A Probabilistic Model for Cushing's Syndrome Screening in At-Risk Populations: A Prospective Multicenter Study

Antonio León-Justel, Ainara Madrazo-Atutxa, Ana I. Alvarez-Rios, Rocio Infantes-Fontán, Juan A. Garcia-Arnés, Juan A. Lillo-Muñoz, Anna Aulinas, Eulàlia Urgell-Rull, Mauro Boronat, Ana Sánchez-de-Abajo, Carmen Fajardo-Montañana, Mario Ortuño-Alonso, Isabel Salinas-Vert, Maria L. Granada, David A. Cano, and Alfonso Leal-Cerro, for the Spanish CRISALIDA Study Group*

Context: Cushing's syndrome (CS) is challenging to diagnose. Increased prevalence of CS in specific patient populations has been reported, but routine screening for CS remains questionable. To decrease the diagnostic delay and improve disease outcomes, simple new screening methods for CS in at-risk populations are needed.

Objective: To develop and validate a simple scoring system to predict CS based on clinical signs and an easy-to-use biochemical test.

Design: Observational, prospective, multicenter.

Setting: Referral hospital.

Patients: A cohort of 353 patients attending endocrinology units for outpatient visits.

Interventions: All patients were evaluated with late-night salivary cortisol (LNSC) and a low-dose dexamethasone suppression test for CS.

Main Outcome Measures: Diagnosis or exclusion of CS.

Results: Twenty-six cases of CS were diagnosed in the cohort. A risk scoring system was developed by logistic regression analysis, and cutoff values were derived from a receiver operating characteristic curve. This risk score included clinical signs and symptoms (muscular atrophy, osteoporosis, and dorsocervical fat pad) and LNSC levels. The estimated area under the receiver operating characteristic curve was 0.93, with a sensitivity of 96.2% and specificity of 82.9%.

Conclusions: We developed a risk score to predict CS in an at-risk population. This score may help to identify at-risk patients in non-endocrinological settings such as primary care, but external validation is warranted. **(*J Clin Endocrinol Metab* 101: 3747–3754, 2016)**

Cushing's syndrome (CS) is caused by prolonged exposure to excess glucocorticoids (1), which causes decreased quality of life (2, 3) and increased morbidity and mortality (1, 4, 5). CS is considered a rare disease (6, 7), but recent studies have suggested a higher prevalence in specific, at-risk populations (8) including patients with type 2 diabetes (9, 10), hypertension (11), and osteoporosis (12).

The diagnosis of CS poses a considerable challenge (1, 13, 14) because there are no pathognomonic symptoms or signs of CS, and most of the symptoms and signs of CS are common in the general population, including obesity, hypertension, bone loss, and diabetes. The clinical practice guidelines recommend relying on clinical suspicion and performing biochemical tests to confirm hypercortisolism, but it

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA

Received March 17, 2016. Accepted July 25, 2016.
First Published Online August 4, 2016

* Author affliations are shown at the bottom of the next page.

Abbreviations: AUC, area under the curve; CI, confidence interval; CS, Cushing's syndrome; DST, dexamethasone suppression test; IDI, integrated discrimination improvement; LNSC, late-night salivary cortisol; LR, likelihood ratio; NRI, net reclassification improvement; ROC, receiver operator characteristic.

is not clearly stated which patients should be screened and what screening test should be performed (14).

The laboratory tests most commonly used for CS diagnosis are 24-hour urine-free cortisol, serum cortisol after low-dose (1 mg) dexamethasone suppression test (DST), late-night salivary cortisol (LNSC), and midnight serum cortisol (14). The Endocrine Society recommends using at least two of these tests to diagnose hypercortisolism (14). However, these biochemical tests remain impractical for screening strategies and may result in an excessive number of false positives. Thus, a systematic approach to test patients for CS remains an unmet need in clinical practice. Ideally, screening tests should be performed in nonspecialist settings such as primary care clinics to avoid unnecessary testing. To this end, the use of a noninvasive, easy-to-use, inexpensive test is of paramount importance. LNSC has recently gained popularity due to its high diagnostic sensitivity and specificity, noninvasive sample collection, and cost effectiveness (15–18).

The aim of this study was to develop and internally validate a screening scoring system able to predict CS in at-risk populations. The predictive model we propose is based on the assessment of clinical symptoms and signs and the use of LNSC.

Subjects and Methods

Study design

The CRISALIDA (*Cri*bado en *Sali*va *d*e *A*lteraciones de Cortisol [screening for cortisol alterations with salivary samples]) study was a prospective multicenter project conducted in 13 university hospitals in Spain under the auspices of the Spanish Society of Endocrinology and Nutrition. The study protocol was approved by the ethics committees of the participating hospitals. Written informed consent was obtained from each participant. A total of 389 patients attending endocrinology units between January 2012 and July 2013 were screened. Subjects with at least two features compatible with CS and a willingness to return for follow-up were invited to participate. These CS features included: obesity (body mass index > 30 kg/m^2), poorly controlled blood pressure (patients treated with more than two drugs and systolic blood pressure > 140 mm Hg and/or diastolic blood pressure > 90 mm Hg), uncontrolled diabetes (glycosylated hemoglobin $> 7.0\%$), virilization syndrome (hirsutism) with menstrual disorders, and osteoporosis (T-score ≥ -2.5 SD). Exclusion criteria included: steroid treatments, severe psychiatric illness (such as schizophrenia and dementia), kidney disease (estimated glomerular filtration rate < 60 mL/min/1.73 m^2), liver disease, and treatment with drugs that may affect cortisol metabolism. All patients were referred from either primary care or specialist consults such as hypertension and rheumatology clinics to endocrinology units for reasons related to poor control of metabolic alterations such as obesity, diabetes, and hypertension or osteoporosis of unknown causes (for the rheumatology clinic). No patients were referred with suspicion of CS. Patients were clinically examined, and the presence or absence of dorsocervical fat pad, purple striae, and proximal muscle weakness was recorded. History examination included psychiatric manifestations (other than those considered exclusion criteria) including depression and anxiety and cerebrovascular, cardiovascular, and respiratory diseases.

Methods

The flowchart for the study is depicted in Supplemental Figure 1. All selected patients underwent LNSC and 1-mg DST tests in outpatient settings. At 11 PM, saliva was collected using a Salivette swab (Sarstedt), and 1 mg of dexamethasone was orally administered. Patients did not eat, drink, or smoke for at least 2 hours before saliva collection. Blood samples were collected the following morning (8 to 9 AM). Cortisol levels were measured using a chemiluminescence method (E170; Roche Diagnostics). The intra-assay coefficient of variation (n = 21) was 5.12% at 4.88 (SD = 0.25) nmol/L, 2.1% at 14.9 (0.32) nmol/L, and 1.3% at 29.4 (0.4) nmol/L. Interassay (total) coefficient of variation (n = 45) was 8.3% at 2.9 (0.24) nmol/L and 3.8% at 31.2 (1.2) nmol/L (16). Subjects were classified as: negative for hypercortisolism (LNSC ≤ 7.5 nmol/L, and DST ≤ 50 nmol/L) or positive for hypercortisolism (LNSC > 7.5 nmol/L, and DST > 50 nmol/L). Patients with discordant results (LNSC ≤ 7.5 nmol/L and DST > 50 nmol/L or LNSC > 7.5 nmol/L and DST ≤ 50 nmol/L) were followed until the end of the study (December 2014), ie, at least 1.5 years (mean, 22.2 months; SD, 5.1). If convincing progression of CS signs and symptoms was observed during this follow-up (eg, new onset of facial plethora, muscle atrophy, or purple striae), patients were biochemically reevaluated; otherwise, patients were classified as subjects without hypercortisolism. For the definitive diagnosis of CS, clinicians followed the same algorithm that they used in their usual clinical practice in each center. Briefly, it involved performing an additional biochemical test (other than LNSC and DST) such as urine-free cortisol or 8-mg DST for confirmation of hypercortisolism. ACTH determinations were performed to determine whether pituitary magnetic resonance imaging or adrenal computerized tomography should be performed. Histological data were available for the three ectopic adenomas, six adrenal adenomas, and 14 (out of 17) pituitary adenomas.

Medicine Department (A.L.-J.), Huelva University Hospital, Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/Consejo Superior de Investigaciones Científicas/Universidad de Sevilla, 41013 Sevilla, Spain; Unidad de Gestión Clínica de Endocrinología y Nutrición (A.M.-A., D.A.C., A.L.-C.), IBiS, Hospital Universitario Virgen del Rocío/Consejo Superior de Investigaciones Científicas/Universidad de Sevilla, 41013 Sevilla, Spain; Department of Clinical Biochemistry (A.I.A.-R.), Virgen del Rocío University Hospital (IBiS/CSIC/SAS/University of Seville), 41013 Sevilla, Spain; Servicio de Bioquímica (R.I.-F.), Sección Hormonas especiales, Hospital Universitario Virgen del Rocio, 41013 Sevilla, Spain; Department of Clinical Endocrinology and Nutrition (J.A.G.-A.), Carlos Haya Hospital, 29010 Málaga, Spain; Hospital Regional Universitario de Málaga (J.A.L.-M.), 29010 Málaga, Spain; Pituitary Disease Research Group/Department Endocrinology/Medicine (A.A.), Hospital Sant Pau, Universitat Autónoma de Barcelona and CIBERER U747, ISCIII, 08025 Bellaterra, Barcelona, Spain; Clinical Biochemistry Department (E.U.-R.), Hospital de Sant Pau, 08025 Barcelona, Spain; Sección de Endocrinología y Nutrición (M.B.), Hospital Universitario Insular, Instituto Universitario de Investigaciones Biomédicas y Sanitarias, Universidad de Las Palmas de Gran Canaria, 35001 Las Palmas de Gran Canaria, Spain; Servicio de Bioquímica Clínica (A.S.-d.-A.), Hospital Universitario Insular de Gran Canaria, 35016 Las Palmas de Gran Canaria, Spain; Hospital Universitario de La Ribera (C.F.-M., M.O.-A.), 46600 Alzira, Valencia, Spain; Servicio Endocrinología y Nutrición (I.S.-V.) and Servicio de Bioquímica (M.L.G.), Hospital Universitari Germans Trias i Pujol, 08916 Badalona, Barcelona, Spain

Statistical analysis

Normality of the data was tested using the Kolmogorov-Smirnov test. Differences were analyzed by the Mann-Whitney *U* test for nonparametric data. For categorical variables, the Pearson's χ^2 was used. A P value $< .05$ was considered significant. All variables with meaningful correlations (P value $< .15$) in the univariate analysis were included into a multivariate logistic regression model. We developed two different models to predict the risk of having CS. One assessed the association of clinical signs and symptoms with CS (clinical model). The other (clinical and LNSC model) included, in addition to clinical signs and symptoms, LNSC values. The predictive capacity of each model was calculated using receiver operator characteristic (ROC) curve analysis, which provides area under the curve (AUC) measures, 95% confidence interval (CI), and plotted coordinates (sensitivity and specificity). Positive and negative likelihood ratios (LRs) were also calculated. A weighted risk diagnostic score was constructed using the coefficients of the multivariate logistic regression model that were converted into scores, rounding the values to the nearest whole number. Adequate fit of the model was determined by the Hosmer-Lemeshow test of goodness of fit. Bootstrapping procedures were performed with 1000 randomly selected replicates. We calculated the integrated discrimination improvement (IDI) and net reclassification improvement (NRI) for the comparison between the combined clinical and LNSC model and LNSC alone as described (19). NRI was based on three a priori risk categories of CS: low (0–0.2), medium (0.201–0.5), and high (0.501–1). Statistical analyses were performed using IBM SPSS software (version 22; SPSS Inc).

Results

Characteristics of the patients

Of 389 subjects enrolled, 36 did not complete the study; the remaining 353 subjects form the basis of our study (Supplemental Figure 1). The characteristics of these subjects are shown in Table 1. Of the 353 subjects, 219 did not show any abnormal results in the biochemical tests and were not further studied. Thirty-five subjects showed abnormal results in both tests. Ninety-nine patients exhibited discordant results in the tests; ie, only one test, either LNSC or 1-mg DST was positive. These patients were followed until the end of the study for a definitive diagnosis. Of these 99 patients, seven showed features suggestive of CS and abnormal results on both biochemical tests at reevaluation. Thus, 42 of the 353 subjects were further evaluated. A definitive diagnosis of CS was established in 26 of them (7.4% of the overall cohort). Twenty were ACTH dependent (17 of pituitary origin, three ectopic), and six were of adrenal origin.

Table 1. Patient Characteristics and Univariate Associations With CS

Variables	Overall Cohort	Absence of CS	CS	OR (95% CI)	*P* Value
n	353	327	26		
Gender, % female	68.6	69.4	57.1	0.6 (0.26–1.35)	.21
Age, y	56 [45–63]	56 [45–63]	51 [39–65]	0.99 (0.96–1.02)	.46
BMI, kg/m^2	36.73 [32.97–41.02]	36.85 [33.00–41.08]	35.37 [30.41–37.64]	0.95 (0.89–1.02)	.16
Waist, cm	117 [105–125]	116 [105–125]	118 [98–124]	0.98 (0.94–1.01)	.21
SBP, mm Hg	140 [130–152]	140 [130–152]	144 [131–151]	1.01 (0.99–1.03)	.47
DBP, mm Hg	84 [75–90]	84 [75–90]	87 [75–96]	1.02 (0.99–1.05)	.11
HbA1c, %	7.2 [5.8–8.4]	7.4 [5.9–8.5]	5.6 [5.4–7.0]	1.61 (0.66–3.94)	.29
Obesity	95.5	96.3	84.6	**0.21 (0.06–0.7)**	**.011**
Type 2 diabetes	64.3	66.7	34.6	**0.26 (0.11–0.61)**	**.002**
Hypertension	78.2	78.0	80.8	1.19 (0.43–3.2)	.74
Virilization syndrome (hirsutism)	21.5	20.8	30.8	1.69 (0.71–4.06)	.24
Osteoporosis	7.4	6.1	23.1	**4.60 (1.66–12.75)**	**.003**
Dyslipidemia	64.0	65.4	46.2	0.45 (0.2–1.00)	.051
Cerebrovascular disease	4.0	4.3	0	0.953 (0.89–1.0)	.164
Cardiovascular disease	15.6	15.0	23.1	1.7 (0.6–4.4)	.278
Respiratory disease	17.3	17.1	19.2	1.15 (0.42–3.17)	.790
Psychiatric manifestations	15.6	15.9	11.5	0.68 (0.2–2.36)	.549
Carbohydrate metabolism disorders (except diabetes)	6.2	5.8	11.5	2.03 (0.56–7.38)	.281
Dorsocervical fat pad	23.8	21.7	50.0	**3.32 (1.48–7.5)**	**.004**
Purple striae	7.4	6.7	15.4	2.38 (0.75–7.52)	.139
Muscular atrophy	3.4	1.8	23.1	**15.2 (4.48–51.25)**	**<.001**
LNSC, nmol/L	5.84 [4.61–8.17]	5.71 [4.46–7.48]	12.15 [9.69–17.85]	**1.26 (1.13–1.39)**	**<.001**
DST, nmol/L	28.59 [20.66–40.83]	27.41 [20.33–36.41]	184.62 [115.97–461.38]	**1.06 (1.04–1.08)**	**<.001**

Abbreviations: BMI, body mass index; DBP diastolic blood pressure; HbA1c, glycosylated hemoglobin; IQR, interquartile range; SBP, systolic blood pressure. Data are presented as median (interquartile range) for continuous variables and as percentages for categorical variables. Data within brackets indicate interquartile range. Boldface data indicate variables that shows statistically significant differences between CS and non-CS patients.

Predictors of CS diagnosis

Univariate analysis revealed five clinical variables significantly associated with CS: muscular atrophy (odds ratio [OR], 15.2), followed by osteoporosis (OR, 4.6), dorsocervical fat pad (OR, 3.32), absence of obesity (OR, 0.21), and absence of type 2 diabetes (OR, 0.26) (Table 1). LNSC values were also significantly related with CS (OR, 1.26).

Development and validation of the model

A multivariate logistic regression analysis revealed that muscular atrophy (OR, 9.04), osteoporosis (OR, 3.62), and dorsocervical fat pad (OR, 3.3) remained as independent variables associated with CS (Table 2). ROC analysis was performed to evaluate the predictive power of this clinical-based model (Figure 1). The AUC of this model was 0.684 ($P = .002$), with a sensitivity of 57.7% and specificity of 72.5%. Positive and negative LRs were 2.1 and 0.58, respectively. These results indicate a moderate predictive performance of the model. We performed ROC analysis to evaluate the performance of LNSC to predict CS. A cutoff level of 9.17 nmol/L provided the best results, with an AUC of 0.893 ($P < .001$), a sensitivity of 88.5%, and specificity of 83.2% (Figure 1). Positive and negative LRs were 5.94 and 0.15, respectively. Next, we included the LNSC values in the clinical-based probabilistic model. This combined clinical and LNSC model showed an AUC of 0.916 ($P < .001$), a sensitivity of 92.3%, and specificity of 65.7% (Figure 1). Positive and negative LRs were 5.8 and 0.08, respectively. The Nagelkerke R Square in the combined clinical and LNSC model increased, and the Hosmer-Lemeshow test indicated a good fitness of the model (Table 2). The clinical and LNSC model had a significant increase in AUC compared to the clinical model as well as to the LNSC test (Figure 1). This prediction model allowed calculation of the probability of risk for CS according to the following formula: Probability (P) = $(1 + e^{-z})$ $\exp^{-1}$, in which z = 1.38 × osteoporosis (yes = 1, no = 0) + 1.32 × dorsocervical fat pad (yes = 1, no = 0) + 2.44 × muscular atrophy (yes = 1, no = 0) + 0.25 × LNSC − 5.02 (Table 2). The diagnostic prediction model was internally validated with bootstrapping techniques. The mean ORs obtained with the bootstrap analysis were similar to those observed in the original model (Table 3). The AUC from the bootstrap analysis with the combined clinical variables and LNSC was 0.916. The prognostic score was calibrated according to the Hosmer-Lemeshow goodness-of-fit test ($P = .88$) and exhibited excellent overall performance (Supplemental Figure 2). We also calculated the IDI and NRI statistics. The IDI was 0.14 (95% CI, 0.031–0.239; $P = .013$), and the NRI was 0.39 (39% of improvement) ($P = .02$), further confirming that the

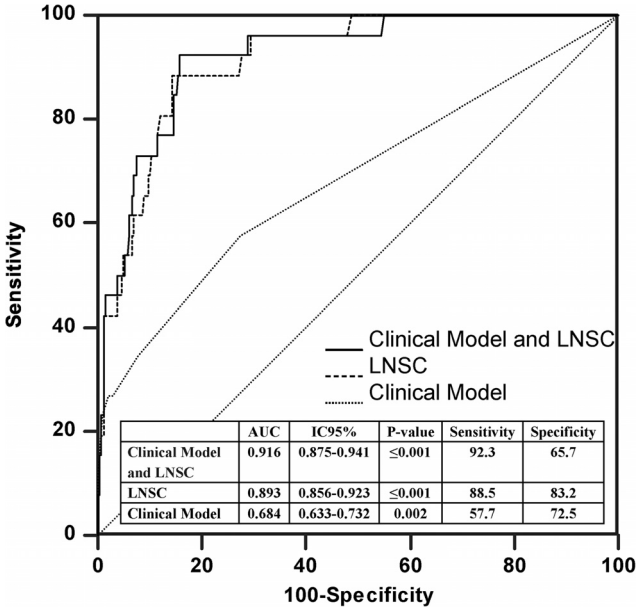

Figure 1. ROC curve of the clinical model, LNSC test, and the clinical model and LNSC in diagnosing CS.

Table 2. Association Between Variables and CS as Determined by Multivariate Logistic Regression Analysis of the Clinical Model and the Combined Clinical and LNSC Model

Clinical Model				Clinical and LNSC Model			
Variables	***B*[a]**	**OR (95% CI)**	***P* Value**	**Variables**	***B***	**OR (95% CI)**	***P* Value**
Osteoporosis	1.29	3.62 (1.16–11.35)	.027	Osteoporosis	1.38	3.97 (1.24–12.75)	.021
Dorsocervical fat pad	1.2	3.3 (1.52–7.17)	.003	Dorsocervical fat pad	1.32	3.75 (1.13–12.44)	.005
Muscular atrophy	2.2	9.04 (2.36–34.65)	<.001	Muscular atrophy	2.44	11.49 (2.6–50.69)	<.001
				LNSC, nmol/L	0.25	1.29 (1.16–1.44)	<.001
Constant	−3.19		<.001	Constant	−5.03		<.001
Nagelkerke R square: 0.167				Nagelkerke R square: 0.387			
P Hosmer-Lemeshow test: 0.618				P Hosmer-Lemeshow: 0.648			

Probabilistic clinical model: probability (P) = $(1 + e^{-z})^{-1}$ in which z = (1.29 × osteoporosis) + (1.2 × dorsocervical fat pad) + (2.2 × muscular atrophy) − 3.19. Probabilistic clinical and LNSC model: P = $(1 + e^{-z})^{-1}$ in which z = (1.38 × osteoporosis) + (1.32 × dorsocervical fat pad) + (2.44 × muscular atrophy) + (0.25 LNSC) − 5.03.

[a] Regression coefficients.

Table 3. Bootstrap Analysis of the Combined Clinical and LNSC Model

	Bootstrap Analysis[a]		
Clinical Variables	Mean OR	95% CI	*P* Value
Osteoporosis	4.21	1.39–12.9	.026
Dorsocervical fat pad	3.66	1.06–12.65	.010
Muscular atrophy	11.99	4.17–34.52	<.001
LNSC	1.31	1.19–1.43	<.001

[a] 1000 samples.

combination of clinical variables and LNSC would improve the predictive ability of the model.

To facilitate the use of the model in clinical practice, a scoring system was developed based on this prediction model using the coefficients of the multivariate logistic regression model and rounding (Table 4). The risk score of CS for individual patients was calculated using the following formula: 2 × presence of osteoporosis (yes = 1, no = 0) + 2 × dorsocervical fat pad (yes = 1, no = 0) + 3 × muscular atrophy (yes = 1, no = 0) + LNSC levels (low = 0, medium = 4, high = 5) (Table 4). ROC analysis showed that a score threshold of 4 (score ≥ 4) resulted in an AUC of 0.93 ($P < .001$), with sensitivity and specificity of 96.2 and 82.9%, respectively. Positive and negative LRs were 5.61 and 0.05, respectively. The number of subjects per score category is shown in Table 5. Selecting this cutoff value of 4, 271 of 327 (83%) subjects without CS were correctly identified whereas only one of 26 CS cases was missed. Our model yielded 56 false positives. Most of them (n = 36) were included due to medium levels of LNSC without any other clinical features included in our model (Supplemental Table 1). Indeed, the LNSC levels of these 56 false-positive cases were significantly increased compared to the non-CS group.

Table 4. Independent Diagnostic Indicators and Risk Score for CS

Variables	Regression Coefficient	*P* Value	Score Points
Osteoporosis	1.53	.004	2
Dorsocervical fat pad	1.81	.001	2
Muscular atrophy	3.4	<.001	3
LNSC			
Medium, 9.17–13.93 nmol/L	3.68	<.001	4
High, ≥13.93 nmol/L	4.93	<.001	5

The final multivariate model is estimated after model validation and adjustment for overfitting. The weighted risk score was constructed using the coefficients of the multivariate logistic regression model that were converted into scores, rounding the values to the nearest whole number.

Table 5. Total Number of Subjects and Prevalence of CS Per Score Category Using the Scoring System Obtained From the Combined Clinical and LNSC Model

	CS	
Score	No	Yes
0	199	0
2	69	1
3	3	0
4	36	8
5	9	2
6	9	2
7	2	6
8	0	0
9	0	3
10	0	1
11	0	1
12	0	2
Total	327	26

Total numbers by category: false positives, 56; false negatives, 1; true positives, 25; and true negatives, 271.

Discussion

In this multicenter study, we developed and internally validated a multivariate prediction model for the diagnosis of CS in at-risk populations. This model allowed the generation of a score system that can be used easily by any physician to determine whether further evaluations are warranted for the diagnosis of CS.

Recent studies have reported a prevalence of CS in specific at-risk populations higher than initially thought (8). However, the results of these systematic screening studies show a broad range for the prevalence of occult CS (8–10, 20–23). The variability in the prevalence of CS likely reflects differences in the selection criteria of the patients but nevertheless point to a higher prevalence of this syndrome in specific populations. In our study, we screened patients with at least two of five nonspecific features of CS, including obesity, poorly controlled blood pressure, uncontrolled diabetes, virilization syndrome with menstrual disorders, and osteoporosis. We chose these features based on the previously reported increased prevalence of CS in patients with these pathologies but, more importantly, because they are not specific to CS, thus eliminating a potential selection bias. We found a 7.4% prevalence of CS in this at-risk population, a value within the range of those reported in previous studies (8) although relatively higher compared to most of these studies. This high prevalence could be due to the stringent inclusion criteria that contemplated multiple combinations of signs and symptoms. A higher prevalence of CS has been reported in patients with obesity, hypertension, and uncontrolled diabetes mellitus (9). In agreement with this notion, our study population exhibited an elevated rate of obesity, hyperten-

sion, and poorly controlled diabetes. Indeed, when we analyzed patients included only due to poorly controlled hypertension and obesity (in the absence of other inclusion criteria), we found a CS prevalence of 9% (eight of 87 patients). Thus, this population that is very prevalent in the general population (even if we considered only poorly controlled hypertension) might be considered an at-high-risk population.

Despite the reported increased prevalence of CS in specific populations, the implementation of general screening procedures for CS remains controversial (8). Several clinical guidelines suggest testing for CS in patients with multiple features compatible with CS, particularly those more discriminatory (14). However, these signs and symptoms of CS might not be obvious, particularly for physicians not familiar with CS. Also, CS patients often only display CS features that are common in the general population such as obesity, hypertension, and diabetes. According to the Endocrine Society, the presence of two abnormal results in biochemical tests with high diagnostic accuracy is required for the diagnosis of overt CS (14), but this strategy is not practical for CS screening in primary care context due to the excessive workload and associated costs. Decades ago, Nugent et al (24) proposed using clinical signs and symptoms to help in the decision of CS diagnosis, and this idea has recently received reappraisal (13). We have reevaluated this approach, trying to develop a diagnostic prediction model based only on clinical signs and symptoms technically easy to obtain in clinical practice. Univariate analysis revealed five clinical variables that significantly associated with CS: muscular atrophy, dorsocervical fat pad, osteoporosis, obesity, and type 2 diabetes. Obesity and type 2 diabetes displayed a negative association with CS. These results might seem paradoxical a priori, but we want to stress that in our analyzed cohort, the prevalence of obesity and diabetes was exceedingly high (likely reflecting the reasons for referral to endocrinology units). Indeed, the prevalence of obesity and diabetes in the CS group was within the range of previously reported rates (4, 7). Thus, for the multivariate regression analysis, we decided to include only the clinical signs showing positive association with CS. However, the specificity and sensitivity values obtained were not acceptable for screening purposes.

Because we aimed to develop a diagnostic model for CS that could be implemented in non-endocrinological settings, we decided to use LNSC due to its ease of sample collection (it can even be collected at home) and decreased costs while maintaining a high diagnostic sensitivity and specificity (15–18). First, we evaluated the diagnostic performance of LNSC. The optimal cutoff value for the diagnosis of CS was of 9.17 nmol/L, with 88.5% sensitivity and 83.2% specificity, results within the range of those reported in the literature (18) although slightly lower compared to pooled data from a recent meta-analysis (25). The 9.17 nmol/L threshold of our study is elevated compared to most of those described in previous studies (cutoff values ranged from 3.6 to 15.2 nmol/L) (25). This may be due to the prominent prevalence of obesity in our cohort, which has been shown to increase cortisol levels. Also, it is important to note that salivary cortisol increases with age, hypertension, and diabetes (26), conditions also very prevalent in our cohort. Although it was not the aim of our study, the data collected allowed us to evaluate the diagnostic performance of the DST. Using the recommended cutoff of 50 nmol/L (14), we obtained a sensitivity of 100% and a specificity of 91.4% (Supplemental Figure 3), indicating an excellent diagnostic performance, in agreement with similar recent studies (18).

In contrast to our results, a recent study has reported a low positive predictive value of LNSC to detect hypercortisolism in type 2 diabetes (27), although no definitive diagnosis of CS was established. The reason for these apparent discrepancies is unclear, but we would like to emphasize the differences between the studied populations and, more importantly, the need for exhaustive follow-up to make a definitive diagnosis of CS.

A statistically significant increase in diagnostic performance was found when we compared the combined clinical model plus LNSC with the LNSC test. It can be argued that despite this statistical significance, the increase in diagnostic performance was modest, but the use of the ROC area has been criticized (28). First, the AUC is an overall measure of discrimination and has no direct clinical interpretation in terms of correct or incorrect diagnostic classifications or absolute patient numbers because a specific diagnostic algorithm uses a specific diagnostic cutpoint. Second, researchers have observed that the increase in AUC is often very small in an absolute sense, certainly when the AUC of the baseline model is large (29). Also, it is important to note that we evaluated our diagnostic model prospectively in a high-risk population, under more stringent conditions than in a case-control study. NRI and IDI have been increasingly adopted in the last few years to quantify the improvement after adding a new variable to a preexisting predictive model (19, 28). Adding clinical variables to LNSC yielded an IDI of 0.14 and an NRI improvement of 39% that can be considered a good diagnostic improvement. Nevertheless, our results highlight that the combination of clinical variables and an easy-to-use biochemical test like LNSC improves the diagnostic performance in a high-risk population, ie, a population in which the diagnosis of CS is particularly challenging.

For clinical use, we developed a simplified scoring system that discriminated satisfactorily between subjects with low and high probabilities of having CS. Applying a cutoff score value of 4, 83% of the subjects without CS were correctly identified, whereas only one of 26 CS cases was missed. In other words, further evaluation tests would have been avoided in 272 of 353 patients with an accurate prediction in 271. Our model yielded an acceptable number of false-positive cases. The LNSC levels of the 56 false-positive cases were significantly increased compared to the non-CS group. Indeed, 64% of the false positives didn't present any typical CS features included in our model and were included due to having medium levels of LNSC. In particular, we found that the proportion of patients with muscle atrophy was very reduced in the false-positive group compared to the CS group. These results once again point to the important role of signs and symptoms in the screening process. Perhaps positive patients with lower scores could be more carefully evaluated for the presence of clinical features before proceeding to further evaluation for CS confirmation.

Our study has a number of strengths. First, the prospective design of our study involved an exhaustive follow-up of patients and resulted in a very reliable diagnosis of CS. Second, the relatively straightforward clinical variables used in our model can be easily acquired during a routine clinical evaluation. However, certain limitations of our study should also be considered. First, external validation of our model is mandatory because our results might not be generalizable to patients of other clinical settings. Second, the cost-effectiveness of a screening program based on our diagnostic model should be determined. Finally, we need to acknowledge that the use of the LNSC test is not widely implemented, particularly in primary care contexts. However, we hope that this study as well as recent studies (18) showing the good diagnostic performance of LNSC in CS screening could contribute to its evaluation and eventual implementation in primary care clinics.

In conclusion, we have developed a scoring system that can be confidently used to predict CS, thus avoiding unnecessary testing in at-risk populations. Although all the assessments were performed by specialists (endocrinologists) in our study, this scoring system could be easily tested in independent cohorts and different settings such as primary care or hypertension clinics. At the very least, our diagnostic prediction model could be used as a framework for future studies and potential improvements in diagnostic performance.

Acknowledgments

We particularly thank Cristina Meseguer (Novartis Oncology) for her support and encouragement throughout the study.

Address all correspondence and requests for reprints to: Alfonso Leal-Cerro, Unidad de Gestión Clínica de Endocrinología y Nutrición, Hospital Universitario Virgen del Rocío, Sevilla. Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/Consejo Superior de Investigaciones Científicas/Universidad de Sevilla, 41013 Sevilla, Spain. E-mail: aleal@neuroendocrinologia.net.

This study was supported by a research grant from Novartis Oncology, Spain. Supplies for salivary cortisol determinations were provided by Roche, Spain. The funders did not have any role in the study design, data collection and analysis, or preparation of the manuscript. The components of the Spanish CRISALIDA study group are listed in the Supplemental Data. D.A.C. is supported by the Nicolás Monardes Program of the Andalusian Ministry of Health, Spain (C-0015-2014).

Disclosure Summary: A.L.-J. and A.L.-C. received a research grant from Novartis Oncology, Spain, for this study. The other authors have nothing to declare.

References

1. **Lacroix A, Feelders RA, Stratakis CA, Nieman LK.** Cushing's syndrome. *Lancet*. 2015;386:913–927.
2. **Santos A, Crespo I, Aulinas A, Resmini E, Valassi E, Webb SM.** Quality of life in Cushing's syndrome. *Pituitary*. 2015;18:195–200.
3. **Lindsay JR, Nansel T, Baid S, Gumowski J, Nieman LK.** Long-term impaired quality of life in Cushing's syndrome despite initial improvement after surgical remission. *J Clin Endocrinol Metab*. 2006; 91:447–453.
4. **Ferraù F, Korbonits M.** Metabolic comorbidities in Cushing's syndrome. *Eur J Endocrinol*. 2015;173:M133–M157.
5. **Arnaldi G, Angeli A, Atkinson AB, et al.** Diagnosis and complications of Cushing's syndrome: a consensus statement. *J Clin Endocrinol Metab*. 2003;88:5593–5602.
6. **Steffensen C, Bak AM, Rubeck KZ, Jørgensen JO.** Epidemiology of Cushing's syndrome. *Neuroendocrinology*. 2010;92(suppl 1):1–5.
7. **Valassi E, Santos A, Yaneva M, et al.** The European Registry on Cushing's syndrome: 2-year experience. Baseline demographic and clinical characteristics. *Eur J Endocrinol*. 2011;165:383–392.
8. **Tabarin A, Perez P.** Pros and cons of screening for occult Cushing syndrome. *Nat Rev Endocrinol*. 2011;7:445–455.
9. **Catargi B, Rigalleau V, Poussin A, et al.** Occult Cushing's syndrome in type-2 diabetes. *J Clin Endocrinol Metab*. 2003;88:5808–5813.
10. **Terzolo M, Reimondo G, Chiodini I, et al.** Screening of Cushing's syndrome in outpatients with type 2 diabetes: results of a prospective multicentric study in Italy. *J Clin Endocrinol Metab*. 2012;97:3467–3475.
11. **Omura M, Saito J, Yamaguchi K, Kakuta Y, Nishikawa T.** Prospective study on the prevalence of secondary hypertension among hypertensive patients visiting a general outpatient clinic in Japan. *Hypertens Res*. 2004;27:193–202.
12. **Chiodini I, Mascia ML, Muscarella S, et al.** Subclinical hypercortisolism among outpatients referred for osteoporosis. *Ann Intern Med*. 2007;147:541–548.
13. **Nieman LK.** Cushing's syndrome: update on signs, symptoms and biochemical screening. *Eur J Endocrinol*. 2015;173:M33–M38.
14. **Nieman LK, Biller BM, Findling JW, et al.** The diagnosis of Cush-

ing's syndrome: an Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab*. 2008;93:1526–1540.
15. **Putignano P, Toja P, Dubini A, Pecori Giraldi F, Corsello SM, Cavagnini F.** Midnight salivary cortisol versus urinary free and midnight serum cortisol as screening tests for Cushing's syndrome. *J Clin Endocrinol Metab*. 2003;88:4153–4157.
16. **León-Justel A, Mangas MA, Infante Fontán R, et al.** Budget impact of using midnight salivary cortisol in the diagnosis of hypercortisolism. *Clin Chim Acta*. 2011;412:2248–2253.
17. **Elias PC, Martinez EZ, Barone BF, Mermejo LM, Castro M, Moreira AC.** Late-night salivary cortisol has a better performance than urinary free cortisol in the diagnosis of Cushing's syndrome. *J Clin Endocrinol Metab*. 2014;99:2045–2051.
18. **Ceccato F, Barbot M, Zilio M, et al.** Screening tests for Cushing's syndrome: urinary free cortisol role measured by LC-MS/MS. *J Clin Endocrinol Metab*. 2015;100:3856–3861.
19. **Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS.** Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008;27:157–172; discussion 207–212.
20. **Reimondo G, Pia A, Allasino B, et al.** Screening of Cushing's syndrome in adult patients with newly diagnosed diabetes mellitus. *Clin Endocrinol (Oxf)*. 2007;67:225–229.
21. **Budyal S, Jadhav SS, Kasaliwal R, et al.** Is it worthwhile to screen patients with type 2 diabetes mellitus for subclinical Cushing's syndrome? *Endocr Connect*. 2015;4:242–248.
22. **Ness-Abramof R, Nabriski D, Apovian CM, et al.** Overnight dexamethasone suppression test: a reliable screen for Cushing's syndrome in the obese. *Obes Res*. 2002;10:1217–1221.
23. **Baid SK, Rubino D, Sinaii N, Ramsey S, Frank A, Nieman LK.** Specificity of screening tests for Cushing's syndrome in an overweight and obese population. *J Clin Endocrinol Metab*. 2009;94:3857–3864.
24. **Nugent CA, Warner HR, Dunn JT, Tyler FH.** Probability theory in the diagnosis of Cushing's syndrome. *J Clin Endocrinol Metab*. 1964;24:621–627.
25. **Carroll T, Raff H, Findling JW.** Late-night salivary cortisol for the diagnosis of Cushing syndrome: a meta-analysis. *Endocr Pract*. 2009;15:335–342.
26. **Liu H, Bravata DM, Cabaccan J, Raff H, Ryzen E.** Elevated late-night salivary cortisol levels in elderly male type 2 diabetic veterans. *Clin Endocrinol (Oxf)*. 2005;63:642–649.
27. **Steffensen C, Thomsen HH, Dekkers OM, Christiansen JS, Rungby J, Jørgensen JO.** Low positive predictive value of midnight salivary cortisol measurement to detect hypercortisolism in type 2 diabetes. *Clin Endocrinol (Oxf)*. 2016;85:202–206.
28. **Cook NR.** Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007;115:928–935.
29. **Tzoulaki I, Liberopoulos G, Ioannidis JP.** Assessment of claims of improved prediction beyond the Framingham risk score. *JAMA*. 2009;302:2345–2352.