Salivary cortisol as a diagnostic tool for Cushing’s syndrome and adrenal insufficiency: improved screening by an automatic immunoassay

Timo Deutschbein, Martina Broecker-Preuss, Jörg Flitsch¹, Andrea Jaeger, Ricarda Althoff, Martin K Walz², Klaus Mann and Stephan Petersenn

Department of Endocrinology and Division of Laboratory Research, Medical Center. University Hospital of Essen, University of Duisburg-Essen, Hufelandstrasse 55, 45122 Essen, Germany, ¹Pituitary Unit, Department of Neurosurgery, University Hospital Hamburg-Eppendorf, Hamburg, Germany and ²Department of Surgery and Center of Minimally Invasive Surgery, Kliniken Essen-Mitte, Essen, Germany

(Correspondence should be addressed to S Petersenn; Email: stephan.petersenn@endoc-med.de)

Abstract

Background: Salivary cortisol is increasingly used to assess patients with suspected hypo- and hypercortisolism. This study established disease-specific reference ranges for an automated electrochemiluminescence immunoassay (ECLIA).

Methods: Unstimulated saliva from 62 patients with hypothalamic–pituitary disease was collected at 0800 h. A peak serum cortisol level below 500 nmol/l during the insulin tolerance test (ITT) was used to identify hypocortisolism. Receiver-operating characteristic (ROC) analysis allowed establishment of lower and upper cutoffs with at least 95% specificity for adrenal insufficiency and adrenal sufficiency. Saliva from 40 patients with confirmed hypercortisolism, 45 patients with various adrenal masses, and 115 healthy subjects was sampled at 2300 h and after low-dose dexamethasone suppression at 0800 h. ROC analysis was used to calculate thresholds with at least 95% sensitivity for hypercortisolism. Salivary cortisol was measured with an automated ECLIA.

Results: When screening for secondary adrenal insufficiency, a lower cutoff of 3.2 nmol/l and an upper cutoff of 13.2 nmol/l for unstimulated salivary cortisol allowed a highly specific diagnosis (i.e. similar to the ITT result) in 26% of patients. For identification of hypercortisolism, cutoffs of 6.1 nmol/l (sensitivity 95%, specificity 91%, area under the curve (AUC) 0.97) and 2.0 nmol/l (sensitivity 97%, specificity 86%, AUC 0.97) were established for salivary cortisol at 2300 h and for dexamethasone-suppressed salivary cortisol at 0800 h.

Conclusions: The newly established thresholds facilitated initial screening for secondary adrenal insufficiency and allowed excellent identification of hypercortisolism. Measurement by an automated immunoassay will allow broader use of salivary cortisol as a diagnostic tool.

European Journal of Endocrinology 166 613–618 ISSN 0804-4643

Introduction

Measurement of cortisol is mandatory in the diagnostic workup of suspected hypo- and hypercortisolemic states and usually involves obtaining blood and/or urine samples (1, 2, 3, 4). In contrast, saliva sampling is noninvasive, painless, stress free, and requires no special equipment or training (5, 6, 7). Accordingly, even nonprofessionals can easily collect saliva themselves, for instance in their home environment. Samples can then be stored at ambient temperature for at least a week and transported to the laboratory by mail without a significant decrease in cortisol levels, thereby reducing costs and inconvenience. Besides, salivary cortisol appears to be independent of transport proteins like albumin and cortisol-binding globulin (CBG) and therefore reflects the bioactive-free molecule. Hence, it has been increasingly chosen for diagnosing adrenal insufficiency and Cushing’s syndrome (8, 9, 10). As with any other biochemical parameter, however, the reliability of salivary cortisol is crucially dependent on the quality and performance of the particular analytical procedure applied. In this context, modern automated immunoassays may offer several advantages: they are widely available, relatively cheap, and easy to use. Furthermore, they have a rapid turnaround time on a large number of samples, require small volumes of saliva, and demonstrate high analytical accuracy. Inadequate standardization and poor interlaboratory performance remain problematic and precise reference ranges are lacking. As a consequence, this study was designed to calculate disease-specific thresholds for a recently introduced automated electrochemiluminescence immunoassay (ECLIA) for salivary cortisol.
Subjects and methods

Subjects
No study participant had a history of alcohol abuse or psychiatric problems, none of the females was on contraceptives or estrogens, and medications known to affect glucocorticoid metabolism were omitted for at least 24 h before testing. The study protocol was approved by the local ethics committee (approval number: 01-187-1787), and all subjects provided written informed consent.

Patients with hypothalamic–pituitary disease
Sixty-two patients (27 females, 35 males; age: 44.3 ± 2.0 years; body mass index (BMI): 28.4 ± 0.7 kg/m²) with a variety of hypothalamic–pituitary diseases were investigated. At the time of enrollment, nine patients suffered from hypothalamic–pituitary impairment but did not have radiologically detectable tumors within the sellar region (8 × pituitary hormone deficiency of unknown etiology, 1 × traumatic brain injury). The remaining 53 patients suffered from sellar masses: ten patients (four prolactinomas, three nonfunctioning adenomas, two GH-secreting adenomas, and one meningioma) had not been operated and 43 patients (24 nonfunctioning adenomas, seven GH-secreting adenomas, three meningiomas, three craniopharyngiomas, three prolactinomas, one Rathke’s cleft cyst, one astrocytoma, and one ependymoma) had already been surgically treated. The latter subjects were generally tested at least 3 months after intervention (median postoperative interval: 9.5 months; range: 3.0–168.0 months). No subject had to be excluded because of potential contraindications to insulin-induced hypoglycemia. Consequently, the insulin tolerance test (ITT) was used as a gold standard for the evaluation of hypothalamic–pituitary–adrenal axis, defining a normal test result as a peak serum cortisol value of ≥500 nmol/l in response to a laboratory blood glucose level of <40 mg/dl.

Patients with hypercortisolism
Forty patients (33 females, seven males; age: 52.3 ± 1.9 years; BMI: 29.1 ± 0.8 kg/m²) with confirmed hypercortisolism were enrolled. The biochemical diagnosis was made with the measurement of ACTH, repeatedly elevated 24 h urinary-free cortisol levels, abnormal midnight serum cortisol, and/or insufficient serum cortisol suppression during the low-dose dexamethasone suppression test. Moreover, most patients presented with typical signs and symptoms of prolonged and inappropriate exposure to excessive concentrations of glucocorticoids, such as obesity or weight gain, facial fullness, purple striae, and hypertension. Twenty-six patients suffered from cortisol-producing adrenal masses but had not been surgically treated before testing. Two patients had ectopic Cushing’s syndrome, and Cushing’s disease was found in 12 patients. Four of the latter patients had undergone transsphenoidal adnomentomy (median postoperative interval: 59.5 months; range: 15.9–257.9 months), but because of residual or recurrent disease, they demonstrated clear clinical and biochemical hypercortisolism at study entry. The remaining patients were untreated.

Patients with adrenal masses
Forty-five patients (25 females, 20 males; age: 50.8 ± 2.2 years; BMI: 26.9 ± 0.7 kg/m²) with adrenal masses served as patient controls. Eighteen of these patients were found to have nonfunctioning adenomas, 15 patients had histologically confirmed pheochromocytomas, and 12 patients with elevated plasma aldosterone concentration to plasma renin activity ratios, positive suppression tests and/or typical clinical symptoms were diagnosed as having aldosterone-producing adenomas.

Control subjects
One hundred and fifteen healthy control subjects (60 females, 55 males; age: 40.1 ± 1.3 years; BMI: 25.7 ± 0.4 kg/m²) without any age or BMI restrictions were recruited from the general population. These subjects had neither signs and symptoms nor a history of severe and/or chronic illness (especially of endocrine origin). Subjects who were currently or in the previous months taking drugs known to interfere with the synthesis or metabolism of endocrine parameters were not included.

Sample collection
All patients were tested on an inpatient basis. Patients with hypothalamic–pituitary diseases underwent an ITT before unstimulated saliva samples were collected at 0800 h. With respect to patients with confirmed hypercortisolism or noncortisol-producing adrenal masses, saliva samples were collected at 2300 h. Afterward, dexamethasone at a dose of 1 mg was administered orally, and saliva samples were taken at 0800 h the next morning. Tests on healthy control subjects were conducted on an outpatient basis. Saliva samples were collected at 2300 h, and a subgroup of 19 control subjects also underwent a low-dose dexamethasone suppression test (as described earlier). All study participants were instructed in the proper conditions of saliva collection before the Salivette sampling device (Sarstedt, Rommelsdorf, Germany) was handed out. In brief, brushing of teeth, smoking, eating, and/or drinking were not allowed during a 30 min interval before collection of saliva samples. Saliva was collected by chewing on the cotton tube for ~2 min, and samples were frozen at −20 °C until thawed for analysis.
**Hormonal evaluation**

All measurements were performed by experienced personnel in a single laboratory. Plasma ACTH was determined by a solid-phase two-site sequential chemiluminescent immunometric assay (Immulite 2000, Siemens, Eschborn, Germany).

**Electrochemiluminescence immunoassay**

An ECLIA (Roche) was used in combination with the automatic ‘Modular Analytics E170’ apparatus. Endogenous salivary cortisol was determined by its ability to compete with cortisol derivatives (ruthenium-labeled complexes) for the binding sites of a biotinylated polyclonal antibody. The lower limit of detection of this assay was 0.6 nmol/l, and the lower limit of quantification was 0.5 nmol/l. Values below the limit of quantification were set to 0.5 nmol/l. The intra-assay variations (both mean ± S.D.) were 11% for 8.9 ± 1.0 nmol/l (n = 14) and 6% for 19.1 ± 1.2 nmol/l (n = 16), while the interassay variations (both mean ± S.D.) were 9% for 9.7 ± 0.9 nmol/l (n = 19) and 5% for 19.9 ± 1.0 nmol/l (n = 19). The cross-reactivity of steroids structurally related to cortisol was as follows: corticosterone, 5.8%; 11-deoxycortisol, 4.1%; 17α-hydroxyprogesterone, 1.5%; 11-deoxycorticosterone, 0.7%; progesterone, 0.4%; cortisone and prednisone, each 0.3%; and dexamethasone, 0.1%.

**RIA**

Salivary cortisol was also assayed using a modification of ‘GammaCoat’ RIA (DiaSorin, Stillwater, MN, USA), decreasing the sample volume from 200 to 100 μl. The lower limit of detection of this assay was 0.6 nmol/l, and the intra- and interassay coefficients of variation were 2.6 and 4.6% respectively, as described previously (10). The antiserum had 100% cross-reactivity to cortisol, 77.0% to corticosterone, 5.8% to 11-deoxycortisol, 6.3% to 17α-hydroxyprogesterone, 6.3% to 11-deoxycorticosterone, 1.2% to cortisone, 1.2% to both 17-hydroxyprogesterone and prednisone, 0.3% to corticosterone, 0.2% to both dexamethasone and dihydrocortisone, and 0.1% to both deoxycorticosterone and tetrahydrocortisone.

**Statistical analysis**

Results are expressed as the mean ± S.E.M. unless otherwise stated. The diagnostic accuracy of unstimulated salivary cortisol at 0800 h, salivary cortisol at 2300 h, and dexamethasone-suppressed salivary cortisol at 0800 h was investigated using receiver-operating characteristic (ROC) analysis and the area under the curve (AUC). The Kruskal–Wallis tests (followed by Dunn’s multiple comparison tests) were performed where appropriate. The Spearman correlation and P values are provided (statistical significance was taken as P < 0.05). GraphPad Prism 5.0 software (GraphPad Software, Inc., San Diego, CA, USA) was used for statistical calculations. Conversion factor for nmol/l to μg/dl: divide by 27.59.

**Results**

**Biochemical workup of the different patient groups before measurement of salivary cortisol**

All patients tested for adrenal insufficiency developed symptomatic hypoglycemia (with blood glucose levels below 40 mg/dl) but did not experience any severe side effects in response to insulin administration during the ITT. Thirty-two patients responded normally to hypoglycemia (i.e. adrenal sufficient), whereas 30 patients had a subnormal response (i.e. adrenal insufficient). In detail, adrenal insufficient patients had a mean peak serum cortisol of 266.5 ± 30.4 nmol/l to insulin-induced hypoglycemia, whereas adrenal sufficient patients demonstrated a mean peak serum cortisol of 586.7 ± 17.9 nmol/l. Hypercortisolemic patients had a mean dexamethasone-suppressed serum cortisol of 480.6 ± 35.2 nmol/l. Furthermore, mean plasma ACTH was 5.3 ± 0.6 pg/ml in patients with cortisol-producing adenals masses (with all patients having ACTH levels of < 10 pg/ml) and 66.8 ± 14.1 pg/ml in patients with ACTH-dependent Cushing’s syndrome. On the contrary, patients with noncortisol-producing adrenal tumors had a mean serum cortisol of 43.6 ± 4.9 nmol/l after dexamethasone.

**Correlation of salivary cortisol concentrations**

A significant correlation between salivary cortisol concentrations derived from the currently evaluated ECLIA and a previously established RIA (‘GammaCoat’ RIA for cortisol; DiaSorin) was detected (r = 0.84, P < 0.0001).

**Individual and mean salivary cortisol concentrations**

Individual salivary cortisol levels at various time points are shown in Fig. 1, whereas means of subgroups are listed in Table 1. Highly significant differences between mean levels from patients with confirmed hypercortisolism and control subjects (i.e. patients with adrenal tumors and healthy controls) were detected (P < 0.01 for both salivary cortisol levels at 2300 h and P < 0.001 for salivary cortisol levels at 0800 h after 1 mg dexamethasone).

**Establishment of cutoffs for unstimulated 0800 h salivary cortisol**

ROC analysis of unstimulated salivary cortisol levels at 0800 h allowed establishment of a lower cutoff of 3.2 nmol/l with ≥ 95% specificity for diagnosing...
adrenal insufficiency (sensitivity 40%, specificity 97%, AUC 0.78) and an upper cutoff of 13.2 nmol/l with ≥ 95% specificity for diagnosing adrenal insufficiency (sensitivity 13%, specificity 97%, AUC 0.78). If these cutoffs were simultaneously applied, the ITT results were confirmed in 16 of 62 patients (26%). The remaining patients had salivary cortisol levels between the lower and the upper cutoff and were therefore thought to require further diagnostic evaluation.

**Establishment of cutoffs for diagnosing hypercortisolism**

ROC analysis was carried out by comparing hypercortisolemic patients with a control group (including patients with noncortisol-producing adrenal masses and healthy control subjects). Thresholds with at least 95% sensitivity for diagnosing hypercortisolism were as follows: 6.1 nmol/l (sensitivity 95%, specificity 91%, AUC 0.97) for salivary cortisol at 2300 h and 2.0 nmol/l (sensitivity 97%, specificity 86%, AUC 0.97) for dexamethasone-suppressed salivary cortisol at 0800 h. Of note, these two highly sensitive cutoffs had AUCs with overlapping 95% confidence intervals, indicating that differences of specificity were not statistically significant. False-positive tests were observed in 9% (for salivary cortisol at 2300 h) and 14% (for salivary cortisol after dexamethasone) of patients.

**Discussion**

As saliva sampling offers several advantages over obtainment of blood and/or urine (5, 6, 7), salivary cortisol has been increasingly chosen for diagnosing adrenal insufficiency and Cushing’s syndrome (8, 9, 10). Our current study was designed to calculate diseasespecific thresholds for a current automated ECLIA. When comparing the results for late-night salivary cortisol with previous studies on this particular assay, our cutoff of 6.1 nmol/l was slightly lower than the 95th percentile of 8.9 nmol/l published by Vogeser et al. (11). This difference may be explained by the fact that Vogeser et al. exclusively investigated healthy control subjects, thereby providing an upper limit of normal instead of a ROC-generated cutoff. A threshold of 9.7 nmol/l (derived from ROC analysis as in our study) was proposed by Beko et al. (12). In their study, 126 patients with symptoms of glucocorticoid excess, obesity, and/or incidentally discovered adrenal masses were consecutively evaluated by measurement of serum cortisol (at 0800 h, at midnight, and after dexamethasone suppression) and plasma ACTH (at 0800 h). By these means, hypercortisolism was confirmed in nine patients. The lower number of patients with Cushing’s syndrome may explain the slightly higher cutoff compared with our study.

In contrast to the well-established measurement of late-night salivary cortisol, less is known about salivary cortisol determination after low-dose dexamethasone suppression. Up to now, only a few groups investigated the validity of this method, providing cutoffs between 1.5 and 3.7 nmol/l (10, 13, 14, 15, 16). Although the
Unstimulated salivary cortisol is consistent with our own experience, demonstrating that cortisol does not reliably distinguish between hypercortisolism, secondary adrenal insufficiency (18, 19). Hence, the optimal diagnostic tool is still a matter of ongoing debate. Although a single threshold for early morning cortisol appears to have similar diagnostic value. However, this is not in line with some other reports published over the last decade (10, 14, 16). All these studies, including our own previous publication, demonstrated that dexamethasone-suppressed salivary cortisol had comparable specificity. This has also been described in a recently published meta-analysis by Elamin et al. (4) who therefore concluded that both screening tests appear to have similar diagnostic value. However, this is not in line with some other reports published over the last decade (10, 14, 16). All these studies, including our own previous publication, demonstrated that dexamethasone-suppressed salivary cortisol (with specificities ranging from 83 to 100%) was preferable to late-night salivary cortisol (with specificities ranging from 69 to 98%). Further research is needed to clarify whether these controversies may be attributed to insufficient physical and/or psychological rest for control subjects before saliva sampling.

In conclusion, patients with suspected or proven secondary adrenal insufficiency by measurement of spontaneous 0800 h salivary cortisol levels. Although the ITT is widely regarded as the gold standard for this purpose, this test is often uncomfortable, limited by numerous contraindications, and sometimes even life threatening. Consequently, alternative means have been evaluated, for instance measurement of cortisol after application of CRH or ACTH. Nevertheless, the CRH test is rather cumbersome and expensive, whereas a meta-analysis of the ACTH test indicated that the AUC for secondary adrenal insufficiency was significantly lower than the AUC for primary adrenal insufficiency (17). This is consistent with our own experience, demonstrating that both tests did not sufficiently identify patients with secondary adrenal insufficiency (18, 19). Hence, the optimal diagnostic tool is still a matter of ongoing debate. Although a single threshold for early morning cortisol does not reliably distinguish between insufficient and sufficient adrenal reserve, we have repeatedly reported on the potential advantages of using both a high (with high specificity for adrenal sufficiency) and a low (with high specificity for adrenal insufficiency) cutoffs for serum and salivary cortisol (9, 19). Of note, such an attempt for prescreening for adrenal insufficiency has also been suggested in a recent review, although published data on lower and upper cutoffs for basal salivary cortisol derived from ROC analysis are still scarce (20).

Our formerly presented thresholds (5.0 and 21.1 nmol/l) for the ‘GammaCoat’ RIA from DiaSorin were somewhat higher than the currently presented ECLIA results (9). Nevertheless, similar discrepancies have been described by other groups. For instance, when testing the same saliva samples in different assays, a wide variability in the absolute cortisol concentrations was detected (21, 22), and these observations were also made for serum (23) and urinary cortisol (24, 25). Accordingly, determination of salivary cortisol requires careful evaluation of the particular testing procedure applied, followed by application of individually established sensitive cutoffs. This is especially true because concentrations are at far lower levels than in serum, being close to the functional limit of detection of most assays.

ECLIA from Roche was not applied in any of these studies, our threshold of 2.0 nmol/l is excellently in line with the previous data derived from multiple non-automated immunoassays. Our current results imply that the highly sensitive thresholds for late-night and dexamethasone-suppressed salivary cortisol had comparable specificity. This has also been described in a recently published meta-analysis by Elamin et al. (4) who therefore concluded that both screening tests appear to have similar diagnostic value. However, this is not in line with some other reports published over the last decade (10, 14, 16). All these studies, including our own previous publication, demonstrated that dexamethasone-suppressed salivary cortisol (with specificities ranging from 83 to 100%) was preferable to late-night salivary cortisol (with specificities ranging from 69 to 98%). Further research is needed to clarify whether these controversies may be attributed to insufficient physical and/or psychological rest for control subjects before saliva sampling.

In addition, we analyzed the performance characteristics of the Roche ECLIA for the investigation of patients with suspected or proven secondary adrenal insufficiency by measurement of spontaneous 0800 h salivary cortisol levels. Although the ITT is widely regarded as the gold standard for this purpose, this test is often uncomfortable, limited by numerous contraindications, and sometimes even life threatening. Consequently, alternative means have been evaluated, for instance measurement of cortisol after application of CRH or ACTH. Nevertheless, the CRH test is rather cumbersome and expensive, whereas a meta-analysis of the ACTH test indicated that the AUC for secondary adrenal insufficiency was significantly lower than the AUC for primary adrenal insufficiency (17). This is consistent with our own experience, demonstrating that both tests did not sufficiently identify patients with secondary adrenal insufficiency (18, 19). Hence, the optimal diagnostic tool is still a matter of ongoing debate. Although a single threshold for early morning cortisol does not reliably distinguish between insufficient and sufficient adrenal reserve, we have repeatedly reported on the potential advantages of using both a high (with high specificity for adrenal sufficiency) and a low (with high specificity for adrenal insufficiency) cutoff for serum and salivary cortisol (9, 19). Of note, such an attempt for prescreening for adrenal insufficiency has also been suggested in a recent review, although published data on lower and upper cutoffs for basal salivary cortisol derived from ROC analysis are still scarce (20).

Our formerly presented thresholds (5.0 and 21.1 nmol/l) for the ‘GammaCoat’ RIA from DiaSorin were somewhat higher than the currently presented ECLIA results (9). Nevertheless, similar discrepancies have been described by other groups. For instance, when testing the same saliva samples in different assays, a wide variability in the absolute cortisol concentrations was detected (21, 22), and these observations were also made for serum (23) and urinary cortisol (24, 25). Accordingly, determination of salivary cortisol requires careful evaluation of the particular testing procedure applied, followed by application of individually established sensitive cutoffs. This is especially true because concentrations are at far lower levels than in serum, being close to the functional limit of detection of most assays.

Of note, when comparing the currently evaluated ECLIA with an in-house RIA, Beko et al. (12) (r=0.98) as well as van Aken et al. (26) (r=0.84) observed an excellent agreement between the two measurement techniques. This is in line with our recent data, emphasizing a good correlation between the Roche ECLIA and a commercially available RIA (r=0.84).

In conclusion, patients with suspected or proven secondary adrenal insufficiency should primarily undergo a stimulation test to assess adrenocortical reserve. However, if basal cortisol levels are available, our current data suggest that the simultaneous application of an upper cutoff with high specificity for adrenal sufficiency and a lower cutoff with high specificity for adrenal insufficiency may obviate dynamic testing in patients who have remarkably low

---

Table 1: Salivary cortisol levels at various time points (i.e. at 0800 h without prior administration of 1 mg dexamethasone, at 0800 h with prior administration of 1 mg dexamethasone). Values are ranges and 95% confidence intervals. Conversion factor for nmol/l to μg/dl: divide by 27.59.

<table>
<thead>
<tr>
<th>Time point</th>
<th>Adrenally insufficient</th>
<th>Adrenally sufficient</th>
<th>Patients with hypercortisolism</th>
<th>Patients with noncortisol-secreting adrenal tumors and healthy controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unstimulated salivary cortisol at 0800 h (nmol/l)</td>
<td>0.7–16.3 (3.5–6.2)*</td>
<td>2.2–35.0 (2.9–28.2)*</td>
<td>–</td>
<td>3.3–109.2 (14.8–30.6)†</td>
</tr>
<tr>
<td>Salivary cortisol at 2300 h (nmol/l)</td>
<td>–</td>
<td>–</td>
<td>0.5–21.2 (2.3–3.1)†</td>
<td>–</td>
</tr>
<tr>
<td>Dexamethasone-suppressed</td>
<td>0.8–86.5 (8.7–21.1)‡</td>
<td>0.5–7.4 (0.8–1.6)‡</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

The symbols indicate significant differences between pairs: * not significant; †P<0.01; ‡P<0.001.
or high basal cortisol levels (usually representing a substantial number of cases, for instance 26% in our current study). As the newly established thresholds also allowed excellent identification of hypercortisolism, measurement of salivary cortisol by an automated immunoassay would allow for broader use of this parameter as a diagnostic tool.

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

Acknowledgements

The authors would like to thank Helmut Tourné for skillful technical support.

References


Received 1 November 2011
Revised version received 18 December 2011
Accepted 3 January 2012