt; 10 pg/ml</td>
<td>5.81</td>
<td>1.44–23.43</td>
<td>0.013</td>
</tr>
<tr>
<td>Age (1 year increase)</td>
<td>1.02</td>
<td>0.95–1.09</td>
<td>0.664</td>
</tr>
<tr>
<td>BMI (1 kg/m² increase)</td>
<td>1.08</td>
<td>0.92–1.27</td>
<td>0.325</td>
</tr>
<tr>
<td>Presence of at least 2 out of O, AH, T2DM, and DL</td>
<td>1.85</td>
<td>0.36–9.6</td>
<td>0.463</td>
</tr>
<tr>
<td>Duration of follow-up (1 month increase)</td>
<td>1.02</td>
<td>0.96–1.08</td>
<td>0.613</td>
</tr>
<tr>
<td>Diameter of the adenoma (1 cm increase)</td>
<td>1.15</td>
<td>0.53–2.52</td>
<td>0.725</td>
</tr>
<tr>
<td>II: 2 out of DST &gt; 1.8 μg/dl, UFC &gt; 70.0 μg/24 h, MSC &gt; 4.0 μg/dl</td>
<td>4.90</td>
<td>1.20–19.80</td>
<td>0.026</td>
</tr>
<tr>
<td>Age (1 year increase)</td>
<td>1.02</td>
<td>0.96–1.09</td>
<td>0.559</td>
</tr>
<tr>
<td>BMI (1 kg/m² increase)</td>
<td>1.01</td>
<td>0.94–1.28</td>
<td>0.233</td>
</tr>
<tr>
<td>Presence of at least 2 out of O, AH, T2DM, and DL</td>
<td>1.79</td>
<td>0.35–9.0</td>
<td>0.468</td>
</tr>
<tr>
<td>Duration of follow-up (1 month decrease)</td>
<td>1.01</td>
<td>0.92–1.08</td>
<td>0.669</td>
</tr>
<tr>
<td>Diameter of the adenoma (1 cm increase)</td>
<td>1.03</td>
<td>0.49–2.35</td>
<td>0.851</td>
</tr>
<tr>
<td>III: DST &gt; 2.0 μg/dl* and MSC &gt; 4.0 μg/dl*</td>
<td>2.10</td>
<td>0.21–5.88</td>
<td>0.904</td>
</tr>
<tr>
<td>Age (1 year increase)</td>
<td>1.04</td>
<td>0.96–1.12</td>
<td>0.339</td>
</tr>
<tr>
<td>BMI (1 kg/m² increase)</td>
<td>1.04</td>
<td>0.96–1.31</td>
<td>0.135</td>
</tr>
<tr>
<td>Presence of at least 2 out of O, AH, T2DM, and DL</td>
<td>1.35</td>
<td>0.30–5.98</td>
<td>0.696</td>
</tr>
<tr>
<td>Duration of follow-up (1 month decrease)</td>
<td>1.01</td>
<td>0.96–1.07</td>
<td>0.782</td>
</tr>
<tr>
<td>Diameter of the adenoma (1 cm increase)</td>
<td>1.13</td>
<td>0.51–2.51</td>
<td>0.77</td>
</tr>
</tbody>
</table>

*Cut-off obtained by ROC analysis (see section Methods).
control group of subjects without AI was not available, thus reducing the evidence of the worsening in the nonoperated group. This study may also help to define the most cost-effective approach for SH identification in patients with AI. Indeed, it confirms that 1 mg-DST may be used as a first-line test as it correlates with the modification of the metabolic outcomes after surgery. Patients with 1 mg-DST < 2.0 or ≥ 5.0 µg/dl have an extremely low or high probability of improvement of ≥ 2 metabolic endpoints after surgery respectively. In patients with 1 mg-DST between these cut-offs, the combination of clinical data and further biochemical tests is required. The presence of MSC ≥ 4.0 µg/dl or ACTH ≤ 10.0 pg/ml combined with the presence of ≥ 2 features of the metabolic syndrome may suggest the usefulness of the surgical approach. Indeed in our study, three out of four patients with this combination experienced improvement of ≥ 2 metabolic endpoints after surgery. With this diagnostic protocol in 81.8% of AI patients, the metabolic outcome after surgery could be correctly predicted before surgery (Fig. 1). However, it must be underlined that we arbitrarily settled a relevant SN and SP to 65% and that the PPVs were not satisfying (with the exception of Criterion V). Furthermore, an important limit of this protocol is that MSC determination requires hospital admission, and, therefore, may be difficult to obtain in the general practice (36). In addition, the midnight salivary cortisol is not considered accurate for diagnosing SH (37).

A limitation of this study is also related to the characteristics of the condition of subtle cortisol excess per se. Indeed, the present data clearly shows that, nowadays, no criterion has enough diagnostic accuracy to be used as a single marker of SH in patients with AI. This was also indirectly found in previous studies showing that the improvement after surgery of some features of metabolic syndrome (i.e. type 2 diabetes and hypertension) was also present in patients without a diagnosis of SH before surgery, therefore suggesting that the currently used SH criteria, though specific, are not sufficiently sensitive (18, 19, 21). This limitation in the diagnosis of SH may be due to the fact that in AI patients cortisol secretion is a continuum and highly variable (36, 38). Therefore, diagnosing SH by arbitrary cut-offs of indexes of cortisol secretion leads to unavoidable mistakes in classifying some patients. These difficulties in diagnosing SH are further overdrawn by the low reliability of almost all markers of the HPA axis activity, particularly ACTH and UFC, used so far (27, 29, 38) and by the possible influence of several widely used concomitant medications on the DST (30). Unfortunately, in this study, the dexamethasone levels were not obtained. Since some patients were affected with diabetes, it is not possible to exclude that the classical 2 day DST could have offered a higher SP, even if the overall evidence in adults indicates that this test has similar or slightly less diagnostic accuracy than the overnight DST (29). These difficulties in diagnosing SH explain the finding of a previous study (39) that the postoperative hypocortisolism cannot be preoperatively ruled out even in the presence of normal HPA parameters. In keeping with this, in this study, the duration of the steroid substitutive therapy was comparable between patients operated on for the size of the adenoma and for the presence of alteration of HPA axis activity before surgery. The results of this and the previous (39) study suggest that more reliable biochemical parameters of pituitary–adrenal axis activity are needed to better define SH.

The intrinsic limits of this study are related to its relatively small sample, and retrospectively to its randomized design. Indeed, some improvements attributed to surgery could have been related to nonspecific changes in medical treatment/lifestyle and not to previous surgery. Moreover, the reasons for which surgery had been performed might have introduced a bias. However, the predictive role of the statistically significant criteria for SH was confirmed even after adjustment for the size of the adenoma. More importantly, when we repeated the analysis in only the subgroup of patients operated on for the size of the adenoma only, we confirmed the results obtained in the whole group. Finally, the criteria able to predict the improvement of the metabolic outcomes in treated patients were also found to be predictive of the worsening of the same outcomes in nontreated patients, in whom the selection bias was not present. Therefore, it seems unlikely that this selection bias might have had a crucial influence. However, a prospective randomized controlled trial is certainly needed to further clarify these issues.

In conclusion, this study shows that i) the UFC–ACTH–DST criterion seems to be the best in predicting before surgery the improvement of ≥ 2 out of body weight, arterial hypertension, fasting glucose, and cholesterol levels after the excision of the adrenal adenoma; ii) this criterion seems to be able also to predict the worsening of > 2 of these metabolic endpoints in nontreated AI patients; and iii) AI patients with 1 mg-DST < 2.0 (< 55 nmol/l) or ≥ 5.0 µg/dl (> 138 nmol/l) before surgery have low or high probability of improvement of ≥ 2 metabolic endpoints after surgery respectively; when 1 mg-DST is between these cut-offs, the presence of MSC ≥ 4.0 µg/dl (> 110 nmol/l) or ACTH ≤ 10.0 pg/ml (< 2.2 pmol/l) combined with the presence of ≥ 2 features of the metabolic syndrome may suggest the usefulness of the surgical approach.

Supplementary data
This is linked to the online version of the paper at http://dx.doi.org/10.1530/EJE-10-0602.
Declaration of interest
The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding
This work was partially supported by grants from Ministero dell’Istruzione Università e Ricerca di Italy (PRIN 2005 to B Ambrosi) and from Fondazione IRCCS Ca’ Granda – Ospedale Maggiore Policlinico (Ricerca Corrente 2008 to P Beck-Peccoz).

References
26 Stewart PM. Is subclinical Cushing’s syndrome an entity or a statistical fallout from diagnostic testing? Consensus surrounding the diagnosis is required before optimal treatment can be defined. Journal of Clinical Endocrinology and Metabolism 2010 95 2618–2620. (doi:10.1210/jc.2010-0633)


Received 28 September 2010
Accepted 29 September 2010

www.eje-online.org