

RESEARCH

Low-dose short synacthen test with salivary cortisol in patients with suspected central adrenal insufficiency

Filippo Ceccato^{1,2,3}, Elisa Selmin¹, Giorgia Antonelli^{1,4}, Mattia Barbot^{1,2,3}, Andrea Daniele¹, Marco Boscaro¹, Mario Plebani^{1,4} and Carla Scaroni^{1,2}

¹Department of Medicine DIMED, University of Padova, Padova, Italy

²Endocrine Disease Unit, University-Hospital of Padova, Padova, Italy

³Department of Neuroscience DNS, University of Padova, Padova, Italy

⁴Laboratory Medicine, University-Hospital of Padova, Padova, Italy

Correspondence should be addressed to F Ceccato: filippo.ceccato@unipd.it

Abstract

Context: The low-dose short synacthen test (LDSST) is recommended for patients with suspected central adrenal insufficiency (AI) if their basal serum cortisol (F) levels are not indicative of an intact hypothalamic–pituitary–adrenal (HPA) axis.

Objective: To evaluate diagnostic threshold for salivary F before and 30 min after administering 1 µg of synacthen, performed before 09:30 h.

Design: A cross-sectional study from 2014 to 2020.

Setting: A tertiary referral university hospital.

Patients: In this study, 174 patients with suspected AI, 37 with central AI and 137 adrenal sufficient (AS), were included.

Main outcome measure: The diagnostic accuracy (sensitivity (SE), specificity (SP)) of serum and salivary F levels measured, respectively, by chemiluminescence immunoassay and liquid chromatography-tandem mass spectrometry.

Results: Low basal serum or salivary F levels could predict AI. For the LDSST, the best ROC-calculated threshold for serum F to differentiate AI from AS was 427 nmol/L (SE 79%, SP 89%), serum F > 500 nmol/L reached SP 100%. A salivary F peak > 12.1 nmol/L after administering synacthen reached SE 95% and SP 84% for diagnosing central AI, indicating a conclusive reduction in the likelihood of AI. This ROC-calculated threshold for salivary F was similar to the 2.5th percentile of patients with a normal HPA axis, so it was considered sufficient to exclude AI. Considering AS those patients with salivary F > 12.1 nmol/L after LDSST, we could avoid unnecessary glucocorticoid treatment: 99/150 subjects (66%) had an inadequate serum F peak after synacthen, but salivary F was >12.1 nmol/L in 79 cases, who could, therefore, be considered AS.

Conclusions: Salivary F levels > 12.1 nmol/L after synacthen administration can indicate an intact HPA axis in patients with an incomplete serum F response, avoiding the need to start glucocorticoid replacement treatment.

Key Words

- ▶ adrenal insufficiency
- ▶ glucocorticoid treatment
- ▶ salivary cortisol
- ▶ liquid chromatography-tandem mass spectrometry (LC-MS/MS)
- ▶ low-dose short synacthen test (LDSST)

Endocrine Connections
(2021) 10, 1189–1199

Introduction

The correct diagnosis of central adrenal insufficiency (AI) is a matter of debate (1), partly because an adequate glucocorticoid (GC) replacement therapy is life-saving, but inappropriate treatment is detrimental (2, 3). Signs and symptoms of AI are often non-specific (fatigue, orthostatic hypotension, nausea, vomiting), and a clinical suspicion needs to be confirmed by biochemical testing.

Baseline morning unstimulated serum cortisol (F) levels are measured in patients with suspected central AI, but the results are affected by variations in binding proteins, and most of the commercially available F assays are not very accurate in the low range of normality (1, 4, 5, 6). The Endocrine Society guidelines only confirm central AI in patients with very low morning basal serum F levels (≤ 83 nmol/L, 3 $\mu\text{g}/\text{dL}$), whereas only high serum F levels (≥ 415 nmol/L) can confirm a normal hypothalamic–pituitary–adrenal (HPA) axis (2). A confirmatory dynamic test is, therefore, required for serum F levels in the range of 83–415 nmol/L (2). The gold standard is the insulin tolerance test (ITT), but it is complicated to perform in the outpatient setting. It demands careful supervision and is not recommended in frail patients because of the possible side effects (7). In clinical practice, the corticotropin stimulation test is widely used to diagnose AI, although there is no consensus on the most suitable dosage (250 or 1 μg) and patient preparation, the timing of blood sampling after the injection (20, 30, or 60 min), or the cut-off for diagnosing AI (500 or 550 nmol/L) (4, 7). Regarding the measurement of serum cortisol levels, the use of modern immunoassays results in lower cortisol concentrations (8), because outdated radioimmunoassays were not able to differentiate compounds with structural similarity to the target molecule as cortisone (E) and F (9, 10). The use of liquid chromatography–tandem mass spectrometry (LC-MS/MS) is increasing, and reduced thresholds to define AI have been reported (8).

In recent years, measuring salivary F has been proposed for patients with hypothalamic–pituitary–adrenal (HPA) axis disease (11). Salivary F reflects serum-free F levels; an altered concentration of binding proteins minimally affects its diagnostic accuracy (12, 13), also in the case of women consuming oral estrogens and suspected hypercortisolism (14). 11 β -hydroxysteroid dehydrogenase type 2 (11 β -HSD2) is strongly expressed in the salivary glands and converts F to E, which has been suggested as a marker of treatment in patients with AI (15). Measurement of salivary F or E is suggested in patients with adrenal insufficiency (11), providing a non-invasive alternative to

serum cortisol levels (16), and recently it has been reported as an adjuvant tool to increase the diagnostic accuracy of the corticotropin stimulation test in adult or pediatric patients with AI (13, 17).

We examined a series of consecutive patients with suspected central AI using the low-dose (1 μg) short synacthen test (LDSST), measuring serum and salivary F at the baseline and 30 min after administering the synacthen. Our aim was to assess the value of salivary F as an adjunctive tool for distinguishing patients with a normal HPA axis from those with AI requiring lifelong GC replacement treatment.

Materials and methods

Patient selection and cortisol measurement

Using a dedicated query in the web-based Padova University Hospital database, we collected all consecutive LDSSTs performed from December 2014 to July 2019 ($n=270$). Combined serum and salivary results were available for 214 tests. After applying the selection criteria listed below and ensuring that a clinical follow-up consultation conducted at least 12 months after the baseline visit was available in all cases enrolled with suspected AI, 174 patients were included in our final analyses.

A LDSST was performed in patients with suspected central AI based on the following criteria:

- evidence of a sellar mass (a pituitary adenoma, with signs of compression/invasion of neighboring structures, or a sellar-parasellar lesion that might cause HPA axis damage);
- a history of pituitary/skull base surgery (at least 3 months before the suspicion of AI);
- a history of radiotherapy (RT, at least 3 months before the suspicion of AI);
- signs or symptoms consistent with AI: orthostatic hypotension (fall in systolic > 20 mmHg and diastolic > 10 mmHg within 3 min upon standing, without medical treatment), unexplained hyponatremia (< 134 nmol/L), unexplained hypoglycemia in patients not using anti-diabetic drugs, salt craving, fatigue;
- HPA axis suppression after remission of endogenous Cushing's syndrome (CS) or after withdrawal of exogenous GC.

After an accurate endocrine and clinical work-up, our cohort was divided into two groups based on their HPA axis and GC treatment:

- Patients with central AI: 37 patients who started chronic GC replacement therapy after their baseline assessment (morning basal unstimulated serum F level ≤ 83 nmol/L, endocrine examination of HPA axis, signs and symptoms of AI, and clinical history);
- Adrenal sufficient (AS) patients: 137 subjects not requiring GC supplementation. From a clinical perspective, we further divided AS subjects into two groups: 106 patients with a normal HPA axis, not requiring any GC treatment; 31 patients with relative adrenal sufficiency (RAS). These patients with RAS were characterized by normal or normal-to-low basal F levels, and they did not reach clearly sufficient F levels in the LDSST (>500 nmol/L). Therefore, their HPA axis was judged adequate for normal life activities, with some limitations in relation to stressful events, when GC treatment was suggested. They were well-educated individuals aware of their incomplete HPA axis response: they were advised to take GC therapy only in the event of illness, body temperature $> 38^{\circ}\text{C}$, major/minor surgery, endoscopic procedures, or other events that might precipitate an adrenal crisis (18). They were registered with a medical alert service and given a steroid alert card and scheduled for annual training sessions by nurses on how to manage their daily medication and any minor or moderate concurrent illnesses. All patients with central AI or RAS were given supplies to enable their self-injection of parenteral GC (19).

In accordance with STARD (standards for reporting diagnostic accuracy studies) criteria, we considered the final diagnosis (based on the previously mentioned criteria) as the reference standard. Patients were then grouped as cases of central AI (on chronic GC treatment, $n=37$), RAS (given stress doses of GC according to need, $n=31$), or normal HPA axis (without any GC replacement therapy, $n=106$).

This observational study was conducted in accordance with the STrengthening the Reporting of OBServational studies in Epidemiology (STROBE) guidelines (20).

The concentration of $1\ \mu\text{g}/\text{mL}$ of synacthen was obtained by diluting a vial of ACTH (available in a $250\ \mu\text{g}/\text{mL}$ ready-to-use formulation, Synacthen®) in $249\ \text{mL}$ of sterile saline physiological solution (NaCl 0.9%), then collecting $1\ \text{mL}$ of the solution (1), injected directly in the catheter hub, to avoid tubing effect (21). Serum and salivary samples were collected simultaneously, at the baseline and 30 min after the injection of $1\ \mu\text{g}$ of ACTH. LDDST was performed in the early morning (before 09:30 h) in all patients.

Serum F was measured by chemiluminescence immunoassay (Immulite 2000, Siemens Healthcare; limit of detection $6\ \text{nmol}/\text{L}$). Saliva was collected in a cotton-based sampling device with or without citric acid (Salivette® green or blank cap commercial device, Sarstedt, Numbrecht, Germany). Patients were advised to soak the absorbent cotton for 2 or 3 min, then the saliva sample was placed in a plastic tube and kept at $+4^{\circ}\text{C}$. Samples were collected at least 30 min before eating or drinking, to avoid any source of food contamination. Patients brushed their teeth at least 30 min before collection. Smoking or eating licorice was forbidden (22). Salivary F and E levels were measured with a LC-MS/MS method, as detailed elsewhere (23).

Written consent was obtained from all participants after fully explaining the purpose of the study and the nature of all the procedures used. The study complied with the principles of the Declaration of Helsinki and was approved by the Ethics Committee at Padova University Hospital (protocol No. 0070140-2020). The clinical data were collected from the Padova University Hospital web-based database.

Statistical analyses

Proportions and rates were calculated for categorical data. Continuous data are reported as means and s.e. Groups were compared with the chi-square test for categorical variables (the raw P values were adjusted with the Bonferroni method to take multiple comparisons into account) and with Student's t -test for quantitative variables.

To measure endocrine serum or salivary response after administering synacthen, we recorded the values 30 min after injection (termed $F^{30\text{LDSST}}$ or $E^{30\text{LDSST}}$, for F and E) and calculated the difference between the stimulated and basal levels (Δ), or their percentage increase ($\Delta\%$). We ran receiver operating curve (ROC) analyses to ascertain the sensitivity (SE), specificity (SP), and their 95% CI. We calculated the likelihood ratio (LR) of the test results as this is independent of disease prevalence: a positive LR (LR^{pos}) and a negative LR (LR^{neg}), respectively, indicate that by how much the probability of HPA axis-related disease increases or decreases if the test result is positive or negative, with the 95% CI calculated using the method proposed by Simel *et al.* (24).

The SPSS 24 software package for Windows (SPSS, Inc.) was used to manage the database and perform the statistical analysis. The significance level was set at $P < 0.05$ for all tests. All data analyzed during this study are included in

the data repositories of the University of Padova – Research Data UniPD (25).

Results

Diagnostic accuracy of serum and salivary F and E

Based on their clinical presentation, we tested patients with a sellar mass ($n = 40$); after brain or pituitary surgery ($n = 59$); after RT ($n = 17$); with symptoms of AI ($n = 40$); after GC withdrawal ($n = 18$). As shown in Table 1, basal F levels were similar among AS patients, and only peak F after synacthen administration was lower in patients with RAS.

Salivary F and E correlated with serum F in all patients, at baseline and after LDSST, as reassumed in Table 2 and depicted in Fig. 1. Poor correlation between serum F and salivary F or E is reported in AS patients (available in data repository (25)).

As shown in Table 3, central AI was diagnosed mainly in patients with tertiary or hypothalamic AI after CS or exogenous GC treatment. In up to 90% of cases, patients with a sellar lesion (with no history of surgery or RT) did not have central AI ($P < 0.001$ with the Bonferroni-adjusted comparison).

Table 4 shows the diagnostic accuracy of serum or salivary F in 37 patients with central AI and 137 AS cases. The diagnostic accuracy of basal unstimulated salivary F or E was similar to that of serum F; the SP of salivary E^0 to indicate central AI was higher than that of salivary F^0 (83% vs 68%, $P < 0.05$). An unstimulated salivary $F^0 > 9.8$ nmol/L presented SP 100% to predict an intact HPA axis. Considering serum $F^{30LDSST} > 500$ nmol/L or salivary $F^{30LDSST} > 12.1$ nmol/L as sufficient thresholds to exclude AI, a basal salivary $F^0 > 9.8$ nmol/L can be used to predict the normal response to LDSST (100% SP).

The salivary F peak in the LDSST (salivary $F^{30LDSST} > 12.1$ nmol/L) showed a good SE and SP (95 and 84%, respectively) in diagnosing central AI, with $LR^{neg} < 0.1$, which indicates a conclusive decrease in the likelihood of disease. Salivary $F^{30LDSST}$ achieved a SE of 100% in diagnosing AI if < 7.2 nmol/L (SP 51%), and a SP of 100% in ruling out AI if > 23.6 nmol/L (SE 59%). The differences between the basal and peak salivary F values (Δ) were accurate in excluding AI if $\Delta^{Salivary_F} > 7.5$ nmol/L (SE 93%, SP 78%).

LDSST in patients with clearly insufficient or sufficient basal serum F^0

The LDSST was performed in 16 patients with basal serum $F^0 \leq 83$ nmol/L. Their mean serum $F^{30LDSST}$ was 136 nmol/L, and their mean salivary F^0 and salivary E^0 were 1 and 4.6 nmol/L, respectively, which rose to a mean salivary $F^{30LDSST}$ and salivary $E^{30LDSST}$ of 2.6 and 8.3 nmol/L, respectively. All patients with low basal serum $F^0 (\leq 83$ nmol/L) reached an inadequate salivary peak (salivary $F^{30LDSST}$ in the range of 0.5–4.7 nmol/L).

The LDSST was also performed in eight patients with basal serum $F^0 \geq 415$ nmol/L. Their mean serum $F^{30LDSST}$ was 612 nmol/L, and their mean salivary F^0 and salivary E^0 were 11.9 and 39.8 nmol/L, respectively, which rose to a mean salivary $F^{30LDSST}$ and salivary $E^{30LDSST}$ of 27.9 and 56.9 nmol/L, respectively. All these patients reached a peak salivary $F^{30LDSST} > 12.1$ nmol/L.

Excluding those 24 patients with clearly insufficient or sufficient basal serum F^0 to, respectively, diagnose AI or define AS, we consider a group of 150 subjects with borderline basal serum F levels (83–415 nmol/L) and the indication to perform a LDSST: 21 with AI and 129 AS patients (31 cases with RAS). Overall, the diagnostic accuracy of serum and salivary F was similar to that obtained in the whole cohort of patients (reported in Table 5).

Table 1 Basal and post-synacthen serum or salivary cortisol (F) and cortisone (E) levels. Data are expressed as means and s.e.

	Adrenal sufficient patients, $n = 137$			Central AI, $n = 37$
	Normal HPA axis, $n = 106$	RAS, $n = 31$		
Serum F^0 (nmol/L)	287.1 (9.7)	265.4 (12.2)		114.7 (15.1) ^{a,b}
Serum $F^{30LDSST}$ (nmol/L)	512.4 (9.6)	426.7 (13.4) ^a		257.6 (24.3) ^{a,b}
Salivary F^0 (nmol/L)	5.82 (0.33)	5.87 (0.48)		2.26 (0.33) ^{a,b}
Salivary $F^{30LDSST}$ (nmol/L)	31.16 (1.5)	21.45 (1.45) ^a		7.54 (1) ^{a,b}
Salivary E^0 (nmol/L)	28.7 (1.26)	25.14 (1.62)		12.85 (2.52) ^{a,b}
Salivary $E^{30LDSST}$ (nmol/L)	63.63 (2.21)	49.6 (3.2) ^a		23.74 (2.92) ^{a,b}
Salivary F/E^0	0.21 (0.01)	0.24 (0.02)		0.32 (0.07)
Salivary F/E^{30}	0.5 (0.02)	0.46 (0.04)		0.44 (0.13)

^a $P < 0.001$ vs no treatment; ^b $P < 0.001$ vs stress dose.

AI, adrenal insufficiency; HPA, hypothalamic–pituitary–adrenal; RAS, relative adrenal sufficiency.

Table 2 Correlation coefficients between salivary cortisol (F) and cortisone (E) and serum F.

	All patients, n = 174	Adrenal sufficient patients, n = 137		Central AI, n = 37
		Normal HPA axis, n = 106	RAS, n = 31	
Salivary F and serum F				
Baseline (before LDSST)	0.75 (0.68–0.81) ^a	0.55 (0.40–0.67) ^a	0.75 (0.55–0.87) ^a	0.73 (0.53–0.85) ^a
30 min after 1 µg ACTH	0.50 (0.38–0.60) ^a	ns	ns	0.62 (0.38–0.79) ^a
Pooled results (before and after LDSST)	0.68 (0.62–0.73) ^a	0.66 (0.57–0.73) ^a	0.71 (0.56–0.82) ^a	0.72 (0.58–0.81) ^a
Salivary E and serum F				
Baseline (before LDSST)	0.67 (0.58–0.75) ^a	0.55 (0.40–0.67) ^a	0.54 (0.23–0.73) ^b	0.65 (0.40–0.81) ^a
30 min after 1 µg ACTH	0.64 (0.54–0.72) ^a	ns	ns	0.74 (0.54–0.86) ^a
Pooled results (before and after LDSST)	0.76 (0.71–0.80) ^a	0.70 (0.63–0.76) ^a	0.67 (0.50–0.72) ^a	0.73 (0.60–0.83) ^a

^aP < 0.001; ^bP < 0.01

AI, adrenal insufficiency; HPA, hypothalamic–pituitary–adrenal; RAS, relative adrenal sufficiency.

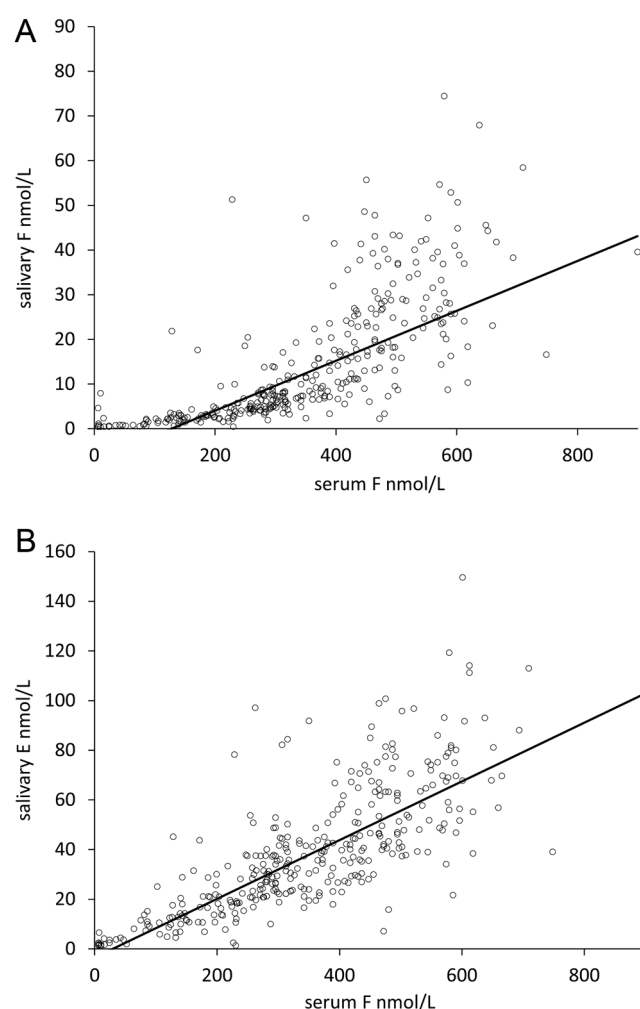


Figure 1 Passing-Bablok regression between serum cortisol (F) and salivary F (panel A: $y = 0.06x - 7.11$, $r = 0.68$, $P < 0.001$) or salivary cortisone (E, panel B: $y = 0.12x - 3.48$, $r = 0.76$, $P < 0.001$) in all patients, for both basal and post-synacthen samples ($n = 348$).

Threshold for serum and salivary F

As shown in Fig. 2, in 49 out of 51 patients with basal serum F⁰ 83–415 nmol/L the LDSST was sufficient to rule out central AI, while RAS was suspected in two patients (evaluated 4 months after pituitary surgery, with serum F^{30LDSST} 516 and 502 nmol/L, respectively). As depicted in Fig. 3, patients with AI had lower F levels and a lower F peak in the LDSST. Serum or salivary peak F levels in the patients with RAS were intermediate between patients with central AI and those with a normal HPA axis. Considering the threshold calculated for salivary F^{30LDSST} (12.1 nmol/L), the patients with RAS would be classified AS, both in the cohort as a whole and in the subset with basal borderline serum F levels (83–415 nmol/L, Fig. 3, panel B).

The cut-off for the salivary F peak in the LDSST for assuming a normal HPA axis was set as the 2.5th percentile of patients with peak serum F > 500 nmol/L levels during the test. Salivary F^{30LDSST} ≥ 12.1 nmol/L was considered sufficient to exclude AI (this threshold is the same as the ROC-calculated cut-off described in Tables 4 and 5). This threshold could be used to prevent unnecessary treatments in most cases of suspected central AI. Overall, 99 out of 150 patients (66%) had an inadequate F^{30LDSST} (≤500 nmol/L), but in 79 of them (80%), the response to synacthen in terms of salivary F^{30LDSST} was >12.1 nmol/L, meaning that these patients could be considered AS (Fig. 2). The HPA axis of these 79 patients was considered normal in 48 cases, RAS in 25, and consistent with central AI in six. An inadequate serum and salivary response in the LDSST were seen in 20 out of 99 patients (Fig. 2). Most of them (15/20, 75%) were considered cases of central AI. A RAS condition was defined in three patients after pituitary surgery (one combined with radiotherapy).

From a clinical perspective, none of the patients in the RAS group experienced an adrenal crisis during the follow-up after the baseline visit (mean 26 months,

Table 3 Type of treatment and biochemical assessment in patients with suspected central adrenal insufficiency (AI), grouped by their clinical presentation. Data are expressed as means and s.e.

Clinical presentation	Adrenal sufficient patients, n = 137		Central AI, n = 37		Salivary F ^{30LDSST}	Salivary F ⁰	Salivary E ^{30LDSST}
	Normal HPA axis, n = 106	RAS, n = 31	Serum F ⁰	Serum F ^{30LDSST}			
Sellar lesion (n = 40)	30 (75%)	5 (12.5%)	245.9 (16)	469.2 (15.2)	29.8 (3.2)	4.8 (0.4)	58.2 (3.9)
Surgery (n = 59)	30 (50.8%)	17 (28.8%)	255.5 (14)	415.8 (19.4) ^a	23.2 (1.7)	5.4 (0.4)	54.3 (3.6)
Surgery + RT (n = 17)	12 (70.6%)	1 (5.9%)	258.6 (28.1)	455.5 (37.6)	22.9 (3.7)	5.1 (0.9)	46.2 (4.8)
Symptoms of AI (n = 40)	30 (75%)	1 (2.5%)	256 (21)	499.5 (21.4) ^c	26 (2.4)	5.4 (0.6)	57.4 (4.1)
GC excess (n = 18)	4 (22.2%)	7 (38.9%)	186.2 (30.1) ^d	336.2 (44.8) ^{b,ef}	14.4 (2.4) ^{b,df}	3.8 (0.7) ^{g,d}	34.3 (5.5) ^{b,df}

^ap < 0.05 vs sellar lesion; ^bp < 0.01 vs sellar lesion; ^cp < 0.01 vs surgery; ^dp < 0.05 vs surgery; ^ep < 0.05 vs surgery+RT; ^fp < 0.05 vs symptoms of AI. E, cortisone; F, cortisol; GC, glucocorticoid; HPA, hypothalamic–pituitary–adrenal; LDSST, low-dose short synacthen test; RAS, relative adrenal sufficiency; RT, radiotherapy.

range 12–54 months, at least 12 months per selection criteria), five needed a GC dose for stress (two during endoscopic procedures, three during infections), and none shifted from the RAS to the central AI group.

We calculated the 2.5th and 97.5th percentiles of the basal salivary F for the 59 patients with a normal HPA axis and basal serum F levels > 415 nmol/L, or serum F^{30LDSST} > 500 nmol/L. The lower threshold was 2 nmol/L (SE 91.2% in detecting AI), and the upper threshold was 16.4 nmol/L (SP 100% in excluding AI).

Discussion

Diagnosing central AI correctly is of the utmost importance because cortisol-related comorbidities can develop if a patient with a normal HPA axis is given unwarranted GC treatment.

We grouped our cohort of patients according to their clinical presentation, endocrine evaluation of their HPA axis, and the type of treatment recommended. Patients with central AI are easy to identify according to the Endocrine Society guidelines (2). On the other hand, the guidelines define those with a normal HPA axis as patients whose basal or stimulated F levels exceed given thresholds, which are high and not always achieved in clinical practice. Other authors have already suggested an intermediate phenotype – what we have called RAS – when assessing a patient’s HPA axis (17). From an endocrine perspective, hormone secretion occurs on a continuum: subclinical or intermediate scenarios are frequently encountered in clinical practice, like subclinical hypercortisolism, hypo- or hyperthyroidism, and so on (26, 27, 28).

Two meta-analyses (4, 7) previously showed and the Endocrine Society guidelines (2) have reiterated that corticotropin test (either standard- or low-dose) is not always enough to judge HPA axis integrity. In both meta-analyses, most studies described the use of standard-dose (250 µg) synacthen test. In our clinical practice, in patients with suspected central adrenal insufficiency, we currently use low dose (1 µg) synacthen test (1). It is suggested that 250 µg is an excessive stimulus, eliciting very high circulating ACTH levels (29); these high ACTH levels are excessive to detect a mild central AI in patients with partial HPA axis impairment (adrenal responsiveness to high doses of ACTH may be preserved, with a resultant false-negative SST). From a clinical perspective, the LDSST suffers from a low SE: an insufficient serum F response to ACTH cannot confirm AI. This means that a complete and accurate

Table 4 Diagnostic accuracy of basal and post-synacthen cortisol (F) in 37 patients with central adrenal insufficiency (AI) and 137 adrenal sufficient patients. The threshold for AI was based on the ROC curve and Youden's J index.

	Threshold level	SE, % (95% CI)	SP, % (95% CI)	LR ^{neg} , % (95% CI)	LR ^{pos} , % (95% CI)	AUC, % (95% CI)
Serum F ⁰	181 nmol/L	86.1 (79.4–90.9)	75.7 (59.8–86.6)	0.18 (0.116–0.289)	3.541 (1.998–6.276)	0.9 (0.847–0.954)
Serum F ^{30LDSST}	427 nmol/L	78.8 (71.3–84.8)	89.2 (75.3–95.7)	0.237 (0.169–0.334)	7.292 (2.878–18.473)	0.921 (0.876–0.967)
Δ _{SerumF}	145 nmol/L	73.7 (65.8–80.4)	56.8 (40.9–71.3)	0.463 (0.311–0.689)	1.705 (1.163–2.499)	0.691 (0.598–0.784)
Δ% _{SerumF}	114%	64.9 (48.8–78.2)	78.1 (70.5–84.2)	0.45 (0.288–0.703)	2.962 (1.995–4.398)	0.704 (0.598–0.809)
Salivary F ⁰	2.7 nmol/L	83.2 (76.1–88.6)	67.6 (51.5–80.4)	0.248 (0.161–0.384)	2.566 (1.602–4.11)	0.837 (0.764–0.91)
Salivary F ^{30LDSST}	12.1 nmol/L	94.9 (89.8–97.5)	83.8 (68.9–92.3)	0.061 (0.029–0.127)	5.852 (2.81–12.184)	0.952 (0.917–0.987)
Δ _{Salivary_F}	7.5 nmol/L	93.4 (87.9–96.5)	78.4 (62.8–88.6)	0.084 (0.044–0.161)	4.321 (2.336–7.993)	0.931 (0.887–0.974)
Δ% _{SalivaryF}	278%	65.7 (57.4–73.1)	73 (57–84.6)	0.47 (0.347–0.637)	2.431 (1.412–4.184)	0.735 (0.641–0.830)
Salivary E ⁰	18.4 nmol/L	78.8 (71.2–84.8)	82.9 (67.3–91.9)	0.255 (0.179–0.365)	4.599 (2.208–9.576)	0.858 (0.78–0.935)
Salivary E ^{30LDSST}	37.3 nmol/L	87.6 (81–92.1)	80 (64.1–90)	0.155 (0.096–0.249)	4.38 (2.251–8.521)	0.91 (0.857–0.963)
Δ _{Salivary_E}	17.9 nmol/L	80.3 (72.8–86.1)	77.1 (61–87.9)	0.255 (0.174–0.375)	3.513 (1.901–6.493)	0.854 (0.785–0.924)
Δ% _{SalivaryE}	207%	20.2 (15.2–28.8)	97.1 (85.5–99.5)	0.812 (0.732–0.9)	7.409 (1.405–52.524)	0.615 (0.511–0.719)
Salivary F/E ⁰	0.08	99.3 (95.9–99.9)	8.6 (3–22.4)	0.085 (0.009–0.794)	1.086 (0.98–1.203)	0.482 (0.363–0.601)
Salivary F/E ³⁰	0.33	79.6 (72.1–85.5)	65.7 (49.1–79.2)	0.311 (0.207–0.468)	2.321 (1.456–3.7)	0.774 (0.686–0.861)

Δ, difference between peak and basal levels; Δ%, increase from basal to peak levels; AUC, area under the curve; E, cortisone; F/E, cortisol-to-cortisone ratio; LDSST, low-dose short synacthen test; LR^{neg}, negative likelihood ratio; LR^{pos}, positive likelihood ratio; SE, sensitivity; SP, specificity.

clinical assessment is needed. A patient with low-to-normal F levels after surgery and RT for a pituitary macroadenoma is more likely to have central AI, for instance, so further dynamic tests are required in most patients. These may include the ITT (often considered the gold standard, but requiring medical supervision (30)), the metyrapone test (not recommended by the guidelines, and requiring an appropriate assay for measuring 11-deoxycortisol levels (30)), or the corticotropin-releasing hormone test (the diagnostic accuracy of which is debated (31)). On the other hand, the LDSST is easy to perform, convenient, and safe.

We examined the diagnostic accuracy of salivary F in 174 consecutive patients with suspected central AI. In our unselected cohort, basal unstimulated serum F levels sufficed to confirm or rule out AI without any further dynamic tests only in 24 cases (≤ 83 or ≥ 415 nmol/L in 16 and 8 patients, respectively). In all these 24 patients,

peak serum or salivary F levels in the LDSST did not improve on the diagnostic value of the basal serum F levels. Dynamic tests could, therefore, only be avoided for a minority of our original cohort of patients (24 out of 174, 14%). We also calculated the accuracy of serum or salivary F in the LDSST in patients with basal serum F levels in the range of 83–415 nmol/L given that, in clinical practice, basal serum F can predict HPA axis function without the need for any further dynamic tests in patients with clearly low (≤ 83 nmol/L) or adequate (≥ 415 nmol/L) basal serum F levels.

In our study, the best ROC-calculated cut-off for basal F levels low enough to pinpoint patients with central AI was 181 nmol/L, that is, higher than the threshold proposed by the Endocrine Society (2), albeit at the expense of SP. On the other hand, the higher peak serum F cut-off that we adopted to rule out central AI (500 nmol/L) confirmed a high SP (100%) (32).

Table 5 Diagnostic accuracy of post-synacthen cortisol (F) in patients with indication to perform LDSST (basal morning serum F 83–415 nmol/L). We selected 150 patients: 21 with central AI and 129 adrenal sufficient subjects). The threshold for AI was based on the ROC curve and Youden's J index.

	Threshold level	SE, % (95% CI)	SP, % (95% CI)	LR ^{neg} , % (95% CI)	LR ^{pos} , % (95% CI)	AUC, % (95% CI)
Serum F ^{30LDSST} (nmol/L)	437	70.5 (62.2–77.7)	95.2 (77.3–92.2)	0.309 (0.233–0.411)	14.814 (2.181–100.64)	0.879 (0.81–0.947)
Salivary F ^{30LDSST} (nmol/L)	12.1	94.6 (89.2–97.4)	71.4 (50.1–86.2)	0.076 (0.035–0.164)	3.31 (1.681–6.517)	0.914 (0.855–0.972)
Salivary E ^{30LDSST} (nmol/L)	42.6	74.4 (66.3–81.2)	80 (58.4–91.9)	0.320 (0.222–0.462)	3.721 (1.54–8.002)	0.842 (0.761–0.924)
Salivary F/E ³⁰	0.47	46.5 (38.1–55.1)	95 (76.4–99.1)	0.563 (0.466–0.681)	9.302 (1.365–63.405)	0.747 (0.649–0.844)

Δ, difference between peak and basal levels; Δ%, increase from basal to peak levels; AUC, area under the curve; E, cortisone; F/E, cortisol-to-cortisone ratio; LDSST, low-dose short synacthen test; LR^{neg}, negative likelihood ratio; LR^{pos}, positive likelihood ratio; SE, sensitivity; SP, specificity.

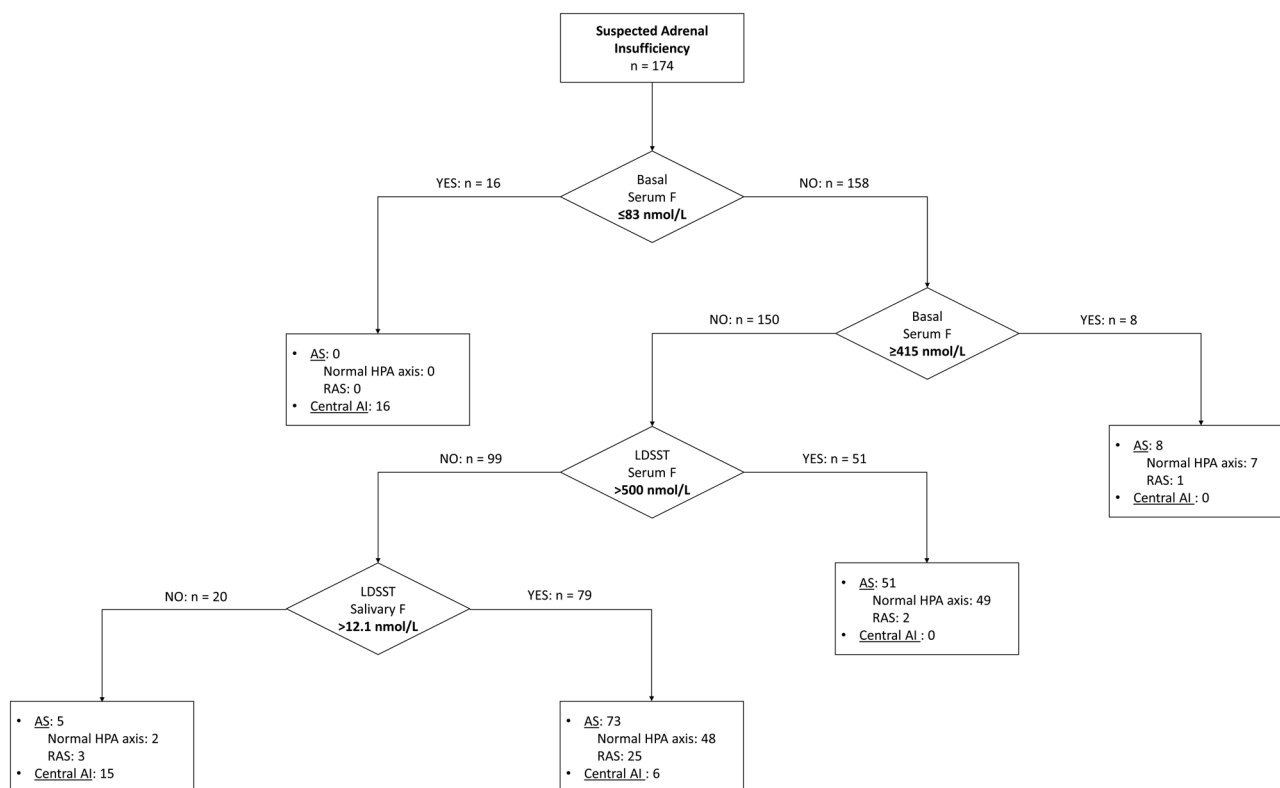


Figure 2

Distribution of patients using the thresholds proposed by the Endocrine Society, combined with our calculated threshold for salivary cortisol. F, cortisol; AS, adrenal sufficient patients; AI, adrenal insufficiency; HPA, hypothalamic–pituitary–adrenal; RAS, relative adrenal sufficiency.

In our cohort as a whole, peak serum F levels alone (>500 nmol/L) were able to rule out central AI in 51 out of 150 cases. According to the Endocrine Society guidelines, GC treatment would be at least considered for the remaining 99 patients, and further basal or dynamic tests would be called for. In patients with HPA axis disease, testing saliva is patient-friendly and suitable for use with outpatients, as well as offering several analytical advantages, because salivary F reflects the amount of serum-free F (11). In our cohort, morning unstimulated salivary F levels were not superior to serum F for diagnostic purposes, as also reported in another study (33). A basal unstimulated salivary F⁰ > 9.8 nmol/L presented SP 100% to predict both an intact HPA axis and a normal response to LDSST (serum F^{30LDSST} > 500 nmol/L or salivary F^{30LDSST} > 12.1 nmol/L), limiting the need for unnecessary LDSST in 69% of patients (11 out of 16 patients with basal serum F⁰ 83–415 nmol/L). On the other hand, measuring peak salivary F (salivary F^{30LDSST} > 12.1 nmol/L) made a reliable contribution to diagnostic accuracy in clinical practice. Most of our patients with an inadequate serum F response to synacthen (<500 nmol/L) were considered AS and were not given chronic GC treatment. The salivary

F threshold of 12.1 nmol/L was able to select 25 patients with RAS (out of 79, 32%) and 48 with a normal HPA axis (out of 79, 61%). Only a minority of our patients had a clinical history, as well as signs or symptoms, consistent with central AI and were started on chronic GC treatment: a careful clinical assessment is hugely important, even after performing adequate dynamic tests (7). The calculated threshold for peak salivary F was the same although it was calculated using two different methods: the ROC-calculated cut-off and the lower percentile of normality (2.5th) of patients with serum F > 500 nmol/L in the LDSST, as recently reported (13). In our cohort, we found a poor correlation between serum F and salivary F or E in AS patients: we performed a short test (2 measurements, baseline and after 30 min), and probably the peak and decline of serum F are different from that of salivary F.

If we had decided to treat patients on the strength of an insufficient serum response to synacthen in the LDSST (<500 nmol/L), we would have had to start GC replacement therapy in 99 patients (57% of the initial cohort, 66% of patients for whom the test was indicated with basal serum F levels in the range of 83–415 nmol/L).

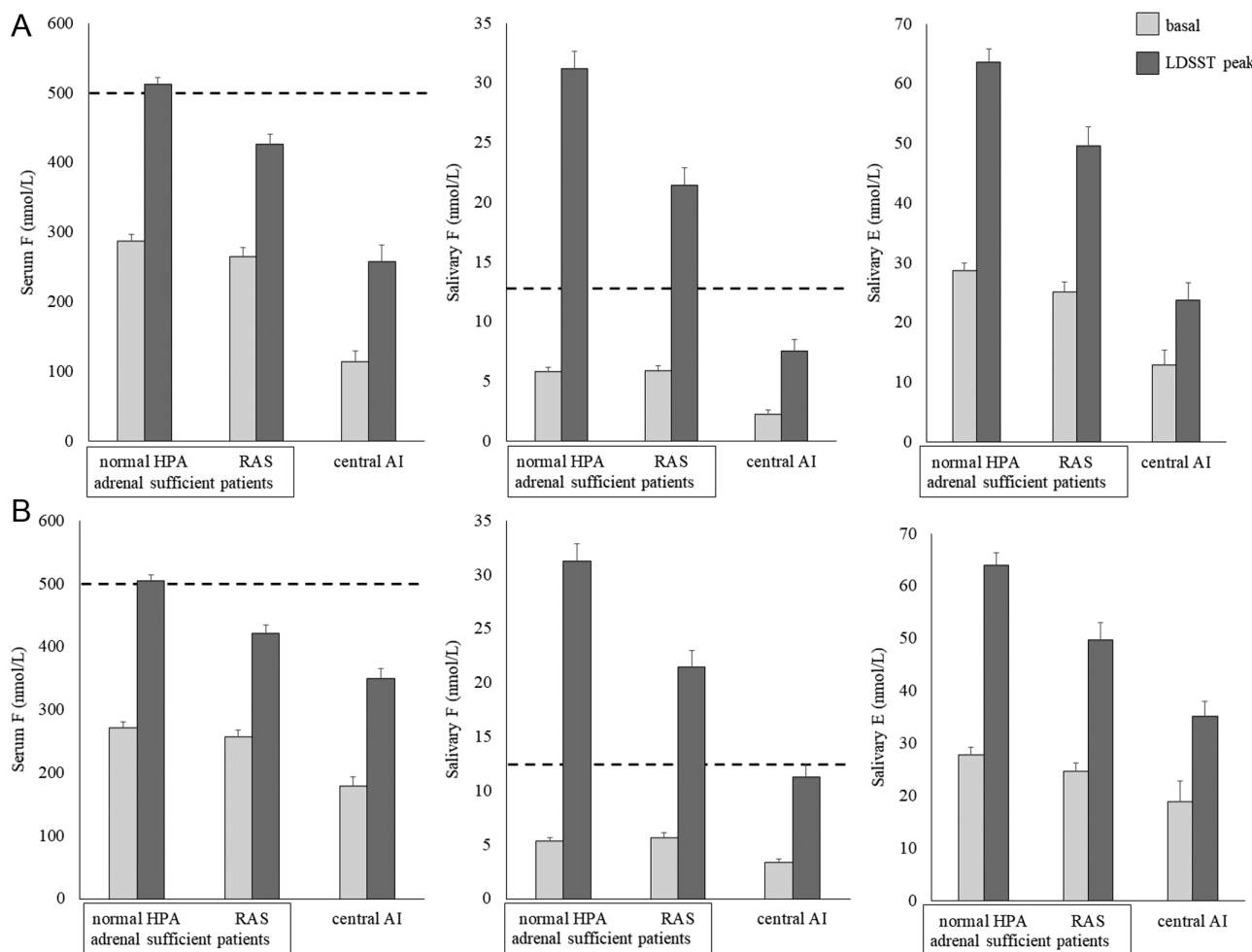


Figure 3

Basal and post-synacthen levels of serum or salivary cortisol (F), and salivary cortisone (E) in the whole cohort of patients ($n = 174$, panel A) and in the group with basal serum F in the range of 83–415 nmol/L ($n = 150$, panel B). Dashed lines indicate the serum F threshold to exclude AI (500 nmol/L) and our calculated threshold for salivary F (12.1 nmol/L).

By applying our calculated salivary F threshold, on the other hand, only 20 patients were AI. This would mean a reduction in the number of further dynamic tests performed and GC treatments administered in up to 80% of cases of suspected AI. Basal unstimulated serum F represents the first screening test to detect primary or secondary AI, according to guidelines and clinical practice (2, 18, 19). Regarding dynamic tests, especially LDSST, serum F might overestimate the number of patients requiring a substitutive treatment: salivary F could be considered to avoid unnecessary long-term GC therapy or to suggest a stress dose. Further studies, ideally prospective, should be considered to establish the diagnostic accuracy of salivary F.

Our work has some strengths (the number of cases, the measurement of salivary F with LC-MS/MS), and some limitations, first of all, the reduced number of patients with

central AI or RAS. Moreover, we did not consider a control group of healthy subjects, because our aim was to provide a threshold that is able to detect central AI in a cohort of patients with suspected HPA axis insufficiency. In this type of study, it is critical to correctly categorize patients. That is usually done with a gold standard (e.g. the ITT) or with a confirmatory test. We preferred to take a clinical approach, based on a close follow-up (no adrenal crises were reported in our patients with RAS during their follow-up). Another limitation lies in that we developed our salivary F threshold in a cross-sectional observational study (we are planning a prospective trial to validate it). We also considered a novel category of patients with partial central AI (what we called RAS). We believe that a 'subclinical' or 'sub-optimal' HPA axis function category should be considered. There is a gray area between normal function and insufficiency, where the 'sick-day' rules could be applied in order to

avoid patients being treated unnecessarily. An additional drawback is related to laboratory analyses, based on clinical practice: we compared serum F measured by a CE-IVD immunoassay with salivary F and E determined by homebrew LC-MS/MS.

To conclude, we have developed a novel cut-off for salivary F in the LDSST that can increase the test's accuracy and help us to treat our patients appropriately.

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 study did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.

Research involving human participants and patient consent

Informed consent was obtained from all participants.

Data availability statement

All data generated or analyzed during this study are included in this published article or in the data repositories listed in the references.

Author contribution statement

All authors contributed equally to the study design, the data acquisition, analysis, and interpretation, and drafting the manuscript. They all approved the final version of the paper.

References

- Ceccato F & Scaroni C. Central adrenal insufficiency: open issues regarding diagnosis and glucocorticoid treatment. *Clinical Chemistry and Laboratory Medicine* 2019 **57** 1125–1135. (<https://doi.org/10.1515/cclm-2018-0824>)
- Fleseriu M, Hashim IA, Karavitaki N, Melmed S, Murad MH, Salvatori R & Samuels MH. Hormonal replacement in hypopituitarism in adults: an Endocrine Society clinical practice guideline. *Journal of Clinical Endocrinology and Metabolism* 2016 **101** 3888–3921. (<https://doi.org/10.1210/jc.2016-2118>)
- Sherlock M, Reulen RC, Alonso AA, Ayuk J, Clayton RN, Sheppard MC, Hawkins MM, Bates AS & Stewart PM. ACTH deficiency, higher doses of hydrocortisone replacement, and radiotherapy are independent predictors of mortality in patients with acromegaly. *Journal of Clinical Endocrinology and Metabolism* 2009 **94** 4216–4223. (<https://doi.org/10.1210/jc.2009-1097>)
- Kazlauskaitė R, Evans AT, Villabona CV, Abdu TA, Ambrosi B, Atkinson AB, Choi CH, Clayton RN, Courtney CH, Gonc EN, *et al.* Corticotropin tests for hypothalamic-pituitary-adrenal insufficiency: a metaanalysis. *Journal of Clinical Endocrinology and Metabolism* 2008 **93** 4245–4253. (<https://doi.org/10.1210/jc.2008-0710>)
- Crowley RK, Argese N, Tomlinson JW & Stewart PM. Central hypoadrenalism. *Journal of Clinical Endocrinology and Metabolism* 2014 **99** 4027–4036. (<https://doi.org/10.1210/jc.2014-2476>)
- Grossman AB. Clinical Review: The diagnosis and management of central hypoadrenalism. *Journal of Clinical Endocrinology and Metabolism* 2010 **95** 4855–4863. (<https://doi.org/10.1210/jc.2010-0982>)
- Ospina NS, Nofal AAI, Bancos I, Javed A, Benkhadra K, Kapoor E, Lteif AN, Natt N & Murad MH. ACTH stimulation tests for the diagnosis of adrenal insufficiency: systematic review and meta-analysis. *Journal of Clinical Endocrinology and Metabolism* 2016 **101** 427–434. (<https://doi.org/10.1210/jc.2015-1700>)
- Grassi G, Morelli V, Ceriotti F, Polledri E, Fustinoni S, D'Agostino S, Mantovani G, Chiodini I & Arosio M. Minding the gap between cortisol levels measured with second-generation assays and current diagnostic thresholds for the diagnosis of adrenal insufficiency: a single-center experience. *Hormones* 2020 **19** 425–431. (<https://doi.org/10.1007/s42000-020-00185-y>)
- El-Farhan N, Pickett A, Ducroq D, Bailey C, Mitchem K, Morgan N, Armston A, Jones L, Evans C & Rees DA. Method-specific serum cortisol responses to the adrenocorticotrophin test: comparison of gas chromatography-mass spectrometry and five automated immunoassays. *Clinical Endocrinology* 2013 **78** 673–680. (<https://doi.org/10.1111/cen.12039>)
- El-Farhan N, Rees DA & Evans C. Measuring cortisol in serum, urine and saliva - are our assays good enough? *Annals of Clinical Biochemistry* 2017 **54** 308–322. (<https://doi.org/10.1177/0004563216687335>)
- Raff H. Utility of salivary cortisol measurements in Cushing's syndrome and adrenal insufficiency. *Journal of Clinical Endocrinology and Metabolism* 2009 **94** 3647–3655. (<https://doi.org/10.1210/jc.2009-1166>)
- Marcus-Perlman Y, Tordjman K, Greenman Y, Limor R, Shenkerman G, Osher E & Stern N. Low-dose ACTH (1 µg) salivary test: a potential alternative to the classical blood test. *Clinical Endocrinology* 2006 **64** 215–218. (<https://doi.org/10.1111/j.1365-2265.2006.02451.x>)
- Langelaan MLP, Kisters JMH, Oosterwerff MM & Boer AK. Salivary cortisol in the diagnosis of adrenal insufficiency: cost efficient and patient friendly. *Endocrine Connections* 2018 **7** 560–566. (<https://doi.org/10.1530/EC-18-0085>)
- Manetti L, Rossi G, Grasso L, Raffaelli V, Scattina I, Del Sarto S, Cosottini M, Iannelli A, Gasperi M, Bogazzi F, *et al.* Usefulness of salivary cortisol in the diagnosis of hypercortisolism: comparison with serum and urinary cortisol. *European Journal of Endocrinology* 2013 **168** 315–321. (<https://doi.org/10.1530/EJE-12-0685>)
- Raff H. Measurement of salivary cortisone to assess the adequacy of hydrocortisone replacement. *Journal of Clinical Endocrinology and Metabolism* 2016 **101** 1350–1352. (<https://doi.org/10.1210/jc.2016-1228>)
- Debono M, Harrison RF, Whitaker MJ, Eckland D, Arlt W, Keevil BG & Ross RJ. Salivary cortisone reflects cortisol exposure under physiological conditions and after hydrocortisone. *Journal of Clinical Endocrinology and Metabolism* 2016 **101** 1469–1477. (<https://doi.org/10.1210/jc.2015-3694>)
- Vaiani E, Lazzati JM, Ramirez P, Costanzo M, Gil S, Dratler G, Zaidman V, Chaler E & Belgorosky A. The low-dose ACTH test: usefulness of combined analysis of serum and salivary maximum cortisol response in pediatrics. *Journal of Clinical Endocrinology and Metabolism* 2019 **104** 4323–4330. (<https://doi.org/10.1210/jc.2019-00304>)
- Bornstein SR, Allolio B, Arlt W, Barthel A, Don-Wauchope A, Hammer GD, Husebye ES, Merke DP, Murad MH, Stratakis CA, *et al.* Diagnosis and treatment of primary adrenal insufficiency: an Endocrine Society clinical practice guideline. *Journal of Clinical Endocrinology and Metabolism* 2016 **101** 364–389. (<https://doi.org/10.1210/jc.2015-1710>)
- Husebye ES, Allolio B, Arlt W, Badenhop K, Bensing S, Betterle C, Falorni A, Gan EH, Hulting AL, Kasperlik-Zaluska A, *et al.* Consensus statement on the diagnosis, treatment and follow-up of patients with primary adrenal insufficiency. *Journal of Internal Medicine* 2014 **275** 104–115. (<https://doi.org/10.1111/joim.12162>)

- 20 von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP & STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Journal of Clinical Epidemiology* 2008 **61** 344–349. (<https://doi.org/10.1016/j.jclinepi.2007.11.008>)
- 21 Wade M, Baid S, Calis K, Raff H, Sinaï N & Nieman L. Technical details influence the diagnostic accuracy of the 1 microg ACTH stimulation test. *European Journal of Endocrinology* 2010 **162** 109–113. (<https://doi.org/10.1530/EJE-09-0746>)
- 22 Ceccato F, Albiger N, Reimondo G, Frigo AC, Ferasin S, Occhi G, Mantero F, Terzolo M & Scaroni C. Assessment of glucocorticoid therapy with salivary cortisol in secondary adrenal insufficiency. *European Journal of Endocrinology* 2012 **167** 769–776. (<https://doi.org/10.1530/EJE-12-0534>)
- 23 Antonelli G, Ceccato F, Artusi C, Marinova M & Plebani M. Salivary cortisol and cortisone by LC–MS/MS: validation, reference intervals and diagnostic accuracy in Cushing’s syndrome. *Clinica Chimica Acta: International Journal of Clinical Chemistry* 2015 **451** 247–251. (<https://doi.org/10.1016/j.cca.2015.10.004>)
- 24 Simel DL, Samsa GP & Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. *Journal of Clinical Epidemiology* 1991 **44** 763–770. ([https://doi.org/10.1016/0895-4356\(91\)90128-V](https://doi.org/10.1016/0895-4356(91)90128-V))
- 25 Ceccato F. Serum + salivary LDSST (repository data @UNIPD), 2020. (available at: <http://researchdata.cab.unipd.it/id/eprint/521>)
- 26 Fassnacht M, Arlt W, Bancos I, Dralle H, Newell-Price J, Sahdev A, Tabarin A, Terzolo M, Tsagarakis S & Dekkers OM. Management of adrenal incidentalomas: European Society of Endocrinology Clinical Practice Guideline in collaboration with the European Network for the study of adrenal tumors. *European Journal of Endocrinology* 2016 **175** G1–G34. (<https://doi.org/10.1530/EJE-16-0467>)
- 27 Lazarus J, Brown RS, Daumerie C, Hubalewska-Dydejczyk A, Negro R & Vaidya B. 2014 European Thyroid Association guidelines for the management of subclinical hypothyroidism in pregnancy and in children. *European Thyroid Journal* 2014 **3** 76–94. (<https://doi.org/10.1159/000362597>)
- 28 Ross DS, Burch HB, Cooper DS, Greenlee MC, Laurberg P, Maia AL, Rivkees SA, Samuels M, Sosa JA, Stan MN, *et al.* 2016 American Thyroid Association guidelines for diagnosis and management of hyperthyroidism and other causes of thyrotoxicosis. *Thyroid* 2016 **26** 1343–1421. (<https://doi.org/10.1089/thy.2016.0229>)
- 29 Dekkers OM, Timmermans JM, Smit JW, Romijn JA & Pereira AM. Comparison of the cortisol responses to testing with two doses of ACTH in patients with suspected adrenal insufficiency. *European Journal of Endocrinology* 2011 **164** 83–87 (<https://doi.org/10.1530/EJE-10-0621>)
- 30 Giordano R, Picu A, Bonelli L, Balbo M, Berardelli R, Marinazzo E, Corneli G, Ghigo E & Arvat E. Hypothalamus–pituitary–adrenal axis evaluation in patients with hypothalamo–pituitary disorders: comparison of different provocative tests. *Clinical Endocrinology* 2008 **68** 935–941. (<https://doi.org/10.1111/j.1365-2265.2007.03141.x>)
- 31 Ceccato F, Tizianel I, Vedolin CK, Boscaro M, Barbot M & Scaroni C. Human corticotropin-releasing hormone tests: 10 years of real-life experience in pituitary and adrenal disease. *Journal of Clinical Endocrinology and Metabolism* 2020 **105** dgaa564. (<https://doi.org/10.1210/clinem/dgaa564>)
- 32 Cross AS, Helen Kemp E, White A, Walker L, Meredith S, Sachdev P, Krone NP, Ross RJ, Wright NP & Elder CJ. International survey on high- and low-dose synacthen test and assessment of accuracy in preparing low-dose synacthen. *Clinical Endocrinology* 2018 **88** 744–751. (<https://doi.org/10.1111/cen.13559>)
- 33 Deutschbein T, Broecker-Preuss M, Flitsch J, Jaeger A, Althoff R, Walz MK, Mann K & Petersenn S. Salivary cortisol as a diagnostic tool for Cushing’s syndrome and adrenal insufficiency: improved screening by an automatic immunoassay. *European Journal of Endocrinology* 2012 **166** 613–618. (<https://doi.org/10.1530/EJE-11-0945>)

Received in final form 10 August 2021

Accepted 23 August 2021

Accepted Manuscript published online 23 August 2021