

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Importance of sampling time and assay cut-offs for routine assessment of adrenal function: an observational longitudinal study.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-019273
Article Type:	Research
Date Submitted by the Author:	22-Aug-2017
Complete List of Authors:	Ortiz-Flores, Andrés; Hospital Universitario Ramón y Cajal, Endocrinology Santacruz, Elisa; Hospital Universitario Ramon y Cajal Jiménez-Mendiguchia, Lucía; Hospital Universitario Ramon y Cajal García-Cano, Ana; Hospital Universitario Ramon y Cajal Nattero-Chávez, Lía; Hospital Universitario Ramon y Cajal Escobar-Morreale, Héctor; Hospital Universitario Ramon y Cajal Luque-Ramírez, Manuel; Hospital Universitario Ramon y Cajal
Primary Subject Heading:	Diabetes and endocrinology
Secondary Subject Heading:	Diagnostics, Patient-centred medicine
Keywords:	Adrenal disorders < DIABETES & ENDOCRINOLOGY, cortisol, biochemical diagnosis, specificity, immunoassay

SCHOLARONE™
Manuscripts

Only

1 **Title** Importance of sampling time and assay cut-offs for routine assessment of adrenal function:
2 an observational longitudinal study.

3
4 **Short title** Cosyntropin test in adrenal insufficiency.

5
6 **Authors** Andrés E. Ortiz-Flores, M.D.,¹ Elisa Santacruz-Cerdá, M.D.,¹ Lucía Jiménez-
7 Mendiguchia, M.D.,² Ana García-Cano, M.D.,² Lia Nattero Chávez, M.D.,¹ Héctor F. Escobar-
8 Morreale, M.D., Ph.D.,^{1,3} Manuel Luque-Ramírez, M.D., Ph.D.^{1,3}

9
10 **Affiliations**

11 ¹ Department of Endocrinology and Nutrition, Hospital Universitario Ramón y Cajal, Madrid
12 (Spain).

13 ² Department of Clinical Biochemistry, Hospital Universitario Ramón y Cajal, Madrid (Spain).

14 ³ Diabetes, Obesity and Human Reproduction Research Group, Hospital Universitario Ramón y
15 Cajal & Universidad de Alcalá & Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS) &
16 Centro de Investigación Biomédica en Red Diabetes y Enfermedades Metabólicas Asociadas
17 (CIBERDEM), Madrid, Spain.

18
19 **Word count:** 3,506; **Word count for abstract:** 288; **Figures:** 4; **Tables:** 2; **Number of**
20 **references:** 21.

21
22 **CORRESPONDENCE & REPRINTS:** Andres E. Ortiz-Flores, MD. Department of
23 Endocrinology and Nutrition, Hospital Universitario Ramón y Cajal, Carretera de Colmenar km
24 9.1, E-28034 Madrid, Spain. Phone/Fax: +34-91 336 9029. E-mail: andres_ortiz_f@yahoo.com

25 ABSTRACT

26 **Objectives** Serum cortisol concentrations after adrenal stimulation with a high-dose cosyntropin
27 bolus is the test of choice for diagnosis of primary and non-acute central adrenal insufficiency.

28 We aim to: i) assess the role of 30-min and 60-min sample timing, and the importance of assay-
29 specific normative cut-offs concentrations for adrenal insufficiency diagnosis, ii) to estimate
30 specificity and positive predictive value of 30-min and 60-min sampling time, and iii) to
31 establish an assay-specific lower limit of normality of serum cortisol concentrations after
32 cosyntropin stimulation.

33 **Design and Setting:** Observational retrospective study performed in a tertiary-level Spanish
34 hospital between 2011 and 2015.

35 **Participants and interventions:** Two groups were evaluated: i) A main study group that m 406
36 patients in whom serum cortisol was measured at 30 and 60 minutes after cosyntropin
37 stimulation, and ii) a confirmative group that included 153 women with a normal hypothalamic-
38 pituitary-adrenal axis in whom a cosyntropin test was conducted for other reasons. Diagnostic
39 agreement between sampling times was analysed considering classic (500 nmol/l) and assay-
40 specific serum cortisol cut-off concentrations.

41 **Results** Diagnostic agreement was greater when applying assay-specific cut-off values instead of
42 those derived from the literature. For suspected primary adrenal insufficiency, serum cortisol
43 measured 30-min after cosyntropin administration was enough to make a diagnosis in over 95%
44 of cases, without missing any necessary treatment. For central adrenal insufficiency suspicion,
45 60-min cortisol concentrations were more specific, establishing diagnosis in over 97% of cases.

46 **Conclusions:** Assay-specific cut-off cortisol concentrations instead of classic literature values
47 improve the diagnostic accuracy of the cosyntropin test. For primary adrenal insufficiency, 30-

1
2
3 48 min subnormal serum cortisol responses are a reliable marker of adrenal function. On the
4
5 49 contrary, when secondary or tertiary adrenal insufficiency is suspected, a 60-min cortisol
6
7 50 measurement improves the diagnostic accuracy of the test.
8
9

10 51

11 52 Strengths and limitations:

12
13
14 53 A large series of subjects with suspected AI evaluated with a standardized dynamic study, in
15
16
17 54 whom, a systematic review of their clinical recordings was performed.
18

19 55 The present study enhances the importance of the use of local normative thresholds for adrenal
20
21 56 function assessment, situation than in the clinical practice is rarely considered among physicians.
22

23 57 Our present results may not be extrapolable to other populations in whom SC has been measured
24
25
26 58 with different immunoassays that would require different local normative data.
27

28 59 Analysis of specificity and positive predictive value has not been challenged against a
29
30
31 60 biochemical gold-standard in most cases and, we have not been able to establish false negative
32
33 61 rates, sensitivity and negative predictive values
34

35 62

36
37 63 **Keywords:** Adrenal insufficiency; biochemical diagnosis; cortisol; immunoassay; specificity.
38
39

40 64

41 65

42 66

43 67

44 68

45 69

46 70

47 71

72

73 Introduction

74 The laboratory diagnosis of adrenal insufficiency (AI) at the clinical setting lies on the
75 finding of an inappropriately low morning circulating serum cortisol (SC) or subnormal SC
76 responses to adrenal stimulation (1). Classic adrenal stimulation protocol consists of SC
77 measurements at 30 and 60 minutes after a single 250 µg intravenous bolus or intramuscular
78 injection of tetracosactide (cosyntropin). A SC value 30 and/or 60-min after cosyntropin
79 administration ≥ 500 nmol/l (18 µg/dl) is considered as a normal response (2). This protocol, also
80 known as a short standard high-dose test (SST), is the dynamic exploration of choice for primary
81 AI diagnosis (1, 3) and it is also used for non-acute central AI (4, 5). In critically ill patients, SST
82 may be performed to rule out a functional form of AI –critical illness-related corticosteroid
83 insufficiency– in subjects showing sustained refractory hypotension and no response to
84 vasopressors (2, 6). Although a definite biochemical definition for this condition is lacking, a SC
85 increase above 248 nmol/l (9 µg/dl) in response to cosyntropin is associated with a good
86 prognosis in septic patients (7).

87 The most appropriate sampling time for SC during the SST is controversial. SC
88 measurements 30 min after SST have been validated against a “gold standard” such as an insulin
89 tolerance test (8). Thus, some authors suggest that a single 30 min SC is enough to establish or
90 rule out a clinically significant AI (4, 9, 10). Other studies show that a 60 min sample may avoid
91 unnecessary overdiagnosis (11-13). Recent clinical practice guidelines recommend further
92 research to clarify whether 60 min SC sampling might be more specific for AI diagnosis (3, 14).

93 Liquid chromatography/mass spectrometry techniques are currently recommended for the
94 accurate measurement of circulating steroids. But in most centres, clinical routine still relies on
95 automated immunoassays for SC measurement (15). Considering that the classic cut-off value

1
2
3 96 for the SST was established for SC as measured by older radioimmunoassays, and that
4
5 97 immunochemiluminescent assays differ in antibody specificity with radioimmunoassays (16),
6
7 98 establishing local assay-specific cut-off values is of paramount importance to properly classify
8
9 99 SC responses to cosyntropin (3, 16, 17). This fact is not a spurious one because despite being
10
11 well established that local assay-specific lower limits of normality (LLN) should be used for
12
13 100 dynamic assessments of the HPA axis (3), in our experience many physicians are still using
14
15 101 classic cut-offs in their routine practice. Also, other factors that may influence SC measurement
16
17 102 include the stimulation of hepatic synthesis and secretion of cortisol binding globulin by
18
19 103 oestrogens, sex and several non-glucocorticoid drugs (18).
20
21 104

22
23
24 105 To provide new insights into still open questions, our study's aims were: i) to assess the
25
26 106 concordance between 30 and 60 min SC concentrations after SST at the clinical setting; ii) to
27
28 107 estimate the diagnostic agreement between both sampling times when using literature or assay-
29
30 108 and sex-specific cut-offs values, taking into account the origin of AI; iii) to estimate the
31
32 109 specificity (Sp) and positive predictive value (PPV) of 30 and 60 min sampling times while
33
34 110 taking into account the origin of AI; and iv) to confirm assay-specific LLN for SC concentration
35
36 111 after SST in a group of subjects with a normal hypothalamic-pituitary-adrenal (HPA) function.
37
38
39
40 112

41 42 113 **Subjects and methods**

43
44 114 From January 1, 2011 to December 31, 2015 we conducted a longitudinal observational
45
46 115 study, performed in a third-level Spanish hospital, where we assessed SC responses to SST in
47
48 116 two study populations:

49
50
51 117 i) A subgroup of adult subjects (n = 451) in whom 0, 30 and 60 min SC concentrations were
52
53 118 assayed during a SST conducted at the clinical setting for suspected AI (main study
54
55 119 population).

1
2
3 120 ii) A group of women with normal HPA axis (n = 153) prospectively recruited from our
4
5 121 Reproductive Endocrinology clinic during the study of functional hyperandrogenism whom 0
6
7 122 and 30 min SC concentrations were obtained during a SST performed for routine screening
8
9
10 123 of non-classic congenital adrenal hyperplasia by a local study protocol including SC values at
11
12 124 those sampling times (confirmative group). Non-classic congenital adrenal hyperplasia
13
14 125 screening was negative in all cases.

16
17 126 Before conducting the study, we obtained approval from the local ethics committee. All
18
19 127 women from our Reproductive Endocrinology clinic had previously signed an informed consent
20
21 128 form for the inclusion of a selection of coded clinical variables in an electronic database for
22
23 129 clinical research purposes that included the SC measurements presented here.
24
25

26 130

28 131 *Main study population*

30 132 Basal and stimulated SC values were extracted from the electronic database of our
31
32 133 Department of Clinical Biochemistry. We collected a minimum dataset in an electronic case
33
34 134 form from the clinical records of the patients including age, sex, weight, height, laboratory
35
36 135 measurements at the dates when the SST was conducted such as circulating electrolytes,
37
38 136 glomerular filtration rate and basal ACTH concentrations, clinical suspicion of primary or central
39
40 137 AI, other dynamic tests performed for the evaluation of adrenal function, history of pituitary
41
42 138 disease, time from hypothalamic-pituitary insult to SC determination, administration of drugs
43
44 139 that may interfere with the HPA axis, and the immunoassay used for SC assay. Baseline
45
46 140 characteristics of study population are shown in **Table 1**.

48
49 141 We considered a clinical suspicion of primary AI in cases when the patient had some known
50
51 142 adrenal disease, had required mineralocorticoid supplementation latter in their follow-up, had
52
53 143 received drugs that may interfere with cortisol biosynthesis, had not clinical suspicion of any
54
55
56
57

144 hypothalamic-pituitary condition, and had not developed such a condition latter in time.
145 Conversely, we suspected a central AI in subjects known to suffer from hypothalamic-pituitary
146 disease, or had received drugs that may suppress the HPA axis.

147 We excluded from analysis those subjects stimulated with cosyntropin doses other than 250
148 μg ($n = 4$), subjects aged below 18-year old ($n = 35$), and subjects from whom we could not
149 obtain enough information from their clinical records as to establish a reason for conducting a
150 SST ($n = 6$). Therefore, the study group finally included in the analyses consisted of 406
151 subjects.

152

153 *Confirmative group*

154 The results of SST from 153 premenopausal women with a normal HPA axis aged from 14 to
155 42 years old were included. Three women showed clearly subnormal SC responses and were
156 consequently excluded from the study: in two, the HPA axis suppressive effect of progestins
157 administered during 10 days before the SST with the aim of inducing a withdrawal bleed could
158 justify the abnormal results; in the other case, we could not establish the cause of the subnormal
159 response with certainty because the patient was lost to follow-up.

160

161 *Assays*

162 During the study period, two immunoassays were used in our centre: i) *Siemens Immulite*
163 *2000*[®] *Cortisol Immunoassay System* from 2011 to July 1, 2013 (immunoassay 1) and ii) *Abbot*
164 *Laboratories Diagnostics Division Architect*[®] *Cortisol Immunoassay System* from 2013 August
165 1, 2013 to December 31, 2015 (immunoassay 2).

166

167 *Agreement among 30 and 60 min sampling times*

1
2
3 168 We analysed diagnostic agreement between the 30 and 60 min SC using two different
4
5 169 stimulated SC LLN: i) the classic ≥ 500 nmol/l (3), and ii) sex and assay-specific cut-off values
6
7
8 170 taking into account concomitant COC use if necessary (14). For immunoassay 1, the reported
9
10 171 LLN (2.5th percentile) was 470 nmol/l (17 μ g/dl) in men and women, and 690 nmol/l (25 μ g/dl)
11
12 172 for women taking COC. For immunoassay 2, the LLNs were 441 nmol/l (16 μ g/dl) for men, 414
13
14 173 nmol/l (15 μ g/dl) for women, and 579 nmol/l (21 μ g/dl) for women taking COC (16).
15
16
17 174

19 175 *Statistical analysis*

21 176 Data are shown as mean \pm standard deviation or 95% confidence interval (CI), median
22
23 177 (minimum-maximum), and raw numbers (percentage) as needed. Normal distribution of
24
25
26 178 continuous variables was assessed by the Kolmogorov-Smirnov test for one sample after a two-
27
28 179 step approach for transforming skewed variables if necessary (19). Comparisons among
29
30
31 180 continuous variables were performed by the Student's t test or a repeated-measure ANOVA.
32
33 181 Comparisons among categorical variables were performed by Fisher's exact or χ^2 tests as
34
35 182 appropriate. Pearson's analysis was used to correlate 30 and 60-min sampling times. Consistency
36
37
38 183 and absolute agreement among both point times of SST were determined by their intraclass
39
40 184 correlation coefficient (ICC) with a two-factor and random-effect model. Quantitative agreement
41
42 185 was graphically assessed by Bland-Altman plots. Biochemical agreement in the diagnosis of
43
44 186 normal or subnormal adrenal was assessed by using the kappa (κ) coefficient. True positives
45
46
47 187 (TP) were defined as SSTs showing subnormal cortisol responses at both time points. True
48
49 188 negatives (TN) were defined as SSTs showing a normal cortisol response at both time points in
50
51 189 patients who did not need glucocorticoid replacement during their follow-up, did not suffer an
52
53
54 190 adrenal crisis, and did not have any other functional HPA test with a subnormal response if
55
56
57
58
59

1
2
3 191 available. False positives (FP) were SSTs showing a subnormal response in only one of the
4
5 192 sampling times but not in the other. We calculated Sp and PPV [$Sp = TN / (TN + FP)$ and $PPV =$
6
7 193 $TP / (TP + FP)$] for each sampling times after the SST. A P value < 0.05 was considered
8
9
10 194 statistically significant.

11
12 195

14 196 **Results**

17 197 *Main study population*

18
19 198 Of 406 cases included (261 tested by immunoassay 1 and 145 tested by immunoassay 2),
20
21 199 168 SSTs were performed to rule out primary AI; 226 to rule out central AI; and 12 SST were
22
23 200 performed in critically ill patients. Regarding all SSTs as a whole, SC concentrations at 30 and
24
25 201 60 min after SST increased when compared to unstimulated values (**Figure 1A**), and SC
26
27 202 concentrations at both 30 and 60 min showed a very strong linear correlation (**Figure 1B**). The
28
29 203 ICC among both sampling times showed a very good consistence index (0.948; 95%CI: 0.937 –
30
31 204 0.957) and a good absolute agreement (0.899, 95%CI: 0.476 – 0.962), although according to the
32
33 205 95%CI lower limit only qualify as fair. The Bland-Altman plot (**Figure 1C**) showed a good
34
35 206 agreement between 30 and 60 min sampling times, with a tendency towards greater differences
36
37 207 with increasing mean values of stimulated SC and only ~ 5% of extreme differences among both
38
39 208 times.

40
41 209 Diagnostic agreement among both times according to classic and to assay-specific cut-off
42
43 210 values is shown in **Figure 2**. With a classic cut-off point, 42 cases (10.3%) had a subnormal
44
45 211 response at 30 min, whereas it was ≥ 500 nmol/l at 60 min. On the contrary, the response was
46
47 212 normal at 30 min in 5 cases (1.2%) but subnormal at 60 min. Using sex- and assay-specific
48
49 213 values, 37 cases (9.1%) had a subnormal response at 30 min but normal at 60 min. In 7 cases
50
51 214 (1.7%), the response was normal at 30 min and subnormal at 60 min.

1
2
3 215 The analysis of the diagnostic agreement as a function of central or primary AI suspicion
4
5 216 is shown in **Figure 3**. As a rule, agreement among sampling times of the SST was better when
6
7 217 primary AI was suspected compared with central AI suspicions. When using classic cut-off
8
9 218 values to rule out primary AI, 8 cases (4.8%) showed a subnormal response at 30 min that
10
11 219 reached normal concentrations at 60 min, whereas no subject with a normal response at 30 min
12
13 220 had a subnormal response at 60 min. Using sex- and assay-specific cut-off values, in 7 cases
14
15 221 (4.2%) the response was subnormal at 30 min but reached normal concentrations at 60 min. Five
16
17 222 of them showed a subnormal SC response after SST that was very close to reaching the cut-off
18
19 223 value. In these subjects, the differences between the cut-off value and the stimulated SC ranged
20
21 224 from 22 to 39 nmol/l (0.8 to 1.4 µg/dl), very small concentrations that are in fact included within
22
23 225 the coefficient of variation of the assay (18,19), thereby suggesting no clinical consequences.
24
25 226 The two remaining patients showed peak SC concentrations of 303 and 360 nmol/l (11 and 13
26
27 227 µg/dl) at the 30 min sampling time: one had received oral glucocorticoid replacement therapy
28
29 228 that did not preclude the patient of responding to cosyntropin by showing a SC of 470 nmol/l (17
30
31 229 µg/dl) at the 60 min sample, and the other subject was submitted to SST because of the presence
32
33 230 of bilateral adrenal hyperplasia and did not show any signs or symptoms of AI nor suffered an
34
35 231 adrenal crisis during follow-up. None of the SSTs showing normal responses at 30 min had a
36
37 232 subnormal response at 60 min.

38
39
40 233 When central AI was suspected and a classic cut-off point was applied, 33 cases (14.6%)
41
42 234 had a normal response at 60 min and subnormal at 30 min. Only 3 subjects (1.3%) presented
43
44 235 with the opposite situation. Regarding specific cut-offs, 30 cases (13.3%) had a normal response
45
46 236 at 60 min but subnormal at 30 min, and in only 5 cases (2.2%) the contrary occurred. These 5
47
48 237 subjects had been evaluated in the context of withdrawal of prolonged glucocorticoid therapy
49
50 238 during the first year after a pituitary injury (surgery and/or pituitary radiotherapy). Three of them

239 showed a complete recovery of their HPA axis throughout the follow-up period, whereas in the
240 other two patients, who had received pituitary radiotherapy, the subnormal response to
241 cosyntropin was maintained over time.

242 The Sp and PPV for different sampling times and cut-off values used here are shown in
243 **Table 2.** SC concentrations at 60 min had a higher Sp and PPV compared with 30 min
244 measurements, particularly when central AI was suspected. Nonetheless, the Sp of the
245 determination at 30 min was as high as 95% when SST had been performed to rule out primary
246 disease both when applying classic or sex- and assay-specific cut-off values.

247 We observed discordant results between classic and specific cut-off values in 50 cases. In
248 47 of these subjects the subnormal response observed considering the classic cut-off value was
249 normal if a sex- and assay-specific cutoff was applied. In 7 of them, SST was performed to rule
250 out a primary AI and the remaining 40 SSTs were performed to rule out a central AI.
251 Glucocorticoid replacement was started in 18 cases, and no subject presented with signs or
252 symptoms of chronic or acute AI. Of the 50 discordant SSTs, 3 were conducted in women under
253 estrogenic therapy and presented a normal response according to the classic cut-off value, but
254 subnormal when considering a sex- and assay-specific cut-off, yet none of them required
255 glucocorticoid therapy.

256 Finally, no discordant results were observed among sampling times and between classic
257 and sex- and assay-specific cut-off values in critically ill patients.

258 259 *Confirmative group*

260 Of these women (n = 150, 97 tested by immunoassay 1 and 53 tested by immunoassay 2),
261 30 (20%) presented with a subnormal response to SST according to classic cut-off values, yet
262 this figure was reduced to 3 (2%) when sex- and assay-specific cut-off values were used

1
2
3 263 [observed agreement: 82%; κ : 0.151 (95%CI: 0.066-0.235)]. The 3 women showing subnormal
4
5 264 response during SST using a sex- and assay-specific cut-off value showed stimulated SC
6
7 265 concentrations of 342 nmol/l (12.4 μ g/dl), 353 nmol/l (12.8 μ g/dl) and 372 nmol/l (13.5 μ g/dl),
8
9
10 266 whereas the LLNs (2.5th percentile) of SC concentrations at 30 min sampling time of SST were
11
12 267 441 and 414 nmol/l for immunoassays 1 and 2, respectively (**Figure 4**). None of these female
13
14 268 controls developed any HPA disease during their follow-up.
15
16
17 269

19 270 **DISCUSSION**

21 271 AI is a clinical condition associated with a high morbidity and mortality. Unstimulated
22
23 272 early morning SC values below 138 nmol/l (5 μ g/dl) show a high PPV for AI, whereas
24
25 273 concentrations over 500 nmol/l predict a normal adrenal response. However, values between 138
26
27 274 and 500 nmol/l are considered indeterminate and require adrenal stimulation to confirm or rule
28
29 275 out a diagnosis (1-3).
30
31
32

33 276 Our data shows that both 30 and 60 min sampling times during SST have an adequate
34
35 277 index of consistency, but the same is not true in terms of absolute agreement, particularly if
36
37 278 central AI is suspected. Overall, a single determination at 60 min during the SST appears to have
38
39 279 the higher Sp and PPV for the diagnosis of subjects presenting with either primary or central AI.
40
41 280 In consonance, after evaluating retrospectively 73 subjects, Zueger et al.(20) reported that
42
43 281 sampling at 30 min of the SST did not provide any additional diagnostic advantage over
44
45 282 performing a single determination at 60 min of the test. Although similar results have been also
46
47 283 reported by others (12, 13), these studies did not take into account the primary or central origin
48
49 284 of AI and did not apply sex- and assay specific cut-off values, a fact of paramount importance
50
51 285 because of the considerable influence that cortisol immunoassays exerts on the final values
52
53 286 observed after SST (16, 17).
54
55
56
57
58
59
60

1
2
3 287 Our results also indicate that SC measurement at 30 min during the SST, when using sex-
4
5 288 and assay-specific cut-off values, are enough to rule out clinically relevant primary AI since only
6
7 289 4.2% of patients in this particular situation showed a subnormal response at 30 min followed by
8
9
10 290 normal response at 60 min. Furthermore, these subjects presented with stimulated SC
11
12 291 concentrations which were very close to the cut-off point values, to the extent that the differences
13
14 292 with these normal limits may be explained by the intrinsic variability of the commercial
15
16 293 immunoassays. Even more important from a practice point of view, these subjects did not require
17
18 294 replacement therapy during their follow-up, did not suffer an acute adrenal crisis, and were not
19
20 295 diagnosed with any adrenal condition strongly suggesting that their HPA function was actually
21
22 296 normal.
23
24

25
26 297 However, in line with abovementioned studies, the 60-min sampling time appears to be
27
28 298 more specific than 30 min measurements when central AI is suspected. In such a case, 12% of
29
30 299 subjects presenting with a subnormal response at 30 min had a normal response at 60 min,
31
32 300 avoiding unnecessary treatments in them. Although a subnormal 30-min response in patients
33
34 301 with suspicion of secondary AI may not translate adrenal replacement needs in a non-critical
35
36 302 scenario, it is more than likely that many physicians feel more convenient with a stimulated
37
38 303 value over the LLN for not beginning that therapy, and 60-min sampling time is mildly better in
39
40 304 gauging that aim. Furthermore, the need of relying mostly on 60-min cortisol responses to
41
42 305 cosyntropin when a central AI is also supported by the fact that in 2 out of 5 cases showing a
43
44 306 subnormal response at 60 min but normal values at 30 min, an AI was confirmed latter in their
45
46 307 follow-up due to former pituitary radiotherapy.
47
48
49

50
51 308 Our present findings also reinforce the need of sex and assay-specific cut-off values to
52
53 309 interpret the results of the SST, in agreement with recent clinical guidelines (3). The use of such
54
55 310 cut-offs leads to a reduction in FP results, higher Sp and PPV, less discordant results among
56
57
58
59

1
2
3 311 sampling times, and fewer unnecessary treatments [20 patients (5%) could have been treated
4
5 312 unnecessarily if a classic cut-off value was applied for diagnosis]. The reliability of sex- and
6
7 313 assay-specific cut-off values was confirmed in our population of premenopausal women with
8
9 314 normal HPA axis, in whom these cut-offs were more appropriate than relying on classic values to
10
11 315 assess the functionality of their HPA axis. In this population, the LLNs for 30 min stimulated SC
12
13 316 were very close to those reported for each immunoassay by the manufacturers, yet reinforcing
14
15 317 the need to establish local normative data in order to improve the diagnostic accuracy of cortisol
16
17 318 measurements during SSTs (16).

19 319 Among the strengths of our study, we would highlight the large series of subjects
20
21 320 suspected of AI who were evaluated with a standardized dynamic study, and the systematic
22
23 321 review of subjects' clinical recordings that followed such evaluations. However, we are aware of
24
25 322 several weaknesses derived from the observational and retrospective design of the study, making
26
27 323 impossible to rule out information bias. Our best efforts might have not been enough to avoid
28
29 324 misclassification of patients according to the suspicion of primary or central AI. Also, the
30
31 325 administration of supraphysiological doses of cosyntropin does not permit ruling out partial
32
33 326 deficiencies either, particularly in those suspected of central HPA defects. Also, and even
34
35 327 considering the large sample of subjects included in our study, our present results may not be
36
37 328 extrapolable to other populations in whom SC has been measured with different immunoassays
38
39 329 that would require different local normative data. Moreover, analysis of Sp and PPV has not
40
41 330 been challenged against a biochemical gold-standard in most cases and, we have not been able to
42
43 331 establish false negative rates, sensitivity and negative predictive values. Nonetheless, besides
44
45 332 those assessments had been unethical in most cases, the lack of a laboratory gold-standard such
46
47 333 as an insulin tolerance test did not override our results, since from a practical point of view, we
48
49 334 are looking for patients needing replacement therapy and not for those with a partial AI who do

335 not require any treatment. Lastly, we could not rule out entirely pre-treatment with progestogens
336 in the context of induction of withdrawal bleeding in our confirmative population. Because these
337 drugs might exert a mild suppressive effect on the HPA axis (18, 21), their administration in a
338 few cases could have, at least in theory, lowered stimulated SC values.

339

340 CONCLUSIONS

341 Compared with the use of classic cut-off values derived from the literature, application of
342 sex- and assay-specific cut-off values of SC responses to cosyntropin results into higher Sp and
343 PPV for establishing a diagnosis of AI, thereby avoiding unnecessary treatments. Measurement
344 of stimulated SC at 30 min after SST may suffice for the correct diagnosis of primary AI, yet 60
345 min measurements might be preferable when central AI is suspected.

346

347 **Conflict of interest:** None.

348

349 **Funding:** This work has been supported by a grant Fondo de Investigación Sanitaria
350 (PI1400649) from Instituto de Salud Carlos III, Spanish Ministry of Economy and
351 Competitiveness. M.L.-R. has a local grant for clinical research from the Instituto Ramón y Cajal
352 de Investigación Sanitaria (IRYCIS). CIBERDEM is also an initiative of Instituto de Salud
353 Carlos III, partially supported by Fondo Europeo de Desarrollo Regional FEDER. There were no
354 other sources of funding.

355 **Data sharing statement:** Individual participant data that underlie the results reported in this
356 article, after deidentification, so as the study protocol would be available immediately after
357 publication to anyone who wishes to access the data to achieve aims in the approved proposal

1
2
3 358 and for individual participant data meta-analysis. Proposals should be directed to
4
5 359 andres_ortiz_f@yahoo.com or to manuel.luque@salud.madrid.org. To gain access, data
6
7 360 requestors will need to sign a data access agreement
8
9

10
11 361
12
13 362 **Contributorship statement:** A.O.-F. y M.L.-R. designed the protocol, and performed the
14
15 363 statistical analysis. A.O.-F. y E.S.-C. reviewed the clinical data using the electronic or written
16
17 364 records if necessary. A.G.-C. and L.J.-M. performed the electronic search of serum cortisol
18
19 365 samples. A.O.-F. y M.L.-R. wrote the first draft of the study. All the authors, including L.N.-C.
20
21 366 and H.F.E.-M, reviewed the manuscript before its submission and contributed to intellectual
22
23 367 content. All the authors have accepted responsibility for the entire content of the manuscript and
24
25 368 approved the final submission.
26
27
28
29

30 369

31 370

32 371

33 372

34 373

35 374

36 375

37 376

38 377

39 378

40 379

41 380

381 **REFERENCES**

- 382 1. Charmandari E, Nicolaides NC, Chrousos GP. Adrenal insufficiency. *Lancet*.
383 2014;383(9935):2152-67.
- 384 2. Neary N, Nieman L. Adrenal insufficiency: etiology, diagnosis and treatment. *Curr Opin*
385 *Endocrinol Diabetes Obes*. 2010;17(3):217-23.
- 386 3. Bornstein SR, Allolio B, Arlt W, Barthel A, Don-Wauchope A, Hammer GD, et al.
387 *Diagnosis and Treatment of Primary Adrenal Insufficiency: An Endocrine Society Clinical*
388 *Practice Guideline*. *J Clin Endocrinol Metab*. 2016;101(2):364-89.
- 389 4. Grossman AB. Clinical Review#: The diagnosis and management of central
390 hypoadrenalism. *J Clin Endocrinol Metab*. 2010;95(11):4855-63.
- 391 5. Agha A, Tomlinson JW, Clark PM, Holder G, Stewart PM. The long-term predictive
392 accuracy of the short synacthen (corticotropin) stimulation test for assessment of the
393 hypothalamic-pituitary-adrenal axis. *J Clin Endocrinol Metab*. 2006;91(1):43-7.
- 394 6. Sprung CL, Annane D, Keh D, Moreno R, Singer M, Freivogel K, et al. Hydrocortisone
395 therapy for patients with septic shock. *N Engl J Med*. 2008;358(2):111-24.
- 396 7. Annane D, Sebille V, Troche G, Raphael JC, Gajdos P, Bellissant E. A 3-level prognostic
397 classification in septic shock based on cortisol levels and cortisol response to corticotropin.
398 *Jama*. 2000;283(8):1038-45.
- 399 8. Lindholm J, Kehlet H. Re-evaluation of the clinical value of the 30 min ACTH test in
400 assessing the hypothalamic-pituitary-adrenocortical function. *Clin Endocrinol (Oxf)*.
401 1987;26(1):53-9.
- 402 9. Hurel SJ, Thompson CJ, Watson MJ, Harris MM, Baylis PH, Kendall-Taylor P. The
403 short Synacthen and insulin stress tests in the assessment of the hypothalamic-pituitary-adrenal
404 axis. *Clin Endocrinol (Oxf)*. 1996;44(2):141-6.

- 1
2
3 405 10. Gonzalbez J, Villabona C, Ramon J, Navarro MA, Gimenez O, Ricart W, et al.
4
5 406 Establishment of reference values for standard dose short synacthen test (250 microgram), low
6
7 407 dose short synacthen test (1 microgram) and insulin tolerance test for assessment of the
8
9 408 hypothalamo-pituitary-adrenal axis in normal subjects. *Clin Endocrinol (Oxf)*. 2000;53(2):199-
10
11 409 204.
- 12
13
14 410 11. Longui CA, Vottero A, Harris AG, Chrousos GP. Plasma cortisol responses after
15
16 411 intramuscular corticotropin 1-24 in healthy men. *Metabolism*. 1998;47(11):1419-22.
- 17
18
19 412 12. Mansoor S, Islam N, Siddiqui I, Jabbar A. Sixty-minute post-Synacthen serum cortisol
20
21 413 level: a reliable and cost-effective screening test for excluding adrenal insufficiency compared to
22
23 414 the conventional short Synacthen test. *Singapore Med J*. 2007;48(6):519-23.
- 24
25
26 415 13. Chitale A, Musonda P, McGregor AM, Dhatariya KK. Determining the utility of the 60
27
28 416 min cortisol measurement in the short synacthen test. *Clin Endocrinol (Oxf)*. 2013;79(1):14-9.
- 29
30
31 417 14. de Miguel Novoa P, Vela ET, Garcia NP, Rodriguez MM, Guerras IS, Martinez de
32
33 418 Salinas Santamaria Mde L, et al. Guidelines for the diagnosis and treatment of adrenal
34
35 419 insufficiency in the adult. *Endocrinol Nutr*. 2014;61 Suppl 1:1-35.
- 36
37
38 420 15. Turpeinen U, Hamalainen E. Determination of cortisol in serum, saliva and urine. *Best*
39
40 421 *Pract Res Clin Endocrinol Metab*. 2013;27(6):795-801.
- 41
42
43 422 16. El-Farhan N, Pickett A, Ducroq D, Bailey C, Mitchem K, Morgan N, et al. Method-
44
45 423 specific serum cortisol responses to the adrenocorticotrophin test: comparison of gas
46
47 424 chromatography-mass spectrometry and five automated immunoassays. *Clin Endocrinol (Oxf)*.
48
49 425 2013;78(5):673-80.
- 50
51
52 426 17. Klose M, Lange M, Rasmussen AK, Skakkebaek NE, Hilsted L, Haug E, et al. Factors
53
54 427 influencing the adrenocorticotropin test: role of contemporary cortisol assays, body composition,
55
56 428 and oral contraceptive agents. *J Clin Endocrinol Metab*. 2007;92(4):1326-33.

- 1
2
3 429 18. Kontula K, Paavonen T, Luukkainen T, Andersson LC. Binding of progestins to the
4
5 430 glucocorticoid receptor. Correlation to their glucocorticoid-like effects on in vitro functions of
6
7 431 human mononuclear leukocytes. *Biochem Pharmacol.* 1983;32(9):1511-8.
8
9
10 432 19. Templeton G. A two step approach for transforming continuous variables to normal:
11
12 433 Implications and recommendations for IS Research. *Communications of the Association for*
13
14 434 *Information Systems.* 2011;28(4):41-58.
15
16
17 435 20. Zueger T, Jordi M, Laimer M, Stettler C. Utility of 30 and 60 minute cortisol samples
18
19 436 after high-dose synthetic ACTH-1-24 injection in the diagnosis of adrenal insufficiency. *Swiss*
20
21 437 *Med Wkly.* 2014;144(23):w13987.
22
23
24 438 21. Bornstein SR. Predisposing factors for adrenal insufficiency. *N Engl J Med.*
25
26 439 2009;360(22):2328-39.
27
28
29 440
30
31 441
32
33 442
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

TABLE 1. Baseline characteristics population as a function as clinical suspicion of primary adrenal disease, secondary adrenal disease or critical patient.

	Suspicion of primary AI (n = 168)		Suspicion of Central AI (n= 226)		Critically-illpatients (n = 12)	
	Female (n = 105)	Male (n = 63)	Female (n =141)	Male (n = 85)	Female (n = 5)	Male (n = 7)
Sex						
Age (years)	53 ± 19	59 ± 15	55 ± 16	55 ± 13	53 ± 21	45 ± 23
Weight(kg)	60 ± 12	71 ± 15	72 ± 14	84 ± 17	59 ± 13	66 ± 14
BMI (kg/m ²)	24 ± 5	24 ± 5	28 ± 5	29 ± 5	23 ± 4	25 ± 4
Na(mmol/l)	138 ± 4	137 ± 5	139 ± 2	140 ± 4	141 ± 8	141 ± 9
K(mmol/l)	4.3 ± 0.6	4.4 ± 0.7	4.1 ± 0.4	4.1 ± 0.3	4.2 ± 0.7	4.2 ± 0.9
Ca (mmol/l)	2.3 ± 0.1	2.3 ± 0.2	2.4 ± 0.1	2.3 ± 0.1	2.2 ± 0.2	2.1 ± 0.2
Cr(μmol/l)	62 (44 – 1114)	80 (44 – 1158)	62 (18 – 875)	80 (44 – 150)	141 (35 – 856)	97 (53 – 283)
GFR (MDRD) (ml/min/1.73m ²)	88 (4 – 137)	85 (4 – 183)	84 (5 – 361)	93 (43 – 163)	37 (4 – 184)	78 (20 – 132)

Abbreviations, BMI, body mass index; F, female; GFR, glomerular filtration rate; M, male; sCa,

total serum calcium; sCr, serum creatinine; sK, serum potassium; sNa, serum sodium.

Data are presented as mean ± standard deviation or median (minimum-maximum) according to their distribution. To convert serum calcium to conventional system units, multiply by 4 (result in mg/dl); to convert serum creatinine to conventional system units by 0.0113 (result in mg/dl).

453 **TABLE 2.** Specificity and positive predictive value (PPV) for diagnosis of adrenal insufficiency,
 454 after short high-dose cosyntropin test, according to serum cortisol cut-off points (classic and
 455 assay-specific), and as a function of suspected level of the defect.

	Classic cut-off values						Sex- and assay-specific cut-off values					
	Global		Clinical suspicion Primary AI		Clinical suspicion Central AI		Global		Clinical suspicion Primary AI		Clinical suspicion Central AI	
Sampling time (min)	30	60	30	60	30	60	30	60	30	60	30	60
Specificity (%)	87	98	95	100	79	98	89	98	95	100	83	97
PPV (%)	69	95	74	100	67	96	67	92	75	100	61	90

457

458

459

460

461

462

463

464

465

466

467 Legend to figures

468 **Figure 1.** Panel A: Mean difference among 30 and 60 min time points after cosyntropin
469 stimulation. Data are shown as mean \pm standard deviation, and mean difference (95%CI).
470 Comparisons among time points were performed by a repeated-measure ANOVA addressing
471 main effects by a Bonferroni's confidence interval adjustment. Panel B: Pearson's correlation
472 analysis between serum cortisol values at 30 and 60 min time points. Panel C: Bland-Altman
473 plot. Solid black line represents the perfect agreement among both time points. Solid blue line is
474 the mean of differences among both time points. Dashed blue lines are 2 standard deviation of
475 the mean of differences. Solid red line is the regression line of mean differences. To convert
476 serum cortisol to metric units, multiply by 0.03625 (result in $\mu\text{g/dl}$).

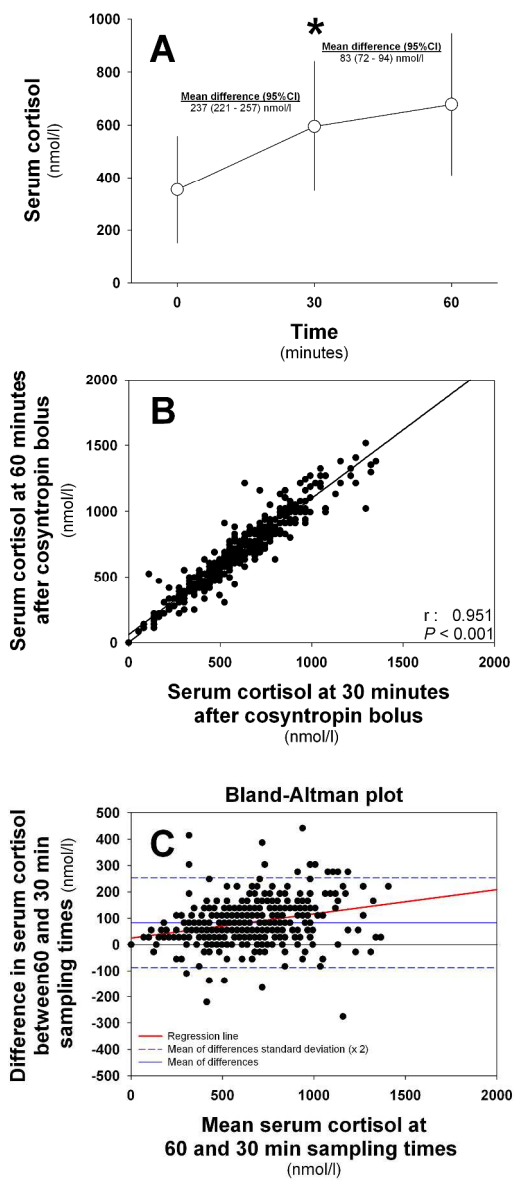
477
478 **Figure 2.** Subgroups of patients according to serum cortisol response after cosyntropin
479 stimulation as a function of classic and assay-specific cut-offs. Figures on top of bars indicate the
480 number of patients included in the subgroups. Diagnostic agreement is shown as the percentage
481 of observed agreements and kappa coefficients (95%CI).

482
483 **Figure 3.** Subgroups of patients according to serum cortisol response after cosyntropin
484 stimulation as a function of cut-off values and suspected primary or central AI. Figures on top of
485 the bars show the number of patients included in the different subgroups. Diagnostic agreement
486 is shown as the percentage of observed agreements and kappa coefficients (95%CI).

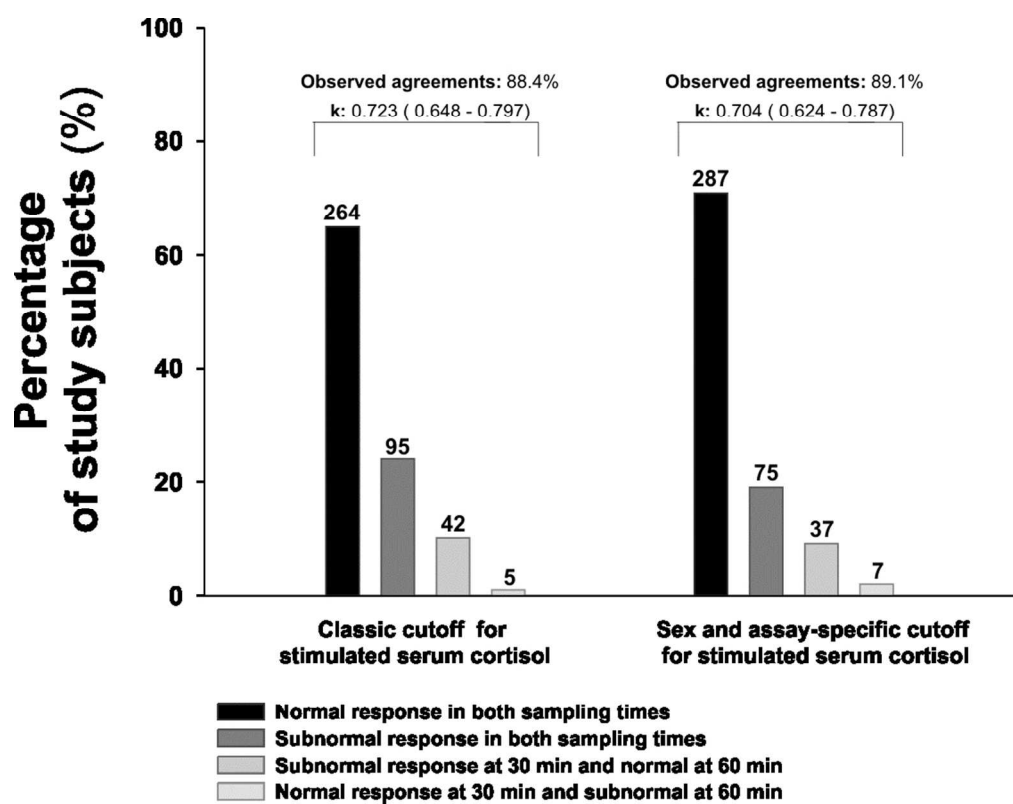
487
488 **Figure 4.** Descriptive statistics and distribution of 30 min stimulated serum cortisol
489 measurement in a population of premenopausal healthy women with evidence of normal HPA
490 axis function. The boundary of the box closest to zero indicates the 25th percentile, the solid and

1
2
3 491 long dash lines within the box marks the median and mean, respectively, and the boundary of the
4
5 492 farthest from zero indicates the 75th percentile. Whiskers above and below the box indicate the
6
7 493 90th and 10th percentiles. The dashed red line indicates the lower limit of normality (2.5th
8
9
10 494 percentile) for each immunoassay.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

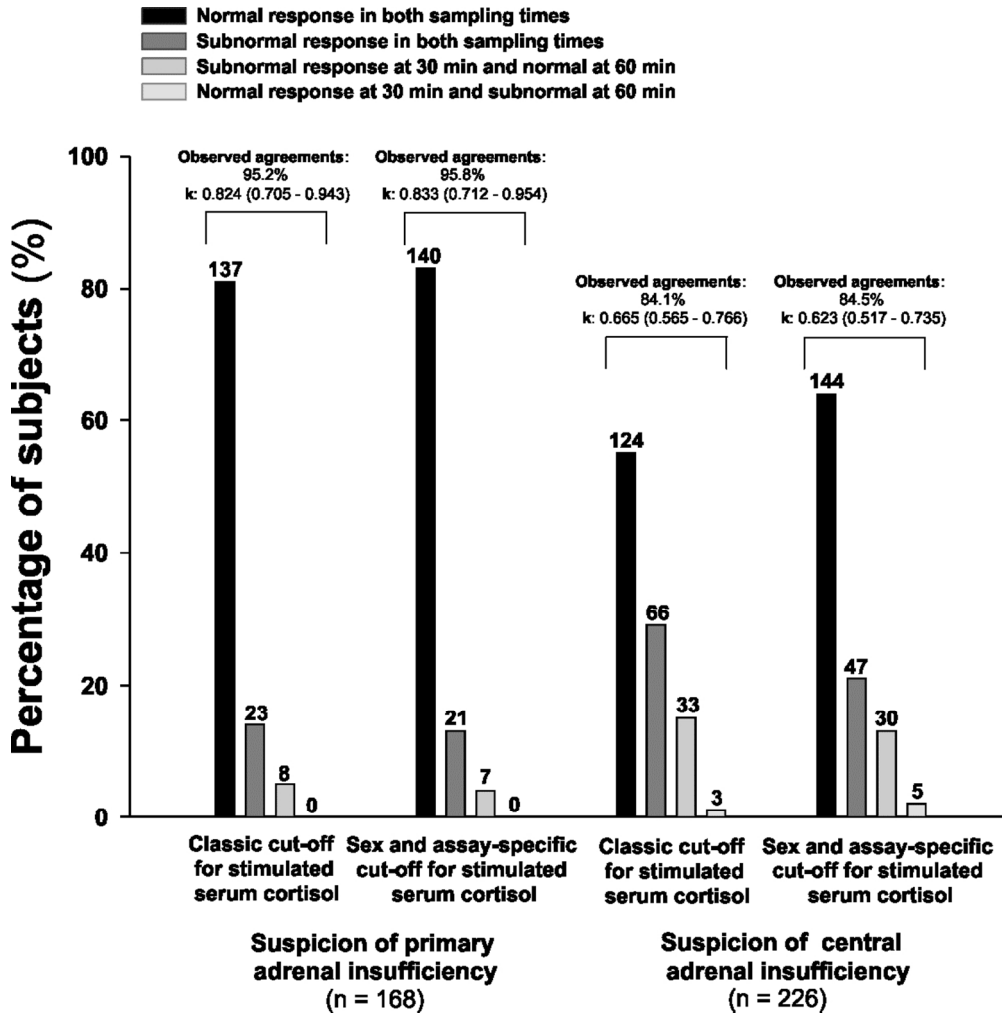


176x410mm (300 x 300 DPI)



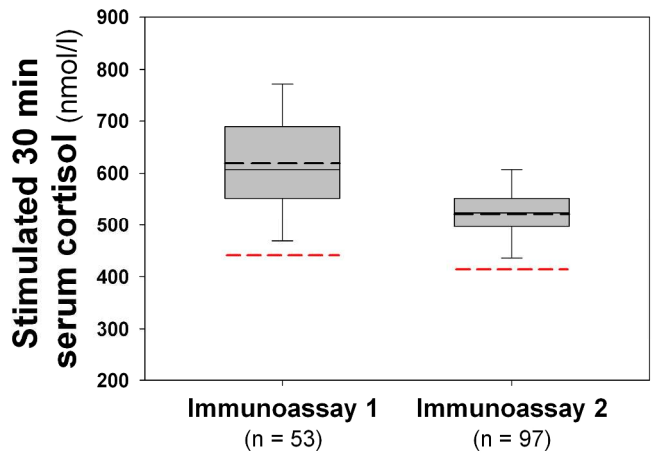
190x149mm (150 x 150 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



201x202mm (150 x 150 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



209x296mm (300 x 300 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

Title: Importance of sampling time and assay cut-offs for routine assessment of adrenal function: an observational longitudinal study.

	Item No	Recommendation
Title and abstract Page: 1	1	(a) Indicate the study's design with a commonly used term in the title or the abstract. Page 1 (b) Provide in the abstract an informative and balanced summary of what was done and what was found. Page 2
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported. Page 4
Objectives	3	State specific objectives, including any prespecified hypotheses. Page 5
Methods		
Study design	4	Present key elements of study design early in the paper. Page 5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection. Page 5
Participants	6	(<i>Cross-sectional study</i>)—Give the eligibility criteria, and the sources and methods of selection of participants. Page: 5-6
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable. Page: 5-7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Page: 5-7
Bias	9	Describe any efforts to address potential sources of bias. Page 5-7
Study size	10	Explain how the study size was arrived at. Page: 5-7
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why. Page: 5-7
Statistical methods Page: 8-9	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses

Continued on next page

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed Page: 5-6 (b) Give reasons for non-participation at each stage Page: 5-6 (c) Consider use of a flow diagram Not considered.
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Page 5 - 6 (b) Indicate number of participants with missing data for each variable of interest. Page: 5-6 <i>Cross-sectional study</i> —Report numbers of outcome events or summary measures Page: 5-6
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included Page 9-11 (b) Report category boundaries when continuous variables were categorized Page 9-11 (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period. Not relevant
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses Page 9-11
Discussion		
Key results	18	Summarise key results with reference to study objectives Page 12 - 15
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias Page 14
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence Page 14
Generalisability	21	Discuss the generalisability (external validity) of the study results Page 15
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based Page 15



PROYECTO DE INVESTIGACIÓN

EVALUACIÓN DE LA PRUEBA DE ESTIMULACIÓN CORTA CON 1-24 CORTICOTROPINA EN EL DIAGNÓSTICO DE LA INSUFICIENCIA SUPRARRENAL

Co-IPs

Manuel Luque Ramírez

Andrés Eduardo Ortiz Flores

Servicio de Endocrinología y Nutrición

Abril 2016

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

BMJ Open: first published as 10.1136/bmjopen-2017-019273 on 5 May 2018. Downloaded from <http://bmjopen.bmj.com/> on April 19, 2024 by guest. Protected by copyright.

INDICE

• Resumen (Objetivos y Metodología)	4
• Introducción.....	5
• Justificación del presente proyecto de investigación	7
• Hipótesis de trabajo	9
• Objetivos específicos	9
• MATERIAL Y MÉTODOS	11
○ Diseño del estudio y variables de análisis	11
○ Presentación de resultados	14
○ Limitaciones del estudio	15
○ Seguridad	15
○ Plan de trabajo	15
• Medios disponibles para la realización del proyecto	16
○ Equipo investigador	16
○ Medios materiales	16
○ Memoria económica	16
• Aspectos éticos	17
○ Declaración de cumplimiento, cumplimiento de requisitos éticos y regulatorios	17
○ Confidencialidad de los pacientes	17



- **Consideraciones acerca del consentimiento informado.....18**
- **Compensación a los investigadores18**
- **Difusión de los resultados18**
- **Bibliografía19**



RESUMEN (Objetivos y Metodología)

1. OBJETIVOS:

Valorar la calidad asistencial de la prueba de estimulación corta con 1-24 ACTH sintética en nuestro centro hospitalario, en el diagnóstico de pacientes con sospecha de insuficiencia suprarrenal tanto primaria como secundaria, analizando los valores de cortisol sérico en el tiempo 30 y 60 tras el inicio del estímulo, y valorar si son equivalentes o no; de igual manera, buscamos establecer el límite inferior de la normalidad con mayor sensibilidad y/o especificidad para el diagnóstico de insuficiencia suprarrenal. Finalmente buscamos conocer el valor predictivo de esta prueba en nuestro centro hospitalario

2. METODOLOGÍA

Es un estudio de diseño retrospectivo, observacional, en el cual se recolectarán aquellos datos correspondientes a las determinaciones de cortisol plasmático no estimulado, y en los tiempos 30 y 60 tras el inicio del estímulo con 250 μg de 1-24 ACTH sintética, prueba indicada tanto para el screening como para el diagnóstico de insuficiencia suprarrenal primaria y central. Para esto se realizará una revisión de aquellos datos que constan en el registro electrónico del Servicio de Análisis Clínicos-Sección de Hormonas (años 2011 a 2015) (n: 536). Luego de esto se obtendrán aquellos datos clínicos y bioquímicos relevantes, mediante el análisis de la historia clínica electrónica, o del archivo en papel en caso de ser necesario. En segundo lugar, buscaremos determinar el punto de corte tras estímulo con mayor sensibilidad y/o especificidad para el diagnóstico de insuficiencia suprarrenal, y finalmente buscaremos determinar la existencia o no de concordancia entre los tiempos 30 y 60 tras estímulo con 1-24 ACTH, buscando de esta manera valorar la calidad asistencial de este test en el diagnóstica de esta entidad clínica.



INTRODUCCIÓN

La insuficiencia suprarrenal (IS) es un trastorno que refleja una síntesis deficiente de glucocorticoides, ya sea por un defecto primario a nivel de las glándulas suprarrenales, o por una secreción inadecuada de corticotropina (ACTH) por la hipófisis que produce una atrofia suprarrenal secundaria (1). Presenta una elevada mortalidad, por lo que es importante, ante la presencia de clínica compatible con la misma, realizar un adecuado diagnóstico, el cual debe ser confirmado mediante las pruebas de laboratorio correspondientes. De ahí la necesidad de disponer de métodos de alta sensibilidad y especificidad que permitan establecer un diagnóstico y tratamiento precoz.

El diagnóstico de IS se basa en la demostración de una producción de cortisol inadecuadamente baja (1). La determinación de cortisol sérico no estimulado a las 8:00 AM permite establecer una aproximación diagnóstica, dado que cifras por debajo de 5 $\mu\text{g/dl}$ sugieren una alta probabilidad de IS, mientras que valores por encima de 18 - 20 $\mu\text{g/dl}$ predicen una respuesta normal en las pruebas de estímulo (hipoglucemia insulínica o prueba de estimulación corta con ACTH sintética). Valores entre 5 a 18 $\mu\text{g/dl}$ se consideran indeterminados y precisan de pruebas que confirmen el diagnóstico (1-2).

La prueba de estimulación corta (EC), se realiza mediante la determinación de los niveles de cortisol sérico a los 30 minutos y/o 60 minutos de la administración de 250 μg de 1-24 ACTH sintética. Cualquier valor por encima de 18 - 20 $\mu\text{g/dl}$ tras la prueba de estímulo, es considerado como una respuesta normal (2). Es la prueba de elección para el diagnóstico de IS primaria (1,2).



Por otra parte, esta prueba también se emplea en el diagnóstico de IS secundaria a una producción endógena insuficiente de ACTH, que conducirá a atrofia suprarrenal y, por lo tanto, a una respuesta pobre en la producción de cortisol (3,4). Sin embargo, no se recomienda su utilización en los primeros momentos tras el insulto hipotálamo-hipofisario, ya que requiere tiempo (4 a 6 semanas) para que se produzca una alteración en la respuesta al estimular la reserva suprarrenal (3, 4).

Debido a que algunos pacientes con insuficiencia suprarrenal central parcial presentan una respuesta normal tras el EC con 250 μg de 1-24 ACTH, pero anormal tras otras pruebas de estímulo, algunos autores han propuesto la realización de la EC con 1 μg (5). No obstante, un porcentaje significativo de individuos sanos presentan respuestas “patológicas” con este estímulo (6), lo que sugiere la existencia de problemas técnicos que influyen a la hora de realizar el EC con 1 μg (6), como una administración incompleta de la dosis bien por errores en la dilución del producto, o bien por adherencia de la ACTH a las paredes del catéter (6). Por este motivo, y dado el mayor grado de evidencia disponible, actualmente también se recomienda la utilización del EC con 250 μg de 1-24 ACTH para el diagnóstico de IS secundaria (1-4).

Finalmente, en las unidades de cuidados intensivos, tanto pediátrica como de adultos, el EC con 1-24 ACTH se utiliza, generalmente en pacientes con sepsis grave, que presentan hipotensión sostenida con falta de respuesta al tratamiento vasopresor o inotrópico, sospechando por lo tanto en un déficit relativo de glucocorticoides, entidad conocida como insuficiencia suprarrenal relativa o del paciente crítico (2,8). No existe una clara definición clínica o bioquímica de IS relativa, por lo que hasta fecha actual se considera como diagnóstico un incremento del pico de cortisol en el tiempo 30 o 60 inferior a 9 $\mu\text{g}/\text{dl}$ respecto a su valor basal (7,8).



JUSTIFICACIÓN DEL PRESENTE PROYECTO DE INVESTIGACIÓN

Actualmente, existen pocos estudios que avalen la necesidad de realizar la determinación de cortisol simultáneamente en los tiempos 30 y 60 min. tras el estímulo con 1-24 ACTH, y aunque ciertos autores sugieren innecesaria la determinación del tiempo 60 (3, 9-10), no existe evidencia suficiente que lo justifique, puesto que este momento de extracción se ha demostrado como más apropiado en otros estudios (11-12). No obstante, las Guías Clínicas más recientemente publicadas sugieren la determinación del cortisol sérico a los 30 minutos de la estimulación (13-14), aunque reconocen que evaluar la posibilidad de que la determinación a los 60 minutos sea más específica requiere ser investigado (14). Estas controversias podrían estar relacionadas con los métodos de determinación de cortisol, vía de administración del estímulo (acceso intravenoso o intramuscular), y la evaluación conjunta de pacientes con déficit primario o central, en los que el momento más adecuado para la determinación del cortisol sérico tras estimulación puede ser también diferente. Pese a estos datos, o como consecuencia de los mismos, durante los últimos años en nuestro Servicio se realizan ambas determinaciones (tiempos 30 y 60 min tras estímulo).

Junto con el momento más adecuado para la extracción de cortisol sérico, otro aspecto importante es la determinación del punto de corte de mayor sensibilidad y especificidad a la hora de establecer el diagnóstico de insuficiencia suprarrenal, dada las importantes implicaciones clínicas de este diagnóstico, y que debería ensayo-específico, puesto que el empleo de diferentes anticuerpos de detección influye en el límite de normalidad tras el estímulo con 1-24 ACTH (14-15). Aunque como ya se ha comentado, clásicamente se considera un valor superior a 18-20 $\mu\text{g}/\text{dl}$ como “normal”, existen pocos datos en población de nuestro entorno que evalúen el valor predictivo positivo de este punto de corte (10), y en nuestro centro en los últimos años se ha cambiado el ensayo para la determinación de cortisol, lo que ya ha demostrado variar el límite de normalidad en estudios previos (15).



1
2
3
4
5
6
7
8
9
10 Por lo tanto, con el objetivo de evaluar y mejorar la calidad asistencial en nuestro
11 Servicio en el diagnóstico de la insuficiencia suprarrenal mediante la estimulación con 1-24
12 ACTH, parece pertinente la revisión sistemática de nuestros resultados que permitiría determinar
13 el momento para la determinación de cortisol y punto de corte más adecuado, mejorandola
14 eficiencia de esta prueba diagnóstica. A modo de ejemplo, la ausencia de diferencias entre ambos
15 tiempos de evaluación podría disminuir el tiempo y número de determinaciones con la
16 consiguiente reducción de costes económicos, mejorar la calidad de atención del paciente y
17 evitar posibles tratamientos innecesarios.
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



HIPÓTESIS DE TRABAJO

- 1) La determinación de cortisol plasmático a los 30 minutos del estímulo con 1-24 ACTH es lo suficientemente sensible para detectar a aquellos pacientes con insuficiente función corticosuprarrenal, obviando la necesidad de la determinación a los 60 minutos.
- 2) Algunos pacientes con insuficiencia suprarrenal central podrían presentar una respuesta tardía al estímulo con 1-24 ACTH que precisaría la determinación a los 60 minutos tras estímulo.
- 3) El límite inferior de la normalidad utilizado para el diagnóstico de insuficiencia suprarrenal tras el estímulo con 1-24 ACTH en nuestro centro (18-20 mcg/dl), no es adecuado con el inmunoensayo utilizado actualmente, sobreestimando el diagnóstico de insuficiencia suprarrenal.

OBJETIVOS ESPECÍFICOS

3. Valorar la calidad asistencial del EC con 1-24 ACTH en nuestro centro hospitalario en el diagnóstico de pacientes con sospecha de IS tanto primaria como secundaria.
4. Establecer una correlación entre los hallazgos de laboratorio obtenidos, con la sospecha clínica que motivó el estudio en estos pacientes, valorando así la eficacia y la calidad asistencial.
5. Conocer el valor predictivo positivo de esta prueba en nuestro centro hospitalario, analizando los resultados patológicos tras la prueba de estímulo, la existencia o no de IS y la necesidad de tratamiento sustitutivo con glucocorticoides.



6. Analizar los valores de cortisol sérico obtenidos en el tiempo 30 y 60 tras el inicio del estímulo, compararlos y determinar si ambos resultados son equivalentes o no, para poder justificar la necesidad de realizar o no ambos tiempos.
7. Establecer el límite inferior de la normalidad con mayor sensibilidad y/o especificidad para el diagnóstico de insuficiencia suprarrenal en nuestro medio.



MATERIAL Y MÉTODOS

Diseño del estudio y variables de análisis

Se trata de un estudio observacional retrospectivo en el que se recogerán los datos correspondientes a las determinaciones de cortisol plasmático de aquellos pacientes con determinación basal, y en los tiempos 30 y 60 minutos tras estímulo con 1-24 ACTH solicitados para el despistaje-diagnóstico de insuficiencia suprarrenal en nuestro centro, y disponibles en el registro electrónico del Servicio de Análisis Clínicos-Sección de Hormonas (años 2011 a 2015) (n: 536). Junto con los valores de cortisol plasmático en los 3 tiempos, los IPs o investigadores colaboradores accederán a la historia clínica electrónica, y si es preciso al archivo en papel con el objetivo de obtener un mínimo conjunto de datos que permita la consecución de los objetivos del estudio:

- Edad en el momento de la determinación.
- Sexo.
- Peso y Talla.
- Na, K y Calcio plasmáticos en el momento del estímulo.
- Cr y TFGe.
- Concentraciones de ACTH basal si están disponibles.
- Motivo de la solicitud. Sospecha de insuficiencia suprarrenal 1^a / Central / no determinada.
- Antecedentes de patología hipotálamo-hipofisaria. Si existen tiempo de evolución desde el insulto hipotálamo-hipofisario y estudio de función suprarrenal.
- Antecedentes de patología autoinmune poliglandular: tiroiditis, diabetes mellitus tipo 1, etc...
- Antecedentes o presencia de tratamiento estrogénico o gestación.



- Antecedentes o presencia de tratamiento con mitotane, ketoconazol, metopirona, fenitoína o carbamacepina.
- Antecedentes de tratamiento glucocorticoideo previo al estudio. Dosis, preparado, tiempo de administración, tiempo de evolución desde su suspensión.
- Realización de otros estudios del eje hipotálamo-hipófisis-suprarrenal: hipoglucemia insulínica, prueba de estímulo con glucagón, prueba de estímulo con CRH, prueba de supresión con metopirona y resultado de la misma.
- Tipo de inmunoensayo: Immulite ® / Architect ®.
- Necesidad de tratamiento glucocorticoideo ± mineralocorticoideo por déficit de producción posterior al estudio con la prueba de estimulación.
- Presencia de cuadros clínicos compatibles con crisis suprarrenal aguda.
- Diagnóstico final de funcionalidad del eje hipotálamo-hipófisis-suprarrenal.

Para determinar el punto de corte tras estímulo con mayor sensibilidad y/o especificidad para el diagnóstico de insuficiencia suprarrenal se definirá como pacientes con eje hipotálamo-hipófisis-suprarrenal normal desde el punto de vista clínico a aquellos pacientes que cumplan todos los siguientes criterios:

- i) Ausencia de necesidad de administración glucocorticoideo ± mineralocorticoideo transitorio o crónico con la indicación de tratamiento sustitutivo de la función corticosuprarrenal.
- ii) Ausencia de cuadros clínicos compatibles con crisis suprarrenal aguda.



- iii) Ausencia de hiperpigmentación cutánea de etiología no filiada.
- iv) Ausencia de hiponatremia o hiperpotasemia de origen no filiado.
- v) Ausencia de hipoglucemia en pacientes sin diabetes mellitus de origen no filiado.
- vi) Ausencia de pérdida de peso no justificada de origen no filiado.
- vii) Ausencia de cuadros compatibles con hipotensión ortostática de origen no filiado.
- viii) Ausencia de elevación de ACTH 2 veces por encima del límite superior de la normalidad (> 300 pg/ml).
- ix) Ausencia de demostración de insuficiencia suprarrenal parcial o completa en alguna otra prueba de valoración del eje hipotálamo-hipófisis-suprarrenal.

Para establecer la concordancia entre la determinación de las concentraciones de cortisol plasmático entre los tiempos 30 y 60 minutos se utilizará como límite inferior de la normalidad:

- i) Un valor de cortisol plasmático tras estímulo con 1-24 ACTH ≥ 18 $\mu\text{g/dl}$ (punto de corte clásico).
- ii) Un valor de cortisol plasmático tras estímulo con 1-24 ACTH específico para inmunoensayo, sexo y si está presente, ingesta de ACO combinados (15).



Presentación y análisis de resultados

Los datos se presentan en forma de media \pm desviación típica o IC95%, mediana (rango intercuatílico) (valor mínimo-máximo), número crudos (porcentaje), de acuerdo a la naturaleza y distribución de los mismos, que se analizará mediante la prueba de Kolmogorov-Smirnov para una muestra. Las comparaciones entre variables continuas se realizarán mediante la prueba U de Mann-Whitney o t de Student según la distribución de las mismas. Las comparaciones entre variables cualitativas se realizarán mediante la prueba exacta de Fisher o χ^2 según sea necesario. La correlación entre los tiempos de determinación 30 y 60 minutos se analizará mediante análisis de Pearson, aplicando transformación logarítmica de las variables si es necesario para asegurar la normalidad de las mismas. La concordancia entre los tiempos de determinación de cortisol plasmático se analizará mediante el coeficiente de correlación intraclase y los métodos de representación gráfica de Bland-Altman. La concordancia entre el diagnóstico bioquímico de funcionalidad suprarrenal normal tras estímulo en los tiempos 30 y 60 minutos se analizará mediante la determinación del coeficiente κ . Los puntos de corte tras estimulación con mayor sensibilidad y especificidad para el diagnóstico de insuficiencia suprarrenal se evaluarán mediante curvas ROC. Se considerará estadísticamente significativo un valor $p < 0,05$. Los análisis se realizarán con el paquete estadístico SPSS 15.0.



Limitaciones del estudio

Las limitaciones del estudio provienen fundamentalmente de su diseño observacional y retrospectivo, con la consiguiente pérdida de datos no recogidos en la historia clínica. Para compensar por estas potenciales pérdidas, los análisis se realizarán tanto con los valores disponibles como mediante análisis de imputación múltiple. Por otro lado, otra de las limitaciones del estudio es el establecimiento del patrón oro con el que comparar los resultados de cortisol plasmático tras estimulación en términos de integridad del eje hipotálamo-hipófisis-suprarrenal. No obstante, como se han especificado previamente hemos establecido una serie de criterios clínicos y bioquímicos, que garantizan con un grado muy elevado de seguridad la ausencia de una insuficiencia suprarrenal completa, y razonablemente, una insuficiencia suprarrenal parcial clínicamente significativa.

Seguridad

El estudio no supone ningún riesgo para los pacientes incluidos en el mismo, dado que no implica ninguna intervención dado su carácter observacional y retrospectivo.

Plan de Trabajo

La recogida de datos del Servicio de Análisis Clínicos y el conjunto mínimo de datos de la historia clínica detallado previamente, se realizará por el Equipo Investigador desde el 1 de abril de 2016 al 30 de junio de 2016. El análisis de los mismos se realizará del 1 de julio de 2016 al 31 de septiembre de 2016. La comunicación de resultados del 1 de octubre al 31 de diciembre de 2016.



Medios disponibles para la realización del proyecto

Equipo Investigador

El Equipo Investigador está compuesto como co-investigadores principales de Manuel Luque Ramírez, Facultativo Especialista de Área y Andrés Eduardo Ortiz Flores, M.I.R. de 3º año del Servicio de Endocrinología y Nutrición.

Como investigadores colaboradores participarán en la adquisición, análisis de datos y reporte de resultados, Elisa Santacruz Cerdá, M.I.R. de 3º del Servicio de Endocrinología y Nutrición, Ana García Cano y Lucía Jiménez Mendiguchia, Facultativos Especialistas de Área del Servicio de Análisis Clínicos.

Medios Materiales

El presente estudio se realizará con el equipo informático y software disponible en los Servicios de Análisis Clínicos y Endocrinología y Nutrición, no precisando la adquisición de ningún equipo ni material fungible.

Memoria Económica

El presente estudio no tiene presupuestado ningún gasto ni va a incurrir en ningún coste extraordinario para el centro, puesto que los investigadores implicados realizarán el trabajo de revisión de historias clínicas fuera de su horario habitual, y el software informático preciso para la realización del mismo está ya disponible, como se ha detallado previamente.



Aspectos Éticos

Declaración de cumplimiento, cumplimiento de los requisitos éticos y regulatorios.

Este estudio se realizará con arreglo al protocolo, a los principios establecidos en la Declaración de Helsinki, a las directrices de buenas prácticas clínicas (BPC) del Comité Internacional de Armonización y a la ley de Investigación Biomédica (14/2007, de 3 de julio).

El protocolo del estudio y los documentos que demuestran la cualificación del investigador son remitidos al Comité Ético de Investigación Clínica para su revisión ética y su aprobación con arreglo a las normativas locales, antes del inicio del estudio.

Las modificaciones de la realización del estudio o de los análisis previstos se documentarán en una enmienda de protocolo y/o del plan de análisis estadístico.

Confidencialidad de los pacientes

Los investigadores preservarán la confidencialidad de todos los pacientes que participen en el estudio, con arreglo a las BPC, Declaración de Helsinki y a la legislación local (Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal).

La información obtenida y difundida por la puesta en marcha del presente estudio es considerada confidencial y deberá ser tratada en todo momento como tal. Las pacientes del estudio se identificarán con un código numérico tanto en el cuaderno de recogida de datos (CRD) en una base informatizada. Sólo aquellos datos de la historia clínica que estén relacionados con el estudio, variables especificadas previamente, serán objeto de comprobación. Esta comprobación se hará en presencia del Investigador Principal / Investigadores Colaboradores, responsables de garantizar la confidencialidad de todos los datos de las historias clínicas pertenecientes a las pacientes participantes en el estudio. Los datos recogidos para el estudio estarán identificados mediante un código y solo el investigador principal / colaboradores podrán relacionar dichos datos con el paciente y con su historia clínica.



El tratamiento de los datos se hará con las medidas de seguridad establecidas en cumplimiento de la Ley Orgánica 15/1999 de Protección de Datos de Carácter Personal. Tanto los investigadores responsables del ensayo clínico, como un representante de las Autoridades Sanitarias y el CEIC del Hospital Universitario Ramón y Cajal tendrán acceso a la información registrada a lo largo del estudio. En la publicación de los resultados del estudio no se revelará la identidad de los participantes.

Consideraciones acerca del consentimiento informado (CI).

El presente proyecto de investigación no contempla en ningún caso entrevistar a los sujetos a los que se practicó la prueba de estimulación con 1-24 ACTH, además de adoptar un método de disociación seguro para evitar manejar datos personales durante el análisis de los resultados. Su carácter retrospectivo, en pacientes que en muchos casos no tienen seguimiento posterior en el centro, y su objetivo de evaluación de calidad asistencial, hace que en consideración de los investigadores del mismo no sea precisa la solicitud de consentimiento informado específico, más allá del consentimiento genérico que otorga el paciente por su atención en un Hospital Universitario del Servicio Madrileño de Salud. No obstante, este punto del protocolo debe ser valorado favorablemente por el CEIC del Centro.

Compensación a los investigadores

Se trata de un estudio de Promoción Interna, que no cuenta con ninguna subvención a cargo de fondos públicos, y en el que no se contempla compensación económica a los investigadores participantes en el estudio.

Difusión de resultados

Los investigadores se comprometen a publicar los resultados derivados del presente estudio, independientemente de los resultados y conclusiones del mismo, respetando siempre la confidencialidad de la identidad de los sujetos participantes.



BIBLIOGRAFIA

1. Charmandari E *et al.* **Adrenal insufficiency.** *Lancet* **2015**; 383: 2152-67.
2. Neary N & Nieman L. **Adrenal Insufficiency- etiology, diagnosis and treatment** *Curr Opin Endocrinol Diabetes Obes* **2010**; 17: 217–223
3. Grossmann A. **The diagnosis and management of central hypoadrenalism.** *J Clin Endocrinol Metab* **2010**; 95: 4855–63.
4. Agha A *et al.* **The long-term predictive accuracy of the short Synacthen (corticotropin) stimulation test for assessment of the hypothalamic-pituitary-adrenal axis.** *J Clin Endocrinol Metab* **2006**; 91: 43–7.
5. Kazlauskaitė R *et al.* **Corticotropin tests for hypothalamic-pituitary-adrenal insufficiency: a metaanalysis.** *J Clin Endocrinol Metab* **2008**; 93: 4245–53.
6. Wade M *et al.* **Technical details influence the diagnostic accuracy of the 1 microg ACTH stimulation test.** *Eur J Endocrinol* **2010**; 162: 109–13.
7. Loriaux DL & Fleseriu M. **Relative adrenal insufficiency.** *Curr Opin Endocrinol Diabetes Obes* **2009**; 16: 392–400.
8. Sprung CL *et al.* **Hydrocortisone therapy for patients with septic shock.** *N Engl J Med* **2008**; 358: 111–24.
9. Hurel SJ *et al.* **The short Synacthen and insulin stress tests in the assessment of the hypothalamic-pituitary-adrenal axis.** *Clin Endocrinol (Oxf)* **1996**; 44: 141-6.
10. Gonzálbez J *et al.* **Establishment of reference values for standard dose short synacthen test (250 microgram), low dose short synacthen test (1 microgram) and insulin tolerance test for assessment of the hypothalamo-pituitary-adrenal axis in normal subjects.** *Clin Endocrinol (Oxf)* **2000**; 53: 199-204.
11. Longui CA *et al.* **Plasma cortisol responses after intramuscular corticotrophin 1-24 in healthy men.** *Metabolism* **1998**; 47: 1419-22.



12. Mansoor S *et al.* **Sixty-minute post-Synacthen serum cortisol level: a reliable and cost-effective screening test for excluding adrenal insufficiency compared to the conventional short Synacthen test.** *Singapore Med J* **2007**; 48: 519-23.
13. de Miguel Pet *al.* **Guía para el diagnóstico y tratamiento de la insuficiencia suprarrenal en el adulto.** *Endocrinol Nutr* **2014**; 61: 1-35.
14. Borstein SR *et al.* **Diagnosis and treatment of primary adrenal insufficiency: an Endocrine Society Clinical Practice Guideline.** *J Clin Endocrinol Metab* **2016**; 101: 364-89.
15. El-Farhan N *et al.* **Method-specific serum cortisol responses to the adrenocorticotrophin test: comparison of gas chromatography-mass spectrometry and five automated immunoassays.** *Clin Endocrinol (Oxf)* **2013**; 78: 673-80.

BMJ Open

Role of sampling times and serum cortisol cut-off concentrations on the routine assessment of adrenal function using the standard cosyntropin test in an academic hospital: a cross-sectional study.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-019273.R1
Article Type:	Research
Date Submitted by the Author:	15-Dec-2017
Complete List of Authors:	Ortiz-Flores, Andrés; Hospital Universitario Ramón y Cajal, Endocrinology Santacruz, Elisa; Hospital Universitario Ramon y Cajal Jiménez-Mendiguchia, Lucía; Hospital Universitario Ramon y Cajal García-Cano, Ana; Hospital Universitario Ramon y Cajal Nattero-Chávez, Lía; Hospital Universitario Ramon y Cajal Escobar-Morreale, Héctor; Hospital Universitario Ramon y Cajal Luque-Ramírez, Manuel; Hospital Universitario Ramon y Cajal
Primary Subject Heading:	Diabetes and endocrinology
Secondary Subject Heading:	Diagnostics, Patient-centred medicine
Keywords:	Adrenal disorders < DIABETES & ENDOCRINOLOGY, cortisol, biochemical diagnosis, specificity, immunoassay

SCHOLARONE™
Manuscripts

Title Role of sampling times and serum cortisol cut-off concentrations on the routine assessment of adrenal function using the standard cosyntropin test in an academic hospital: a cross-sectional study.

Short title Cosyntropin test and adrenal insufficiency

Authors Andrés E. Ortiz-Flores, M.D.,¹ Elisa Santacruz-Cerdá, M.D.,¹ Lucía Jiménez-Mendiguchia, M.D.,² Ana García-Cano, M.D.,² Lia Nattero-Chávez, M.D.,¹ Héctor F. Escobar-Morreale, M.D., Ph.D.,^{1,3} Manuel Luque-Ramírez, M.D., Ph.D.^{1,3}

Affiliations

¹ Department of Endocrinology and Nutrition, Hospital Universitario Ramón y Cajal, Madrid (Spain).

² Department of Clinical Biochemistry, Hospital Universitario Ramón y Cajal, Madrid (Spain).

³ Diabetes, Obesity and Human Reproduction Research Group, Hospital Universitario Ramón y Cajal & Universidad de Alcalá & Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS) & Centro de Investigación Biomédica en Red Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), Madrid, Spain.

Word count: 5094; **Word count for abstract:** 300; **Figures:** 5; **Tables:** 3; **Number of references:** 26.

CORRESPONDENCE & REPRINTS

1
2
3 24 Andres E. Ortiz-Flores, MD. Department of Endocrinology and Nutrition, Hospital Universitario
4
5 25 Ramón y Cajal, Carretera de Colmenar Km 9.1, E-28034 Madrid, Spain. Phone/Fax: +34-91 336
6
7 26 9029.
8
9
10
11 27 E-mail: andres_ortiz_f@yahoo.com
12

13
14 28
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

29 ABSTRACT

30 **Objectives** With the final aim of validating the use of a single post-stimulus sampling protocol
31 for the cosyntropin test (SST) in our institution, our primary objectives were: i) to determine the
32 concordance between 30 and 60 min serum cortisol measurements (SC) during the SST; ii) to
33 evaluate the diagnostic agreement between both sampling times when using classic or assay- and
34 sex-specific cut-offs values for SC. Secondary objectives included :i) estimating the specificity
35 and positive predictive value of 30 and 60 min sampling times while considering the suspected
36 origin of adrenal insufficiency; iv) to obtain assay-specific cut-offs for SC after SST in a group
37 of subjects with a normal hypothalamic-pituitary-adrenal (HPA) axis.

38 **Design and setting** Cross-sectional study conducted in an Academic Hospital between 2011 and
39 2015.

40 **Participants and interventions** Two groups were evaluated: i) a main study group including
41 370 patients in whom SC was measured at 30 and 60 minutes during the SST; and ii) a
42 confirmative group that included 150 women presenting with a normal HPA axis in whom the
43 SST was conducted to rule out late onset congenital adrenal hyperplasia. Diagnostic agreement
44 between both sampling times was assessed by considering both classic (500 nmol/l) and sex- and
45 assay-specific SC cut-off concentrations.

46 **Results** Diagnostic agreement between both sampling times was greater when applying sex- and
47 assay-specific cut-off values instead of classic cut-offs. For suspected primary adrenal
48 insufficiency, SC measured at 30-min was enough to make a diagnosis in over 95% of cases,
49 without missing any necessary treatment. When the suspicion was central adrenal insufficiency,
50 the 60 min SC measurement was more specific, establishing diagnosis in over 97% of cases.

1
2
3 51 **Conclusions:** The use of sex- and assay-specific SC cut-off values improve the diagnostic
4
5 52 accuracy of SST for the evaluation of suspected adrenal insufficiency. For primary disease, a
6
7 53 subnormal SC response at 30 min is a reliable marker of adrenal dysfunction. On the contrary,
8
9 54 when central adrenal insufficiency is suspected, the 60 min SC measurement improves the
10
11 55 diagnostic accuracy of the test.
12
13
14
15 56

16
17 57 **Strengths and limitations:**

- 18
19 58 • We assessed a very large series of well-characterized subjects with a suspicion of adrenal
20
21 59 insufficiency and a minimum clinical follow up of 12 months after the cosyntropin test.
22
23
24 60 • We used a pre-test distinction between primary and central adrenal insufficiency based on
25
26 61 clinical data.
27
28 62 • We used a local cohort of women with definitely normal cortisol secretion to validate our
29
30 63 findings.
31
32
33 64 • Our results were not challenged against a biochemical gold-standard and, therefore, false
34
35 65 negative rates, sensitivity, and negative predictive values were not established.
36
37
38 66 • The confirmatory group was comprised only by premenopausal women, and cosyntropin-
39
40 67 stimulated SC concentrations were only obtained at the 30 min sampling time in these
41
42 68 subjects.
43
44
45 69

46
47 70 **Keywords:** Adrenal insufficiency; biochemical diagnosis; cosyntropin test; immunoassay;
48
49 71 reference values; sampling times; serum cortisol; specificity.
50
51
52 72

73 Introduction

74
75 The laboratory diagnosis of adrenal insufficiency (AI) at the clinical setting relies on the
76 finding of an inappropriately low morning circulating serum cortisol (SC) or subnormal SC
77 responses to adrenal stimulation [1]. However, the diagnosis of AI diagnosis should not be made
78 according only to laboratory tests, since analytical results must always be interpreted in the
79 context of the whole clinical picture of the individual patient [1–3]. The most widely used
80 adrenal stimulation protocol consists of measuring SC in samples obtained 30 and 60 min after a
81 single 250 µg intravenous bolus or intramuscular injection of tetracosactide (cosyntropin). The
82 normal response consists of a SC value ≥ 500 nmol/l (18 µg/dl) 30 at any time after cosyntropin
83 administration. This protocol, also known as a short standard high-dose test (SST), is the
84 dynamic exploration of choice for primary AI diagnosis [1,3] and it is also used for non-acute
85 central AI [4,5]. In critically ill patients, SST may be performed to rule out a functional form of
86 AI –critical illness-related corticosteroid insufficiency– in subjects showing sustained refractory
87 hypotension and no response to vasopressor drugs [2,6]. Clinical guidelines suggest that this
88 condition may be best diagnosed by a random SC below 276 nmol/l (10 µg/dl) or when the
89 increase in SC after cosyntropin is less than 248 nmol/l (9 µg/dl) [7,8].

90 The issue of which sampling time - 30 min or 60 min – of the SST is the most appropriate
91 is controversial. The 30 min SC measurements have been validated against a “gold standard”
92 such as the insulin tolerance test (ITT) [9]. Hence, some authors [4,10,11] suggest that a single
93 SC measurement 30 min after cosyntropin administration is enough to establish or rule out
94 clinically relevant AI. Other studies show that a 60 min sample may avoid unnecessary
95 overdiagnosis [12–14]. Recent clinical practice guidelines recommend further research to clarify
96 whether 60 min SC might be more specific than 30 min measurements for AI diagnosis [3,15].

1
2
3 97 Even though liquid chromatography/mass spectrometry techniques are currently
4
5 98 recommended for the accurate measurement of circulating steroids, in most centres clinical
6
7
8 99 routine still relies on automated immunoassays for SC [16]. Considering that the classic cut-off
9
10 100 value for the SST was established for SC as measured by older radioimmunoassays, and that
11
12 101 immunochemiluminescent assays differ in antibody specificity with these earlier assays [17],
13
14 102 establishing local assay-specific cut-off values is of paramount importance to properly classify
15
16 103 SC responses to cosyntropin [3,17,18]. This issue is not inconsequential because, despite the
17
18 104 recommendation of using local assay-specific lower limits of normality (LLN) for the dynamic
19
20 105 assessment of the hypothalamic-pituitary-adrenal (HPA) axis [3], in our experience many
21
22 106 physicians still apply classic cut-off values in their routine practice. Also, other factors that may
23
24 107 influence SC measurement include the stimulation of hepatic synthesis and secretion of cortisol
25
26 108 binding globulin by oestrogens, sex and several non-glucocorticoid drugs [18,19].

27
28
29
30 109 To provide new insights into these still open questions, and while validating the use of a
31
32 110 single post-stimulus sampling protocol for the routine cosyntropin test (SST) in our institution,
33
34 111 our primary goals were: i) to assess the concordance between 30 and 60 min SC concentrations
35
36 112 after cosyntropin stimulation at the clinical setting; ii) to estimate the diagnostic agreement
37
38 113 between both sampling times when using classic cut-offs derived from the literature or assay-
39
40 114 and sex-specific cut-offs values, taking into account the suspected origin of AI. As secondary
41
42 115 objectives, we aimed to :i) estimate the specificity (Sp) and positive predictive value (PPV) of 30
43
44 116 and 60 min sampling times while taking into account the origin of AI; and ii) confirm assay-
45
46 117 specific LLN for SC concentration after cosyntropin in a group of subjects with a normal HPA
47
48 118 function.

49
50
51
52
53
54 119

55 56 120 **Subjects and methods**

1
2
3 121 From January 1, 2011 to December 31, 2015 we conducted a cross-sectional study in an
4
5 122 academic hospital from Spain. We assessed SC responses during a SST in two study populations:

6
7
8 123 i) Main study population: Four hundred fifty one adults in whom SC concentrations at 0, 30
9
10 124 and 60 min during a SST conducted at the clinical setting for suspected AI.

11
12 125 ii) Confirmative group: One hundred fifty three women with normal HPA axis recruited
13
14 126 from our Reproductive Endocrinology clinic during the study of functional
15
16
17 127 hyperandrogenism in whom SC concentrations were obtained at 0 and 30 min during a SST
18
19 128 performed for the routine screening of non-classic congenital adrenal hyperplasia (NCAH).
20
21 129 NCAH had been ruled out in all those women because cosyntropin-stimulated 17-
22
23 130 hydroxyprogesterone and 11-deoxycortisol concentrations were below 10 ng/ml and 21
24
25 131 ng/ml, respectively [20]. None of these women were using combined contraceptives or any
26
27
28 132 other hormonal therapy at the time of sampling.

29
30 133 Before conducting the study, we obtained approval from the local ethics committee. All
31
32 134 women from our Reproductive Endocrinology clinic had previously signed an informed consent
33
34 135 form for the inclusion of a selection of coded clinical variables in an electronic database for
35
36 136 clinical research purposes that included the SC measurements presented here.
37
38
39

40 137
41
42 138 *Main study population*

43
44 139 Basal and stimulated SC values were extracted from the electronic database of our
45
46 140 Department of Clinical Biochemistry. We collected a minimum dataset in an electronic case
47
48 141 form from the clinical records of the patients including age, sex, weight, height, laboratory
49
50 142 measurements at the dates when the SST was conducted such as circulating electrolytes,
51
52 143 glomerular filtration rate and basal ACTH concentrations at the time of SST, clinical suspicion
53
54 144 of primary or central AI, other dynamic tests performed for the evaluation of adrenal function,
55
56
57

1
2
3 145 history of pituitary disease, time from hypothalamic-pituitary insult to SC determination,
4
5 146 administration of drugs that may interfere with the HPA axis, time of follow-up, and the
6
7 147 immunoassay used for SC assay. Baseline characteristics of study population are shown in Table
8
9
10 148 1.

11
12 149 We considered a clinical suspicion of potential primary AI in cases when the patients were
13
14 150 known to have adrenal disease, had required mineralocorticoid supplementation during follow-
15
16 151 up, had received drugs that may interfere with cortisol biosynthesis, had not clinical suspicion of
17
18 152 any hypothalamic-pituitary condition, and had not developed such a condition later in time.
19
20 153 Conversely, we suspected a potential central AI in subjects known to suffer from hypothalamic-
21
22 154 pituitary disease, had received drugs that may suppress the HPA axis, or when her/his referring
23
24 155 physician reported a clinical suspicion of central AI in the clinical record. All patients included
25
26 156 here had a minimum follow-up of 12 months after obtaining the SST.

27
28
29
30 157 We excluded from analysis: i) seven subjects submitted to dynamic tests other than SST
31
32 158 such as the insulin tolerance test (n = 2), corticotrophin-releasing hormone test (n = 2), oral
33
34 159 glucose tolerance test (n = 2) and glucagon stimulation test (n = 1); ii) thirty six subjects aged
35
36 160 below 18 years; iii) twenty subjects with a follow-up shorter than 12 months; iv) twelve subjects
37
38 161 in whom critically-ill related AI was suspected; and v) six subjects from whom we could not
39
40 162 obtain enough information from their clinical records as to explain the reason for conducting a
41
42 163 SST. Therefore, the study group finally included in the analyses consisted of 370 subjects.
43
44
45
46
47
48

49 165 *Confirmative group*

50
51 166 The results of SST from 153 premenopausal women with a normal HPA axis aged from 14 to
52
53 167 42 years old were included. Three women who showed a clearly subnormal SC response were
54
55 168 excluded from the analysis. In two of these women the suppressive effect on the HPA axis of the

1
2
3 169 progestins administered during 10 days before the SST with the aim of inducing a withdrawal
4
5 170 vaginal bleeding could justify the abnormal results; in the other case, we could not establish the
6
7 171 cause of the subnormal response with certainty because the patient was lost to follow-up.
8
9

10 172

11
12 173 *Assays*

13
14 174 During the study period, two immunoassays were used in our centre: i) from 2011 to July 1,
15
16 175 2013 the *Siemens Immulite 2000[®] Cortisol Immunoassay System* (immunoassay 1) was used and
17
18 176 had 6.0% and 7.8% intra- and inter-assay coefficient of variation (CV) respectively; and ii) from
19
20 177 2013 August 1, 2013 to December 31, 2015, the *Abbot Laboratories Diagnostics Division*
21
22 178 *Architect[®] Cortisol Immunoassay System* (immunoassay 2) was used, showing 3.2% and 3.4%
23
24 179 intra- and inter-assay CVs, respectively.
25
26
27

28 180

29
30 181 *Analysis of the agreement between the 30 and 60 min sampling times*

31
32 182 We analysed diagnostic agreement between the 30 and 60 min SC in patients of the main
33
34 183 study population – in the confirmation subgroup the 60 min measurement was not obtained –
35
36 184 considering two different LLN for cosyntropin-stimulated SC: i) the classic ≥ 500 nmol/l (3), and
37
38 185 ii) sex- and assay-specific cut-off values taking into also account the use of combined oral
39
40 186 contraceptives (COC) by 8 women [18]. For immunoassay 1, the reported LLN (2.5th percentile)
41
42 187 was 470 nmol/l (17 μ g/dl) in men and women, and 690 nmol/l (25 μ g/dl) for women taking
43
44 188 COC. For immunoassay 2, the LLNs were 441 nmol/l (16 μ g/dl) for men, 414 nmol/l (15 μ g/dl)
45
46 189 for women, and 579 nmol/l (21 μ g/dl) for women taking COC [17].
47
48
49

50 190

51
52 191 *Statistical analysis*

1
2
3 192 Data are shown as mean \pm standard deviation or 95% confidence interval (CI), median
4
5 193 (minimum-maximum), and raw numbers (percentage) as appropriate. The normal distribution of
6
7 194 continuous variables was assessed by the Kolmogorov-Smirnov test for one sample after a two-
8
9
10 195 step approach for transforming skewed variables if necessary [21]. Comparisons among
11
12 196 continuous variables were performed by repeated-measures ANOVA. Comparisons among
13
14 197 categorical variables were performed by Fisher's exact or χ^2 tests as appropriate. Pearson's
15
16 198 analysis served to correlate SC at 30 and 60 min samples. Consistency and absolute agreement
17
18 199 among both point times of SST were determined by their intra-class correlation coefficient (ICC)
19
20 200 with a two-factor and random-effect model. Quantitative agreement was graphically assessed by
21
22 201 Bland-Altman plots. Biochemical agreement in the diagnosis of normal or subnormal adrenal
23
24 202 was assessed by using the kappa (κ) coefficient. True positives (TP) were defined as SSTs
25
26 203 showing subnormal cortisol responses at both time points in patients who required adrenal
27
28 204 replacement therapy. True negatives (TN) were defined as SSTs showing a normal cortisol
29
30 205 response at both time points in patients who did not need glucocorticoid replacement during their
31
32 206 follow-up, did not suffer an adrenal crisis, and, when submitted to other dynamic HPA test,
33
34 207 showed normal responses. False positives (FP) for one of the sampling times consisted of the
35
36 208 finding of a subnormal response in one of the sampling times but not in the other. We calculated
37
38 209 Sp and PPV [Sp = TN / (TN + FP) and PPV = TP / (TP + FP)] for each SC sampling times
39
40 210 during the SST. A *P* value < 0.05 was considered statistically significant.
41
42
43
44
45
46
47
48

49 212 **Results**

50 213 *Main study population*

1
2
3 214 Of 370 SSTs including 30 and 60 min sampling times, SC was assayed by immunoassay
4
5 215 1 in 227 cases and by immunoassay 2 in the remaining 143 tests. Basal and cosyntropin-
6
7 216 stimulated SC concentrations, ACTH levels when available, and the median duration of follow-
8
9 217 up in patients with either normal or insufficient responses are shown in **Table 2**.

10
11
12 218 SC concentrations when patients in the main study group were analyzed as a whole are
13
14 219 represented in **Figure 1A**. SC concentrations at 30 and 60 min during the SST increased when
15
16 220 compared to baseline values (**Figure 1, panel A**), and showed a very strong linear correlation
17
18 221 (**Figure 1, panel B**). Baseline SC concentrations correlated with 30 min SC measurements ($r =$
19
20 222 0.735 , $P = 0.001$), and with 60 min SC values ($r = 0.660$, $P = 0.001$).

21
22
23 223 Similar results were observed when analyzing separately the 150 SSTs performed with
24
25 224 the aim of to ruling out primary AI (correlation between baseline SC and 30 min SC: $r = 0.720$, P
26
27 225 $= 0.001$), and correlation between baseline SC and 60 min SC: $r = 0.640$, $P = 0.001$) and the 220
28
29 226 SSTs conducted to exclude central AI (correlation between baseline SC and 30 min SC: $r =$
30
31 227 0.723 , $P = 0.001$, and correlation between baseline SC and 60 min SC: $r = 0.644$ ($P = 0.001$).

32
33 228 The ICC among SC concentrations as assayed at both sampling times showed a very
34
35 229 good consistence index (0.940; 95%CI: 0.928 – 0.952) and a good absolute agreement (0.889,
36
37 230 95%CI: 0.465 – 0.957), even though the latter only qualifies as fair according to the lower limit
38
39 231 of the 95%CI. The Bland-Altman plot (**Figure 1, panel C**) showed a good agreement between
40
41 232 SC assayed at 30 and 60 min, with a slight tendency towards greater percentage differences with
42
43 233 decreasing mean values of stimulated SC.

44
45 234 **Figure 2 and Table 2** show SC concentrations as a function of the clinical suspicion and
46
47 235 whether or not the result of the SST was normal. The diagnostic agreement among both sampling
48
49 236 times according to classic and to sex – and assay-specific cut-off values is shown in **Figure 3**.
50
51 237 Disagreements between both sampling times were as follows. When relying on the classic SC

1
2
3 238 cut-off point (≥ 500 nmol/l), 39 cases (10.5%) had a subnormal response at 30 min that reached
4
5 239 normal values at 60 min whilst, in 3 patients (0.8%), a normal response at 30 min ended being
6
7 240 subnormal at 60 min. Using sex- and assay-specific values, 34 cases (9.2%) showed subnormal
8
9 241 responses at 30 min but normal SC concentrations at 60 min, whereas in 5 cases (1.3%), the
10
11 242 response was normal at 30 min but subnormal at 60 min.

12
13
14 243 The analysis of the diagnostic agreement as a function of the suspicion of primary versus
15
16 244 central AI is shown in **Figure 4**. As a rule, agreement among both sampling times of the SST
17
18 245 was better when primary AI was suspected compared with a suspicion of central AI. When using
19
20 246 classic cut-off values to rule out primary AI, 7 cases (4.7%) showed a subnormal response at 30
21
22 247 min that reached normal concentrations at 60 min, whereas no subject with a normal response at
23
24 248 30 min had a subnormal response at 60 min. Using sex- and assay-specific cut-off values, in 6
25
26 249 cases (4.0%) the response was subnormal at 30 min but reached normal concentrations at 60 min.
27
28 250 Four of them showed a subnormal SC responses to cosyntropin that were very close to the cut-
29
30 251 off value. In these subjects, the differences between the cut-off value and the stimulated SC
31
32 252 ranged from 22 to 39 nmol/l (0.8 to 1.4 $\mu\text{g}/\text{dl}$), very small concentrations that are, in fact,
33
34 253 included within the CV of the assays, thereby suggesting no clinical relevance. The two
35
36 254 remaining patients showed peak SC concentrations of 320 and 364 nmol/l (11,6 and 13,2 $\mu\text{g}/\text{dl}$) at
37
38 255 the 30 min sampling time: one had received oral glucocorticoid replacement therapy that did not
39
40 256 preclude the patient of responding to cosyntropin by showing a SC of 470 nmol/l (17 $\mu\text{g}/\text{dl}$) at
41
42 257 the 60 min sample, and the other subject was submitted to SST because of the presence of
43
44 258 bilateral adrenal hyperplasia and did not show any signs or symptoms of AI nor suffered an
45
46 259 adrenal crisis during follow-up. None of the SSTs showing normal responses at 30 min had a
47
48 260 subnormal response at 60 min.

1
2
3 261 When central AI was suspected and the classic cut-off point was applied, 32 cases
4
5 262 (14.5%) had a normal SC response at 60 min but a subnormal SC value at 30 min. Only 3
6
7 263 subjects (1.4%) presented with the opposite situation. Using sex- and assay-specific cut-off
8
9 264 concentrations, 28 cases (12.7%) showed a normal response at 60 min but a subnormal result at
10
11 265 30 min, yet in only 5 cases (2.3%) the contrary occurred. These 5 subjects had been evaluated in
12
13 266 the context of withdrawal of prolonged glucocorticoid therapy during the first year after a
14
15 267 pituitary insult (surgery and/or pituitary radiotherapy). Three of them showed a complete
16
17 268 recovery of their HPA axis throughout the follow-up period, whereas in the other two patients,
18
19 269 who had received pituitary radiotherapy, the subnormal response to cosyntropin was maintained
20
21 270 over time.

22
23
24
25
26 271 The Sp and PPV for different sampling times and cut-off values used here are shown in
27
28 272 Table 3. SC concentrations at 60 min had a higher Sp and PPV compared with 30 min
29
30 273 measurements, particularly when central AI was suspected. Nonetheless, the Sp of the
31
32 274 determination at 30 min was as high as 95% when SST had been performed to rule out primary
33
34 275 disease both when applying classic or sex- and assay-specific cut-off values.

35
36
37 276 We observed discordant results between classic and sex- and assay-specific cut-off
38
39 277 concentrations in 50 cases. In 47 of these subjects, a subnormal response using the classic cut-off
40
41 278 value turned into a normal response had sex- and assay-specific cut-offs been used. In 7 of them,
42
43 279 SST was performed to rule out primary AI and in the remaining 40 subjects the SSTs were
44
45 280 conducted to rule out central AI. Glucocorticoid replacement was started in 18 cases, and no
46
47 281 subject presented with signs or symptoms of chronic or acute AI. In addition, from the 50
48
49 282 discordant SSTs, 3 were conducted in women under estrogenic therapy and presented a normal
50
51 283 response according to the classic cut-off value, but subnormal when considering sex- and assay-
52
53 284 specific cut-offs, yet none of them required glucocorticoid therapy.

285

286 *Confirmative group*

287 Thirty (20%) of these women presented with a subnormal response to SST according to

288 classic cut-off values, yet this figure was reduced to only 3 (2%) when sex- and assay-specific

289 cut-off values were used [observed agreement: 82%; κ : 0.151 (95%CI: 0.066-0.235)]. The three

290 women showing a subnormal response during SST using a sex- and assay-specific cut-off value

291 showed stimulated SC concentrations of 342 nmol/l (12.4 μ g/dl), 353 nmol/l (12.8 μ g/dl) and

292 372 nmol/l (13.5 μ g/dl), whereas the LLNs (2.5th percentile) of SC concentrations at 30 min

293 sampling time of SST were 436 nmol/l (15.8 μ g/dl) and 411 nmol/l (14.9 μ g/dl) for

294 immunoassays 1 and 2, respectively. The 5th percentiles for both immunoassays were 450 nmol/l

295 (16.3 μ g/dl) and 414 nmol/l (15.0 μ g/dl), respectively, showing minimal differences with the

296 LLNs (**Figure 5**). None of these female controls developed any HPA disease during their follow-

297 up.

299 DISCUSSION

300 AI is a clinical condition associated with a high morbidity and mortality. Unstimulated

301 early morning SC values below 138 nmol/l (5 μ g/dl) show a high PPV for AI, whereas

302 concentrations over 500 nmol/l predict a normal adrenal response. However, values between 138

303 and 500 nmol/l are considered indeterminate and require adrenal stimulation to confirm or rule

304 out a diagnosis, always in consonance with the clinical picture [1–3].

305 Baseline SC concentrations showed stronger linear correlations with cosyntropin-

306 stimulated SC levels at 30 and 60 min samples of the SST, in agreement with previous reports

307 [22]. Our data also show that both 30 and 60 min SC measurements during a SST have an

308 adequate index of consistency, but the same is not true in terms of absolute agreement,

1
2
3 309 particularly when a central AI is suspected. Furthermore, a single determination at 60 min during
4
5 310 the SST appears to have the higher Sp and PPV for the diagnosis of subjects presenting with
6
7 311 either primary or central AI. In consonance, after evaluating retrospectively 73 subjects, Zueger
8
9 312 et al. [23] reported that sampling at 30 min of the SST did not provide any additional diagnostic
10
11 313 advantage over performing a single determination at 60 min of the test. Although similar results
12
13 314 have been also reported by others [13,14], these studies did not take into account the primary or
14
15 315 central origin of AI and did not apply sex- and assay specific cut-off values, a fact of paramount
16
17 316 importance because of the considerable influence that cortisol immunoassays exerts on the final
18
19 317 values observed after cosyntropin-stimulation [17,18].

20
21
22
23
24 318 Our results also indicate that SC measurement at 30 min during the SST, when using sex-
25
26 319 and assay-specific cut-off values, are enough to rule out clinically relevant primary AI since only
27
28 320 4% of patients in this particular situation showed a subnormal response at 30 min followed by
29
30 321 normal response at 60 min. Furthermore, these subjects presented with stimulated SC
31
32 322 concentrations which were very close to the cut-off concentrations, to the extent that the
33
34 323 differences with these normal limits may be explained by the analytical variability of these
35
36 324 commercial immunoassays used here. Even more important from a clinical point of view, none
37
38 325 of these subjects required replacement therapy during their follow-up, suffered an acute adrenal
39
40 326 crisis, nor were diagnosed with any adrenal condition during follow-up, strongly suggesting that
41
42 327 their HPA function was actually normal at the time the SST was performed. The use of sex- and
43
44 328 assay-specific cut-off values appears to be essential, since other authors have suggested that
45
46 329 some healthy individual may have a delayed response to SST using classic reference values [24].

47
48
49
50
51 330 On the other hand, 60 min samples appears to be more specific than 30 min
52
53 331 measurements when central AI is suspected. In such a case, 12.7% of the subjects presenting
54
55 332 with a subnormal response at 30 min actually had a normal response at 60 min, avoiding

1
2
3 333 unnecessary treatments in them. Although a subnormal response 30 min after cosyntropin-
4
5 334 stimulation in patients with suspicion of secondary AI may not translate into the need of adrenal
6
7 335 replacement in a non-critical scenario, it is likely that most physicians would feel more confident
8
9
10 336 with not starting replacement therapy after obtaining a cosyntropin-stimulated SC concentration
11
12 337 above the LLN, favoring the use of 60 min samples over 20 min determinations for this
13
14 338 particular reason. Furthermore, relying mostly on 60 min SC responses to cosyntropin when
15
16 339 suspecting a central origin of AI is also supported by the fact that, in 2 out of the 5 patients in our
17
18 340 series who showed a subnormal response at 60 min preceded by normal SC values at 30 min, AI
19
20 341 was actually confirmed during follow-up because of former pituitary radiotherapy.
21
22

23
24 342 Our present findings also reinforce the need of sex- and assay-specific cut-off values to
25
26 343 interpret the results of the SST, in agreement with recent clinical guidelines[3]. The use of such
27
28 344 cut-off values lead in our study to a reduction in FP results, higher Sp and PPV, less discordant
29
30 345 results among sampling times of the SST, and fewer unnecessary treatments [20 patients (5%)
31
32 346 could have been treated unnecessarily if classic cut-off values were applied for diagnosis]. The
33
34 347 reliability of sex- and assay-specific cut-off values was confirmed in our population of
35
36 348 premenopausal women with normal HPA axis, in whom these cut-offs were more appropriate
37
38 349 than relying on classic values to assess the functionality of their HPA axis. In this population, the
39
40 350 LLNs for stimulated SC at 30 min were very close to those reported for each immunoassay by
41
42 351 the manufacturers, which relied on the 2.5th percentile [17], yet reinforcing the need to establish
43
44 352 local normative data in order to improve the diagnostic accuracy of cortisol measurements during
45
46 353 SSTs [17,25].
47
48
49
50

51 354 Among the strengths of our study, we would highlight the large series of subjects
52
53 355 suspected of suffering AI who were evaluated with a standardized dynamic study, and the careful
54
55 356 review of subjects' medical records that followed such evaluations. However, we are aware of
56
57
58
59

1
2
3 357 several weaknesses derived from the observational and retrospective design of the study, making
4
5 358 impossible to rule out information bias. Our best efforts might have not been enough to avoid
6
7 359 misclassification of patients according to the suspicion of primary or central AI. Also, the
8
9
10 360 administration of supraphysiological doses of cosyntropin does not permit ruling out partial
11
12 361 deficiencies either, particularly in those suspected of central HPA defects. Also, and even
13
14 362 considering the large sample of subjects included in our study, our present results may not be
15
16 363 extrapolable to other populations in whom SC has been measured with different immunoassays
17
18 364 that would require specific local normative data. Moreover, analysis of Sp and PPV has not been
19
20 365 challenged against a biochemical gold-standard in most cases and, as a consequence, we have not
21
22 366 been able to establish false negative rates, sensitivity and negative predictive values.
23
24 367 Nonetheless, besides those assessments had been unethical in most cases, the lack of a laboratory
25
26 368 gold-standard such as an ITT did not override our results, since from a practical point of view,
27
28 369 we are looking for patients needing replacement therapy and not for those with a partial AI who
29
30 370 do not require any treatment. Another limitation was that the confirmation group is not fully
31
32 371 representative of our main study population since was only comprised of premenopausal women
33
34 372 and stimulated SC was only available at the 30 min sampling time. Lastly, we could not rule out
35
36 373 entirely pre-treatment with progestogens in the context of induction of withdrawal bleeding in
37
38 374 our confirmative population. Because these drugs might exert a mild suppressive effect on the
39
40 375 HPA axis [19,26], their administration in a few cases could have, at least in theory, lowered
41
42 376 stimulated SC values.
43
44
45
46
47
48
49
50

377

51 378 **CONCLUSIONS**

53
54 379 Compared with the use of classic cut-off values derived from the literature, application of
55
56 380 sex- and assay-specific cut-off values of SC responses to cosyntropin results into higher Sp and
57

1
2
3 381 PPV for establishing a diagnosis of AI, thereby avoiding unnecessary treatments. Measurement
4
5 382 of stimulated SC at 30 min after cosyntropin-stimulation may suffice for the correct diagnosis of
6
7 383 primary AI, yet 60 min measurements might be preferable when central AI is suspected.
8
9

10 384

11
12 385 **Conflict of interest:** None.
13

14 386

15
16
17 387 **Funding:** This work has been supported by grant PI1400649) from Fondo de Investigación
18
19 388 Sanitaria, Instituto de Salud Carlos III, Spanish Ministry of Economy and Competitiveness.
20
21 389 M.L.-R. has a local grant for clinical research from the Instituto Ramón y Cajal de Investigación
22
23 390 Sanitaria (IRYCIS). CIBERDEM is also an initiative of Instituto de Salud Carlos III, partially
24
25 391 supported by FondoEuropeo de Desarrollo Regional FEDER. There were no other sources of
26
27 392 funding.
28
29

30 393

31
32
33
34 394 **Data sharing statement:** Individual participant data that underlie the results reported in this
35
36 395 article, after deidentification, so as the study protocol would be available immediately after
37
38 396 publication to anyone who wishes to access the data to achieve aims in the approved proposal
39
40 397 and for individual participant data meta-analysis. Proposals should be directed to
41
42 398 andres_ortiz_f@yahoo.com or to manuel.luque@salud.madrid.org. To gain access, data
43
44 399 requestors will need to sign a data access agreement.
45
46
47

48 400

49
50
51 401 **Contributorship statement:** A.O.-F. y M.L.-R. designed the protocol, and performed the
52
53 402 statistical analysis. A.O.-F. y E.S.-C. reviewed the clinical data using the electronic or written
54
55
56
57
58
59

1
2
3 403 records if necessary. A.G.-C. and L.J.-M. performed the electronic search of serum cortisol
4
5 404 samples. A.O.-F. y M.L.-R. wrote the first draft of the study. All the authors, including L.N.-C.
6
7 405 and H.F.E.-M, reviewed the manuscript before its submission and contributed to intellectual
8
9
10 406 content. All the authors have accepted responsibility for the entire content of the manuscript and
11
12 407 approved the final submission.
13
14
15 408

For peer review only

409 **REFERENCES**

- 410 1. Charmandari E, Nicolaidis NC, Chrousos GP. Adrenal insufficiency. *The Lancet*.
411 2014;383(9935):2152-67.
- 412 2. Neary N, Nieman L. Adrenal Insufficiency- etiology, diagnosis and treatment. *Curr Opin*
413 *Endocrinol Diabetes Obes*. 2010;17(3):217-23.
- 414 3. Bornstein SR, Allolio B, Arlt W, Barthel A, Don-Wauchope A, Hammer GD, et al.
415 *Diagnosis and Treatment of Primary Adrenal Insufficiency: An Endocrine Society Clinical*
416 *Practice Guideline*. *J Clin Endocrinol Metab*. 2016;101(2):364-89.
- 417 4. Grossman AB. The Diagnosis and Management of Central Hypoadrenalism. *J Clin*
418 *Endocrinol Metab*. 2010;95(11):4855-63.
- 419 5. Agha A, Tomlinson JW, Clark PM, Holder G, Stewart PM. The Long-Term Predictive
420 Accuracy of the Short Synacthen (Corticotropin) Stimulation Test for Assessment of the
421 Hypothalamic-Pituitary-Adrenal Axis. *J Clin Endocrinol Metab*. 2006;91(1):43-7.
- 422 6. Sprung CL, Annane D, Keh D, Moreno R, Singer M, Freivogel K, et al. Hydrocortisone
423 Therapy for Patients with Septic Shock. *N Engl J Med*. 2008;358(2):111-24.
- 424 7 Marik PE, Pastores SM, Annane D, *et al*. Recommendations for the diagnosis and
425 management of corticosteroid insufficiency in critically ill adult patients: consensus
426 statements from an international task force by the American College of Critical Care
427 Medicine. *Crit Care Med* 2008;36:1937-49.

- 1
2
3 428 8. Annane D, Sébille V, Troché G, Raphaël J-C, Gajdos P, Bellissant E. A 3-Level Prognostic
4
5 429 Classification in Septic Shock Based on Cortisol Levels and Cortisol Response to
6
7 430 Corticotropin. *JAMA*. 2000;283(8):1038-45.
9
10
11 431 9. Lindholm J, Kehlet H. Re-Evaluation of the Clinical Value of the 30 Min Acth Test in
12
13 432 Assessing the Hypothalamic-Pituitary-Adrenocortical Function. *Clin Endocrinol (Oxf)*.
14
15 433 1987;26(1):53-9.
17
18
19 434 10. Hurel SJ, Thompson CJ, Watson MJ, Harris MM, Baylis PH, Kendall-Taylor P. The short
20
21 435 Synacthen and insulin stress tests in the assessment of the hypothalamic-pituitary-adrenal
22
23 436 axis. *Clin Endocrinol (Oxf)*. 1996;44(2):141-6.
25
26
27 437 11. Gonzálbez, Villabona, Ramón, Navarro, Giménez, Ricart, et al. Establishment of reference
28
29 438 values for standard dose short synacthen test (250 µg), low dose short synacthen test (1 µg)
30
31 439 and insulin tolerance test for assessment of the hypothalamo-pituitary-adrenal axis in
32
33 440 normal subjects. *Clin Endocrinol (Oxf)*. 2000;53(2):199-204.
35
36
37 441 12. Longui CA, Vottero A, Harris AG, Chrousos GP. Plasma cortisol responses after
38
39 442 intramuscular corticotropin 1-24 in healthy men. *Metab - Clin Exp*. 1998;47(11):1419-22.
40
41
42 443 13. Mansoor S, Islam N, Siddiqui I, Jabbar A. Sixty-minute post-Synacthen serum cortisol level:
43
44 444 a reliable and cost-effective screening test for excluding adrenal insufficiency compared to
45
46 445 the conventional short Synacthen test. *Singapore Med J*. 2007;48(6):519-23.
47
48
49
50 446 14. Chitale A, Musonda P, McGregor AM, Dhatariya KK. Determining the utility of the 60 min
51
52 447 cortisol measurement in the short synacthen test. *Clin Endocrinol (Oxf)*. 2013;79(1):14-9.
53
54
55
56
57
58
59

- 1
2
3 448 15. de Miguel Novoa P, Vela ET, García NP, Rodríguez MM, Guerras IS, Martínez de Salinas
4
5 449 Santamaría M de LÁ, et al. Guidelines for the diagnosis and treatment of adrenal
6
7 450 insufficiency in the adult. *Endocrinol Nutr Organo Soc Espanola Endocrinol Nutr.* 2014;61
8
9 451 (Suppl 1):1-35.
- 11
12
13 452 16. Turpeinen U, Hämäläinen E. Determination of cortisol in serum, saliva and urine. *Best Pract*
14
15 453 *Res Clin Endocrinol Metab.* 2013;27(6):795-801.
- 17
18 454 17. El-Farhan N, Pickett A, Ducroq D, Bailey C, Mitchem K, Morgan N, et al. Method-specific
19
20 455 serum cortisol responses to the adrenocorticotrophin test: comparison of gas
21
22 456 chromatography-mass spectrometry and five automated immunoassays. *Clin Endocrinol*
23
24 457 *(Oxf).* 2013;78(5):673-80.
- 25
26
27
28 458 18. Klose M, Lange M, Rasmussen AK, Skakkebaek NE, Hilsted L, Haug E, et al. Factors
29
30 459 Influencing the Adrenocorticotropin Test: Role of Contemporary Cortisol Assays, Body
31
32 460 Composition, and Oral Contraceptive Agents. *J Clin Endocrinol Metab.* 2007;92(4):1326-33.
- 33
34
35
36 461 19. Kontula K, Paavonen T, Luukkainen T, Andersson LC. Binding of progestins to the
37
38 462 glucocorticoid receptor: Correlation to their glucocorticoid-like effects on in vitro functions
39
40 463 of human mononuclear leukocytes. *Biochem Pharmacol.* 1983;32(9):1511-8.
- 41
42
43
44 464 20. Escobar-Morreale HF, Sanchón R, San Millán JL. A prospective study of the prevalence of
45
46 465 nonclassical congenital adrenal hyperplasia among women presenting with hyperandrogenic
47
48 466 symptoms and signs. *J Clin Endocrinol Metab.* 2008;93(2):527-33.

- 1
2
3 467 21. Templeton G. A Two-Step Approach for Transforming Continuous Variables to Normal:
4
5 468 Implications and Recommendations for IS Research. *Commun Assoc Inf Syst.* 2011;28(1).
6
7 469 Available at: <http://aisel.aisnet.org/cais/vol28/iss1/4>
9
10
11 470 22 Dorin RI, Qiao ZG, Bouchonville M, *et al.* Characterization of Cortisol Secretion Rate in
12
13 471 Secondary Adrenal Insufficiency. *J Endocr Soc* 2017;1:945–56.
14
15
16 472 23. Zueger T, Jordi M, Laimer M, *et al.* Utility of 30 and 60 minute cortisol samples after high-
17
18 473 dose synthetic ACTH-1-24 injection in the diagnosis of adrenal insufficiency. *Swiss Med*
19
20 474 *Wkly* 2014;144:w13987.
21
22
23
24 475 24. Karaca Z, Lale A, Tanriverdi F, *et al.* The comparison of low and standard dose ACTH and
25
26 476 glucagon stimulation tests in the evaluation of hypothalamo-pituitary-adrenal axis in healthy
27
28 477 adults. *Pituitary* 2011;14:134–40.
29
30
31
32 478 25. Clark PM, Neylon I, Raggatt PR, *et al.* Defining the normal cortisol response to the short
33
34 479 Synacthen test: implications for the investigation of hypothalamic-pituitary disorders. *Clin*
35
36 480 *Endocrinol (Oxf)* 1998;49:287–92.
37
38
39
40 481 26. Bornstein SR. Predisposing Factors for Adrenal Insufficiency. *N Engl J Med.*
41
42 482 2009;360(22):2328-39.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

TABLE 1. Baseline characteristics population as a function as clinical suspicion of primary or secondary adrenal disease

	Clinical suspicion of primary AI (n = 150)		Clinical suspicion of central AI (n= 220)	
	Female (n = 98)	Male (n = 52)	Female (n =139)	Male (n = 81)
Sex				
Age (years)	53 ± 19	56 ± 15	55 ± 16	55 ± 13
Weight(kg)	59 ± 12	72 ± 14	72 ± 14	84 ± 16
BMI (kg/m ²)	24 ± 5	24 ± 5	28 ± 5	29 ± 5
Na (mmol/l)	138 ± 3	137 ± 5	139 ± 2	140 ± 4
K (mmol/l)	4.3 ± 0.6	4.4 ± 0.7	4.1 ± 0.4	4.1 ± 0.3
Ca (mmol/l)	2.3 ± 0.1	2.3 ± 0.2	2.4 ± 0.1	2.3 ± 0.1
Cr (µmol/l)	62 (44 – 1114)	80 (44 – 1158)	62 (18 – 875)	80 (44 – 150)
eGFR (MDRD) (ml/min/1.73m ²)	87 (4 – 137)	86 (4 – 183)	84 (5 – 361)	93 (43 – 163)

Abbreviations, BMI, body mass index; Ca, total serum calcium; Cr, serum creatinine; eGFR, estimated glomerular filtration rate; K, serum potassium; Na, serum sodium.

Data are presented as mean ± SD or median (minimum-maximum) as appropriate.

TABLE 2. Basal and cosyntropin-stimulated serum cortisol concentrations as a function of the presence of a normal or abnormal result during the SST, and mean follow-up of the patients in each subgroup. From the whole sample, ACTH measurements were available for 342 samples.

	Normal responses at both times (n = 307)	Confirmed primary AI (n = 18)	Confirmed secondary AI (n = 45)
Basal ACTH (pmol/l)	4 (1 – 43)	6 (1 – 71) *	3 (1 - 11)
Basal SC (nmol/l)	386 ± 166	165 ± 110	138 ± 83
SC at 30 min (nmol/l)	662 ± 193	248 ± 110	276 ± 110
SC at 60 min (nmol/l)	745 ± 221	304 ± 138	304 ± 110
Follow-up (months)	37 ± 17	43 ± 18	36 ± 15

Data are presented as mean ± SD or median (minimum-maximum) as appropriate. To convert SC to metric units, multiply nmol/l by 0.03625 (result in µg/dl). To convert ACTH to metric units, multiply pmol/l by 4.54545 (result in pg/ml). * Despite not having any hypothalamic-pituitary condition at diagnosis or throughout their follow-up, and not having received drugs that suppress the HPA axis, seven patients with clinical suspicion of primary disease who required replacement therapy presented with normal ACTH levels. Three of them had begun glucocorticoid therapy at the time of SST. In another 4 cases, there is a strong suspicion of that was the case, although the possibility of an inadequate sample processing also existed (i.e.: sample transport at room temperature). *Abbreviations, ACTH, adrenocorticotropin hormone; AI, adrenal insufficiency; SC, serum cortisol.*

TABLE 3. Specificity and positive predictive value (PPV) of the short high-dose cosyntropin test, for the diagnosis of adrenal insufficiency (AI), according to serum cortisol cut-off concentrations (classic and sex- and assay-specific), and as a function of the suspected origin of the disease.

	Classic cut-off values						Sex- and assay-specific cut-off values					
	Global		Clinical suspicion Primary AI		Clinical suspicion Central AI		Global		Clinical suspicion Primary AI		Clinical suspicion Central AI	
	30	60	30	60	30	60	30	60	30	60	30	60
Sampling time (min)												
Specificity (%)	86	99	95	100	79	98	89	98	96	100	84	97
PPV (%)	68	97	74	100	66	96	65	93	75	100	61	90

Legend to figures

Figure 1. Panel A, Serum cortisol levels at different sampling times. Data are shown as mean (95%CI) and mean differences (MD) (95%CI). Comparisons among time points were performed by a repeated-measure ANOVA addressing main effects by a Bonferroni's confidence interval adjustment. * P value < 0.001. **Panel B**, Pearson's correlation analysis between serum cortisol values at 30 and 60 min sampling times. Solid red line represents the simple linear regression and dotted black lines represent the 95%CI of the regression line. **Panel C**, Bland-Altman plot. Solid black line represents the perfect agreement among both time points. Solid blue line is the mean of the percentage difference among both sampling times, and dashed blue lines are ± 2 standard deviation (SD) of that mean. Solid red line is the regression line of the percentage differences.

Figure 2. Baseline and stimulated serum cortisol concentrations as a function of clinical suspicion and response to cosyntropin test. Data are shown as mean and 95%CI. *Abbreviations:*

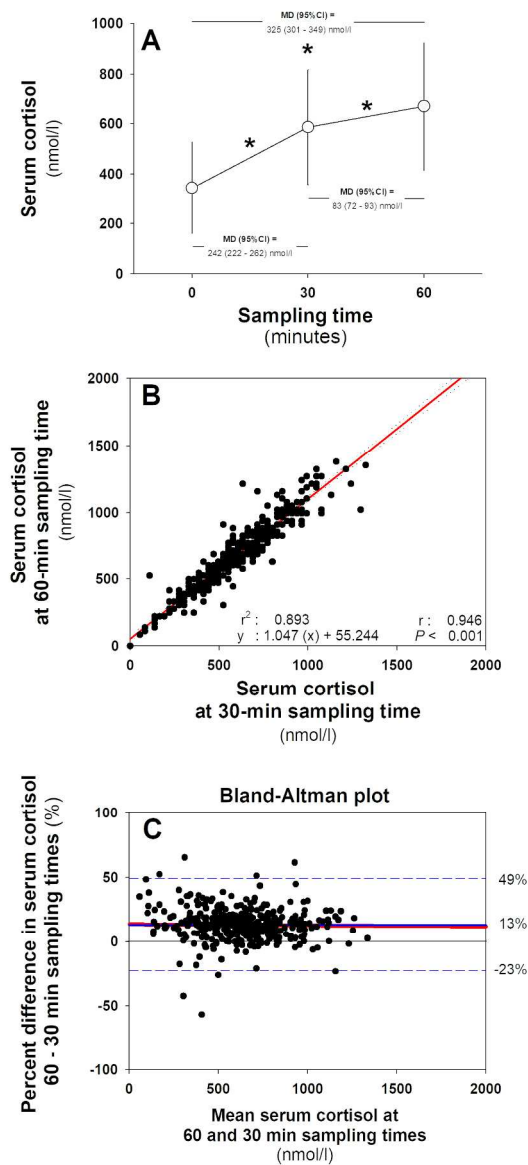
AI: Adrenal Insufficiency

Figure 3. Subgroups of patients according to serum cortisol responses to cosyntropin-stimulation as a function of classic and sex- and assay-specific cut-offs. Figures on top of the bars indicate the number of patients included in each subgroup. Diagnostic agreement is shown as the percentage of observed agreements and kappa coefficients (95%CI).

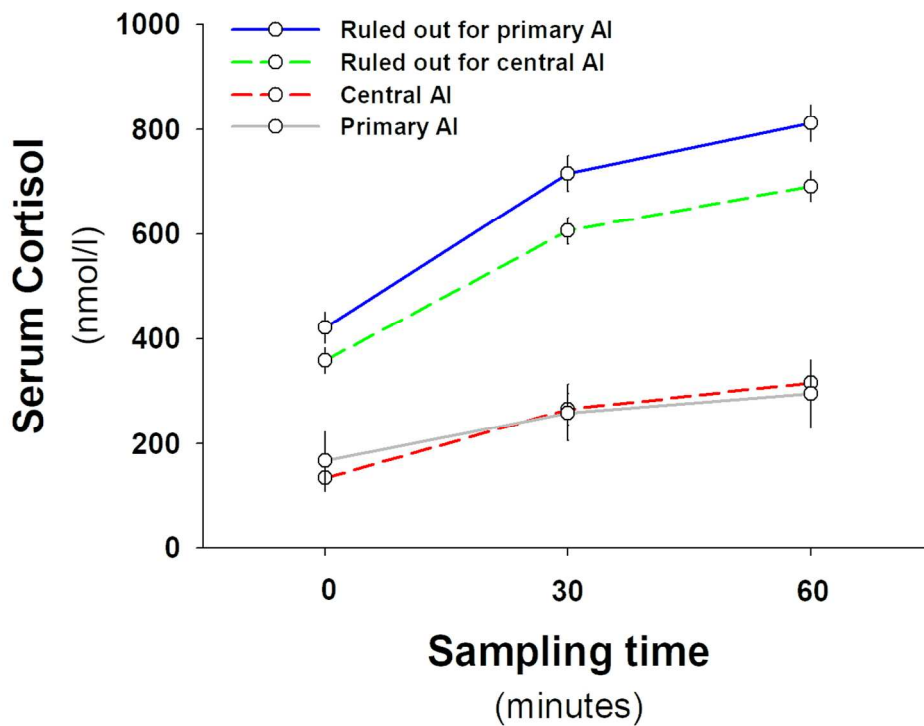
Figure 4. Subgroups of patients according to serum cortisol responses to cosyntropin-stimulation as a function of cut-off values and clinical suspicion of primary or central AI. Figures on top of

1
2
3 the bars show the number of patients included in the different subgroups. Diagnostic agreement
4
5 is shown as the percentage of observed agreements and kappa coefficients (95%CI).
6
7
8
9

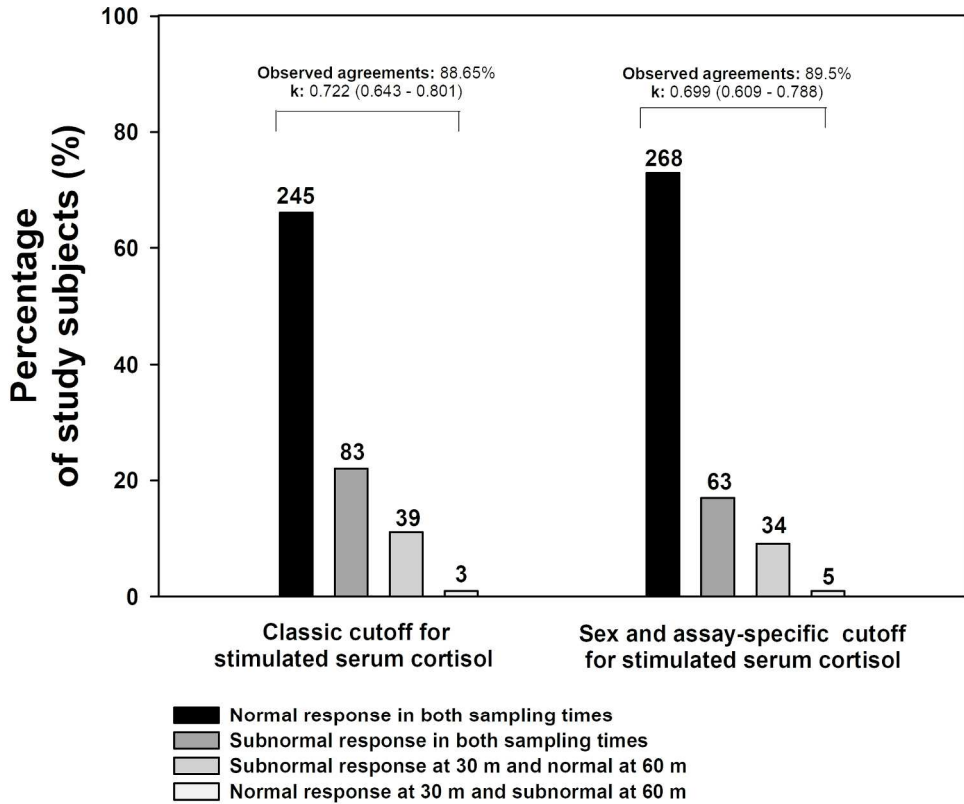
10 **Figure 5.** Descriptive statistics and distribution of 30 min cosyntropin-stimulated serum cortisol
11 concentrations in a population of premenopausal healthy women with evidence of normal HPA
12 axis function. The boundary of the box closest to zero indicates the 25th percentile, the solid and
13 long dash lines within the box marks the median and mean, respectively, and the boundary of the
14 farthest from zero indicates the 75th percentile. Whiskers above and below the box indicate the
15 90th and 10th percentiles. Black circles represent the 5th percentile and the dashed red line
16 indicates the lower limit of normality (2.5th percentile) for each immunoassay.
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



199x282mm (300 x 300 DPI)



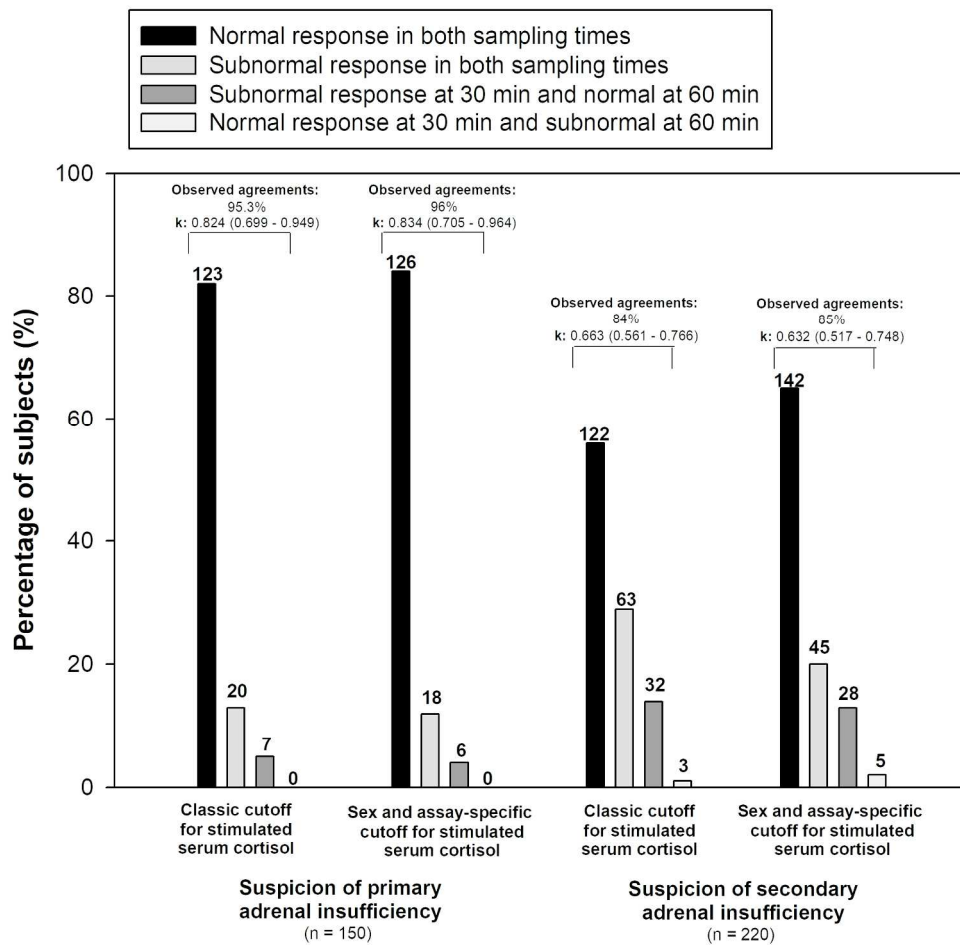
118x97mm (300 x 300 DPI)



214x180mm (300 x 300 DPI)

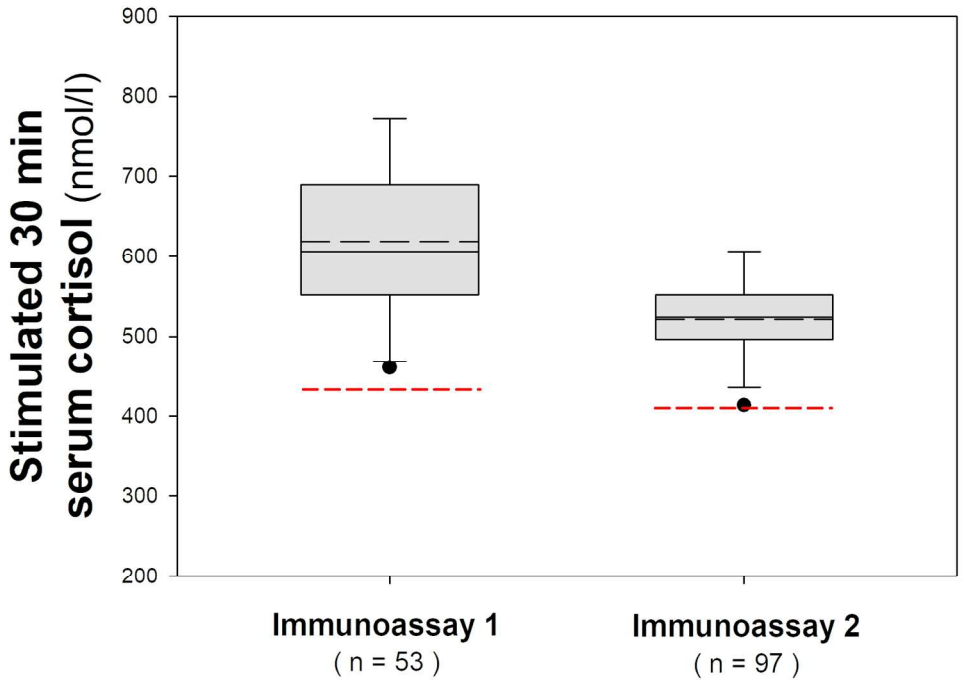
Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



208x195mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



160x120mm (300 x 300 DPI)

View Only

STROBE Statement—checklist of items that should be included in reports of observational studies

Title: Importance of sampling time and assay cut-offs for routine assessment of adrenal function: an observational longitudinal study.

	Item No	Recommendation
Title and abstract Page: 1	1	(a) Indicate the study's design with a commonly used term in the title or the abstract. Page 1 (b) Provide in the abstract an informative and balanced summary of what was done and what was found. Page 2
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported. Page 4
Objectives	3	State specific objectives, including any prespecified hypotheses. Page 5
Methods		
Study design	4	Present key elements of study design early in the paper. Page 5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection. Page 5
Participants	6	(<i>Cross-sectional study</i>)—Give the eligibility criteria, and the sources and methods of selection of participants. Page: 5-6
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable. Page: 5-7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Page: 5-7
Bias	9	Describe any efforts to address potential sources of bias. Page 5-7
Study size	10	Explain how the study size was arrived at. Page: 5-7
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why. Page: 5-7
Statistical methods Page: 8-9	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses

Continued on next page

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed Page: 5-6 (b) Give reasons for non-participation at each stage Page: 5-6 (c) Consider use of a flow diagram Not considered.
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Page 5 - 6 (b) Indicate number of participants with missing data for each variable of interest. Page: 5-6 <i>Cross-sectional study</i> —Report numbers of outcome events or summary measures Page: 5-6
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included Page 9-11 (b) Report category boundaries when continuous variables were categorized Page 9-11 (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period. Not relevant
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses Page 9-11
Discussion		
Key results	18	Summarise key results with reference to study objectives Page 12 - 15
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias Page 14
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence Page 14
Generalisability	21	Discuss the generalisability (external validity) of the study results Page 15
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based Page 15

BMJ Open

Role of sampling times and serum cortisol cut-off concentrations on the routine assessment of adrenal function using the standard cosyntropin test in an academic hospital from Spain: a retrospective chart review.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-019273.R2
Article Type:	Research
Date Submitted by the Author:	23-Jan-2018
Complete List of Authors:	Ortiz-Flores, Andrés; Hospital Universitario Ramón y Cajal, Endocrinology Santacruz, Elisa; Hospital Universitario Ramon y Cajal Jiménez-Mendiguchia, Lucía; Hospital Universitario Ramon y Cajal García-Cano, Ana; Hospital Universitario Ramon y Cajal Nattero-Chávez, Lía; Hospital Universitario Ramon y Cajal Escobar-Morreale, Héctor; Hospital Universitario Ramon y Cajal Luque-Ramírez, Manuel; Hospital Universitario Ramon y Cajal
Primary Subject Heading:	Diabetes and endocrinology
Secondary Subject Heading:	Diagnostics, Patient-centred medicine
Keywords:	Adrenal disorders < DIABETES & ENDOCRINOLOGY, cortisol, biochemical diagnosis, specificity, immunoassay

SCHOLARONE™
Manuscripts

Title Role of sampling times and serum cortisol cut-off concentrations on the routine assessment of adrenal function using the standard cosyntropin test in an academic hospital from Spain: a retrospective chart review.

Short title Cosyntropin test and adrenal insufficiency

Authors Andrés E. Ortiz-Flores, M.D.,¹ Elisa Santacruz-Cerdá, M.D.,¹ Lucía Jiménez-Mendiguchia, M.D.,² Ana García-Cano, M.D.,² Lia Nattero-Chávez, M.D.,¹ Héctor F. Escobar-Morreale, M.D., Ph.D.,^{1,3} Manuel Luque-Ramírez, M.D., Ph.D.^{1,3}

Affiliations

¹ Department of Endocrinology and Nutrition, Hospital Universitario Ramón y Cajal, Madrid (Spain).

² Department of Clinical Biochemistry, Hospital Universitario Ramón y Cajal, Madrid (Spain).

³ Diabetes, Obesity and Human Reproduction Research Group, Hospital Universitario Ramón y Cajal & Universidad de Alcalá & Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS) & Centro de Investigación Biomédica en Red Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), Madrid, Spain.

Word count: 5515 (including tables and figure legends); **Word count for abstract:** 300;

Figures: 5; **Tables:** 3; **Number of references:** 26.

CORRESPONDENCE & REPRINTS

1
2
3 24 Andres E. Ortiz-Flores & Manuel Luque-Ramírez. Department of Endocrinology and Nutrition,
4
5 25 Hospital Universitario Ramón y Cajal, Carretera de Colmenar Km 9.1, E-28034 Madrid, Spain.
6
7
8 26 Phone/Fax: +34-91 336 9029.
9
10
11 27 E-mail: andres_ortiz_f@yahoo.com / manuel.luque@salud.madrid.org
12
13
14 28

15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

29 ABSTRACT

30 **Objectives** Aiming to validate the use of a single post-stimulus sampling protocol for
31 cosyntropin test (SST) in our institution, our primary objectives were: i) to determine the
32 concordance between 30 and 60-min serum cortisol (SC) measurements during SST; ii) to
33 evaluate diagnostic agreement between both sampling times when using classic or assay- and
34 sex-specific SC cut-offs values. Secondary objectives included: i) estimating specificity and
35 positive predictive value of 30 and 60 min sampling times while considering the suspected origin
36 of adrenal insufficiency (AI); iv) to obtain assay-specific cut-offs for SC after SST in a group of
37 subjects with normal hypothalamic-pituitary-adrenal (HPA) axis.

38 **Design and setting** Retrospective chart review study conducted at an Spanish Academic
39 Hospital from 2011 to 2015

40 **Participants and interventions** Two groups were evaluated: i) a main study group including
41 370 patients in whom SC was measured at 30 and 60 minutes during SST; and ii) a confirmative
42 group that included 150 women presenting with a normal HPA axis in whom SST was conducted
43 to rule out late onset congenital adrenal hyperplasia. Diagnostic agreement between both
44 sampling times was assessed by considering both classic (500 nmol/l) and assay-specific SC cut-
45 off concentrations.

46 **Results** Diagnostic agreement between both sampling times was greater when applying sex- and
47 assay-specific cut-off values instead of classic cut-offs. For suspected primary AI, 30-min SC
48 determination was enough to establish diagnosis in over 95% of cases, without missing any
49 necessary treatment. When central AI is suspected, 60 min SC measurement was more specific,
50 establishing diagnosis in over 97% of cases.

1
2
3 51 **Conclusions:** Sex- and assay-specific SC cut-off values improve diagnostic accuracy of SST.
4
5 52 For primary disease, a subnormal SC response at 30 min is a reliable marker of adrenal
6
7 53 dysfunction. On the contrary, when central AI is suspected, 60-min SC measurement improves
8
9 54 diagnostic accuracy of the test.
10
11
12
13 55

14
15 56 **Strengths and limitations:**

- 16
17 57 • We assessed a very large series of well-characterized subjects with a suspicion of adrenal
18
19 58 insufficiency and a minimum clinical follow up of 12 months after the cosyntropin test.
20
21 59 • We used a pre-test distinction between primary and central adrenal insufficiency based on
22
23 60 clinical data.
24
25 61 • We used a local cohort of women with definitely normal cortisol secretion to validate our
26
27 62 findings.
28
29 63 • Our results were not challenged against a biochemical gold-standard and, therefore, false
30
31 64 negative rates, sensitivity, and negative predictive values were not established.
32
33 65 • The confirmatory group was comprised only by premenopausal women, and cosyntropin-
34
35 66 stimulated SC concentrations were only obtained at the 30 min sampling time in these
36
37 67 subjects.
38
39
40
41
42
43
44

45 69 **Keywords:** Adrenal insufficiency; biochemical diagnosis; cosyntropin test; immunoassay;
46
47 70 reference values; sampling times; serum cortisol; specificity.
48
49
50
51
52
53
54
55
56
57
58
59
60

72 Introduction

73
74 The laboratory diagnosis of adrenal insufficiency (AI) at the clinical setting relies on the
75 finding of an inappropriately low morning circulating serum cortisol (SC) or subnormal SC
76 responses to adrenal stimulation [1]. However, the diagnosis of AI diagnosis should not be made
77 according only to laboratory tests, since analytical results must always be interpreted in the
78 context of the whole clinical picture of the individual patient [1–3]. The most widely used
79 adrenal stimulation protocol consists of measuring SC in samples obtained 30 and 60 min after a
80 single 250 µg intravenous bolus or intramuscular injection of tetracosactide (cosyntropin). The
81 normal response consists of a SC value ≥ 500 nmol/l (18 µg/dl) 30 at any time after cosyntropin
82 administration. This protocol, also known as a short standard high-dose test (SST), is the
83 dynamic exploration of choice for primary AI diagnosis [1,3] and it is also used for non-acute
84 central AI [4,5]. In critically ill patients, SST may be performed to rule out a functional form of
85 AI –critical illness-related corticosteroid insufficiency– in subjects showing sustained refractory
86 hypotension and no response to vasopressor drugs [2,6]. Clinical guidelines suggest that this
87 condition may be best diagnosed by a random SC below 276 nmol/l (10 µg/dl) or when the
88 increase in SC after cosyntropin is less than 248 nmol/l (9 µg/dl) [7,8].

89 The issue of which sampling time - 30 min or 60 min – of the SST is the most appropriate
90 is controversial. The 30 min SC measurements have been validated against a “gold standard”
91 such as the insulin tolerance test (ITT) [9]. Hence, some authors [4,10,11] suggest that a single
92 SC measurement 30 min after cosyntropin administration is enough to establish or rule out
93 clinically relevant AI. Other studies show that a 60 min sample may avoid unnecessary
94 overdiagnosis [12–14]. Recent clinical practice guidelines recommend further research to clarify
95 whether 60 min SC might be more specific than 30 min measurements for AI diagnosis [3,15].

1
2
3 96 Even though liquid chromatography/mass spectrometry techniques are currently
4
5 97 recommended for the accurate measurement of circulating steroids, in most centres clinical
6
7 98 routine still relies on automated immunoassays for SC [16]. Considering that the classic cut-off
8
9 99 value for the SST was established for SC as measured by older radioimmunoassays, and that
10
11 100 immunochemiluminescent assays differ in antibody specificity with these earlier assays [17],
12
13 101 establishing local assay-specific cut-off values is of paramount importance to properly classify
14
15 102 SC responses to cosyntropin [3,17,18]. When local validation is not feasible, published assay-
16
17 103 specific cut-off values should be considered [17]. This issue is not inconsequential because,
18
19 104 despite the recommendation of using local assay-specific lower limits of normality (LLN) for the
20
21 105 dynamic assessment of the hypothalamic-pituitary-adrenal (HPA) axis [3], in our experience
22
23 106 many physicians still apply classic cut-off values in their routine practice. Also, other factors that
24
25 107 may influence SC measurement include the stimulation of hepatic synthesis and secretion of
26
27 108 cortisol binding globulin by oestrogens, sex and several non-glucocorticoid drugs [18,19].

28
29
30
31
32
33 109 To provide new insights into these still open questions, and while validating the use of a
34
35 110 single post-stimulus sampling protocol for the routine cosyntropin test (SST) in our institution,
36
37 111 our primary goals were: i) to assess the concordance between 30 and 60 min SC concentrations
38
39 112 after cosyntropin stimulation at the clinical setting; ii) to estimate the diagnostic agreement
40
41 113 between both sampling times when using classic cut-offs derived from the literature or assay-
42
43 114 and sex-specific cut-offs values, taking into account the suspected origin of AI. As secondary
44
45 115 objectives, we aimed to: i) estimate the specificity (Sp) and positive predictive value (PPV) of 30
46
47 116 and 60 min sampling times while taking into account the origin of AI; and ii) confirm assay-
48
49 117 specific LLN for SC concentration after cosyntropin in a group of subjects with a normal HPA
50
51 118 function.
52
53
54
55

56 119

120 **Subjects and methods**

121 We conducted a retrospective chart review study addressing SC responses during SST in two
122 study populations from January 1, 2011 to December 31, 2015 at an academic hospital from
123 Spain:

- 124 i) Main study population: Four hundred fifty one adults in whom SC concentrations at 0, 30
125 and 60 min during a SST conducted at the clinical setting for suspected AI.
- 126 ii) Confirmative group: One hundred fifty three women with normal HPA axis recruited from
127 our Reproductive Endocrinology clinic during the study of functional hyperandrogenism in
128 whom SC concentrations were obtained at 0 and 30 min during a SST performed for the
129 routine screening of non-classic congenital adrenal hyperplasia (NCAH). NCAH had been
130 ruled out in all those women because cosyntropin-stimulated 17-hydroxyprogesterone and
131 11-deoxycortisol concentrations were below 10 ng/ml and 21 ng/ml, respectively [20]. None
132 of the women in the confirmative group was using combined contraceptives or any other
133 hormonal therapy at the time of sampling.

134 Before conducting the study, we obtained approval from the local ethics committee. All
135 women from our Reproductive Endocrinology clinic had previously signed an informed consent
136 form for the inclusion of a selection of coded clinical variables in an electronic database for
137 clinical research purposes that included the SC measurements presented here.

138

139 *Main study population*

140 Basal and stimulated SC values were extracted from the electronic database of our
141 Department of Clinical Biochemistry. We collected a minimum dataset in an electronic case
142 form from the clinical records of the patients including age, sex, weight, height, laboratory
143 measurements at the dates when the SST was conducted such as circulating electrolytes,

1
2
3 144 glomerular filtration rate and basal ACTH concentrations at the time of SST, clinical suspicion
4
5 145 of primary or central AI, other dynamic tests performed for the evaluation of adrenal function,
6
7 146 history of pituitary disease, time from hypothalamic-pituitary insult to SC determination,
8
9
10 147 administration of drugs that may interfere with the HPA axis, time of follow-up, and the
11
12 148 immunoassay used for SC assay. Baseline characteristics of study population are shown in Table
13
14
15 149 1.

16
17 150 We considered a clinical suspicion of potential primary AI in cases when the patients were
18
19 151 known to have adrenal disease, had required mineralocorticoid supplementation during follow-
20
21 152 up, had received drugs that may interfere with cortisol biosynthesis, had not clinical suspicion of
22
23 153 any hypothalamic-pituitary condition, and had not developed such a condition later in time.
24
25
26 154 Conversely, we suspected a potential central AI in subjects known to suffer from hypothalamic-
27
28 155 pituitary disease, had received drugs that may suppress the HPA axis, or when her/his referring
29
30 156 physician reported a clinical suspicion of central AI in the clinical record. According to their
31
32
33 157 clinical records, all patients included here had a minimum 12-month follow-up after obtaining
34
35 158 the SST at any outpatient or in-patient facility of our centre. We actively reviewed these records
36
37 159 looking for any latter diagnosis of AI.

38
39
40 160 We excluded from analysis: i) seven subjects submitted to dynamic tests other than SST
41
42 161 such as the insulin tolerance test (n = 2), corticotrophin-releasing hormone test (n = 2), oral
43
44 162 glucose tolerance test (n = 2) and glucagon stimulation test (n = 1); ii) thirty six subjects aged
45
46
47 163 below 18 years; iii) twenty subjects with a follow-up shorter than 12 months; iv) twelve subjects
48
49 164 in whom critically-ill related AI was suspected; and v) six subjects from whom we could not
50
51 165 obtain enough information from their clinical records as to explain the reason for conducting a
52
53 166 SST. Therefore, the study group finally included in the analyses consisted of 370 subjects.
54
55

56 167
57
58
59
60

1
2
3 168 *Confirmative group*
4

5 169 The results of SST from 153 premenopausal women with a normal HPA axis aged from 14 to
6
7 170 42 years old were included. Three women who showed a clearly subnormal SC response were
8
9
10 171 excluded from the analysis. In two of these women the suppressive effect on the HPA axis of the
11
12 172 progestins administered during 10 days before the SST with the aim of inducing a withdrawal
13
14 173 vaginal bleeding could justify the abnormal results; in the other case, we could not establish the
15
16
17 174 cause of the subnormal response with certainty because the patient was lost to follow-up.
18

19 175

20
21 176 *Assays*
22

23
24 177 During the study period, two immunoassays were used in our centre: i) from 2011 to July 1,
25
26 178 2013 the *Siemens Immulite 2000*[®] *Cortisol Immunoassay System* (immunoassay 1) was used and
27
28 179 had 6.0% and 7.8% intra- and inter-assay coefficient of variation (CV) respectively; and ii) from
29
30
31 180 2013 August 1, 2013 to December 31, 2015, the *Abbot Laboratories Diagnostics Division*
32
33 181 *Architect*[®] *Cortisol Immunoassay System* (immunoassay 2) was used, showing 3.2% and 3.4%
34
35 182 intra- and inter-assay CVs, respectively. Plasma ACTH concentrations were measured by the
36
37 183 *Siemens Immulite 2000*[®] *ACTH Immunoassay System* with an analytical sensitivity of 1.1 pmol/l,
38
39
40 184 and intra- and interassay CVs below 10%. The upper limit of normality for healthy subjects was
41
42 185 10 pmol/l.
43

44 186

45
46
47 187 *Analysis of the agreement between the 30 and 60 min sampling times*
48

49 188 We analysed diagnostic agreement between the 30 and 60 min SC in patients of the main
50
51 189 study population – in the confirmation subgroup the 60 min measurement was not obtained –
52
53
54 190 considering two different LLN for cosyntropin-stimulated SC: i) the classic ≥ 500 nmol/l (3), and
55
56 191 ii) sex- and assay-specific cut-off values derived from the estimated lower reference limit for the
57
58
59

1
2
3 192 SC response at 30-min to cosyntropin, taking also into account the concurrent use by 7 women of
4
5 193 combined oral contraceptives (COC) [18]. For immunoassay 1, the reported LLN (2.5th
6
7 194 percentile) was 470 nmol/l (17 µg/dl) in men and women, and 690 nmol/l (25 µg/dl) for women
8
9 195 taking COC. For immunoassay 2, the LLNs were 441 nmol/l (16 µg/dl) for men, 414 nmol/l (15
10
11 196 µg/dl) for women, and 579 nmol/l (21 µg/dl) for women taking COC [17].
12
13
14
15
16

17 198 *Statistical analysis*

19 199 Data are shown as mean ± standard deviation or 95% confidence interval (CI), median
20
21 200 (minimum-maximum), and raw numbers (percentage) as appropriate. The normal distribution of
22
23 201 continuous variables was assessed by the Kolmogorov-Smirnov test for one sample after a two-
24
25 202 step approach for transforming skewed variables if necessary [21]. Comparisons among
26
27 203 continuous variables were performed by repeated-measures ANOVA. Comparisons among
28
29 204 categorical variables were performed by Fisher's exact or χ^2 tests as appropriate. Pearson's
30
31 205 analysis served to correlate SC at 30 and 60 min samples. Consistency and absolute agreement
32
33 206 among both point times of SST were determined by their intra-class correlation coefficient (ICC)
34
35 207 with a two-factor and random-effect model. Quantitative agreement was graphically assessed by
36
37 208 Bland-Altman plots. Biochemical agreement in the diagnosis of normal or subnormal adrenal
38
39 209 was assessed by using the kappa (κ) coefficient. True positives (TP) were defined as SSTs
40
41 210 showing subnormal cortisol responses at both time points in patients who required adrenal
42
43 211 replacement therapy. True negatives (TN) were defined as SSTs showing a normal cortisol
44
45 212 response at both time points in patients who did not need glucocorticoid replacement during their
46
47 213 follow-up, did not suffer an adrenal crisis, and, when submitted to other dynamic HPA test,
48
49 214 showed normal responses. False positives (FP) for one of the sampling times consisted of the
50
51
52
53
54
55
56
57
58
59
60

215 finding of a subnormal response in one of the sampling times but not in the other. We calculated
216 Sp and PPV [$Sp = TN / (TN + FP)$ and $PPV = TP / (TP + FP)$] for each SC sampling times
217 during the SST. A P value < 0.05 was considered statistically significant.

218

219 Results

220 Main study population

221 Of 370 SSTs including 30 and 60 min sampling times, SC was assayed by immunoassay
222 1 in 227 cases and by immunoassay 2 in the remaining 143 tests. Basal and cosyntropin-
223 stimulated SC concentrations, ACTH levels when available, and the median duration of follow-
224 up in patients with either normal or insufficient responses are shown in **Table 2**.

225 SC concentrations when patients in the main study group were analyzed as a whole are
226 represented in **Figure 1A**. SC concentrations at 30 and 60 min during the SST increased when
227 compared to baseline values (**Figure 1, panel A**), and showed a very strong linear correlation
228 (**Figure 1, panel B**). Baseline SC concentrations correlated with 30 min SC measurements ($r =$
229 0.735 , $P = 0.001$), and with 60 min SC values ($r = 0.660$, $P = 0.001$).

230 Similar results were observed when analyzing separately the 150 SSTs performed with
231 the aim of ruling out primary AI (correlation between baseline SC and 30 min SC: $r = 0.720$, $P =$
232 0.001), and correlation between baseline SC and 60 min SC: $r = 0.640$, $P = 0.001$) and the 220
233 SSTs conducted to exclude central AI (correlation between baseline SC and 30 min SC: $r =$
234 0.723 , $P = 0.001$, and correlation between baseline SC and 60 min SC: $r = 0.644$ ($P = 0.001$).

235 The ICC among SC concentrations as assayed at both sampling times showed a very
236 good consistence index (0.940; 95%CI: 0.928 – 0.952) and a good absolute agreement (0.889,
237 95%CI: 0.465 – 0.957), even though the latter only qualifies as fair according to the lower limit
238 of the 95%CI. The Bland-Altman plot (**Figure 1, panel C**) showed a good agreement between

239 SC assayed at 30 and 60 min, with a slight tendency towards greater percentage differences with
240 decreasing mean values of stimulated SC.

241 **Figure 2 and Table 2** show SC concentrations as a function of the clinical suspicion and
242 whether or not the result of the SST was normal. The diagnostic agreement among both sampling
243 times according to classic and to sex – and assay-specific cut-off values is shown in **Figure 3**.
244 Disagreements between both sampling times were as follows. When relying on the classic SC
245 cut-off point (≥ 500 nmol/l), 39 cases (10.5%) had a subnormal response at 30 min that reached
246 normal values at 60 min whilst, in 3 patients (0.8%), a normal response at 30 min ended being
247 subnormal at 60 min. Using sex- and assay-specific values, 34 cases (9.2%) showed subnormal
248 responses at 30 min but normal SC concentrations at 60 min, whereas in 5 cases (1.3%), the
249 response was normal at 30 min but subnormal at 60 min.

250 The analysis of the diagnostic agreement as a function of the suspicion of primary versus
251 central AI is shown in **Figure 4**. As a rule, agreement among both sampling times of the SST
252 was better when primary AI was suspected compared with a suspicion of central AI. When using
253 classic cut-off values to rule out primary AI, 7 cases (4.7%) showed a subnormal response at 30
254 min that reached normal concentrations at 60 min, whereas no subject with a normal response at
255 30 min had a subnormal response at 60 min. Using sex- and assay-specific cut-off values, in 6
256 cases (4.0%) the response was subnormal at 30 min but reached normal concentrations at 60 min.
257 Four of them showed a subnormal SC responses to cosyntropin that were very close to the cut-
258 off value. In these subjects, the differences between the cut-off value and the stimulated SC
259 ranged from 22 to 39 nmol/l (0.8 to 1.4 $\mu\text{g}/\text{dl}$), very small concentrations that are, in fact,
260 included within the CV of the assays, thereby suggesting no clinical relevance. The two
261 remaining patients showed peak SC concentrations of 320 and 364 nmol/l (11,6 and 13,2 $\mu\text{g}/\text{dl}$) at
262 the 30 min sampling time: one had received oral glucocorticoid replacement therapy that did not

1
2
3 263 preclude the patient of responding to cosyntropin by showing a SC of 470 nmol/l (17 µg/dl) at
4
5 264 the 60 min sample, and the other subject was submitted to SST because of the presence of
6
7 265 bilateral adrenal hyperplasia and did not show any signs or symptoms of AI nor suffered an
8
9
10 266 adrenal crisis during follow-up. None of the SSTs showing normal responses at 30 min had a
11
12 267 subnormal response at 60 min.

14 268 When central AI was suspected and the classic cut-off point was applied, 32 cases
15
16 269 (14.5%) had a normal SC response at 60 min but a subnormal SC value at 30 min. Only 3
17
18 270 subjects (1.4%) presented with the opposite situation. Using sex- and assay-specific cut-off
19
20 271 concentrations, 28 cases (12.7%) showed a normal response at 60 min but a subnormal result at
21
22 272 30 min, yet in only 5 cases (2.3%) the contrary occurred. These 5 subjects had been evaluated in
23
24 273 the context of withdrawal of prolonged glucocorticoid therapy during the first year after a
25
26 274 pituitary insult (surgery and/or pituitary radiotherapy). Three of them showed a complete
27
28 275 recovery of their HPA axis throughout the follow-up period, whereas in the other two patients,
29
30 276 who had received pituitary radiotherapy, the subnormal response to cosyntropin was maintained
31
32 277 over time.

33 278 The Sp and PPV for different sampling times and cut-off values used here are shown in
34
35 279 Table 3. SC concentrations at 60 min had a higher Sp and PPV compared with 30 min
36
37 280 measurements, particularly when central AI was suspected. Nonetheless, the Sp of the
38
39 281 determination at 30 min was as high as 95% when SST had been performed to rule out primary
40
41 282 disease both when applying classic or sex- and assay-specific cut-off values.

42 283 We observed discordant results between classic and sex-and assay-specific cut-off
43
44 284 concentrations in 50 cases. In 47 of these subjects, a subnormal response using the classic cut-off
45
46 285 value turned into a normal response had sex- and assay-specific cut-offs been used. In 7 of them,
47
48 286 SST was performed to rule out primary AI and in the remaining 40 subjects the SSTs were

1
2
3 287 conducted to rule out central AI. Glucocorticoid replacement was started in 18 cases, and no
4
5 288 subject presented with signs or symptoms of chronic or acute AI. In addition, from the 50
6
7 289 discordant SSTs, 3 were conducted in women under estrogenic therapy and presented a normal
8
9
10 290 response according to the classic cut-off value, but subnormal when considering sex- and assay-
11
12 291 specific cut-offs, yet none of them required glucocorticoid therapy.
13

14 292

17 293 *Confirmative group*

18
19 294 Thirty (20%) of these women presented with a subnormal response to SST according to
20
21 295 classic cut-off values, yet this figure was reduced to only 3 (2%) when sex- and assay-specific
22
23 296 cut-off values were used [observed agreement: 82%; κ : 0.151 (95%CI: 0.066-0.235)]. The three
24
25 297 women showing a subnormal response during SST using a sex- and assay-specific cut-off value
26
27 298 showed stimulated SC concentrations of 342 nmol/l (12.4 μ g/dl), 353 nmol/l (12.8 μ g/dl) and
28
29 299 372 nmol/l (13.5 μ g/dl), whereas the LLNs (2.5th percentile) of SC concentrations at 30 min
30
31 300 sampling time of SST were 436 nmol/l (15.8 μ g/dl) and 411 nmol/l (14.9 μ g/dl) for
32
33 301 immunoassays 1 and 2, respectively. The 5th percentiles for both immunoassays were 450 nmol/l
34
35 302 (16.3 μ g/dl) and 414 nmol/l (15.0 μ g/dl), respectively, showing minimal differences (~10%) with
36
37 303 the LLNs previously described (**Figure 5**). None of these female controls developed any HPA
38
39 304 disease during their follow-up.
40
41

42
43
44 305 We performed a sensitivity analysis of the results in the main study population, after
45
46 306 excluding women taking oral contraceptive therapy, using the LLNs derived from the women
47
48 307 with a normal HPA axis that composed our confirmatory group. Both sampling times showed a
49
50 308 similar agreement than that observed earlier when using LLNs derived from the literature
51
52 309 [observed agreement: 92%; κ : 0.724 (95%CI: 0.632-0.816)]. In the whole group of subjects, 4
53
54 310 out of 286 individuals (1.4%) with a normal response at 30-min sampling time showed a
55
56
57
58
59

1
2
3 311 subnormal response at 60-min. Conversely, 26 out of 77 subjects (34%) with a subnormal
4
5 312 response at 30-min had a normal response at 60-min. Then, we analyzed those data as a function
6
7 313 of the suspected reason for screening AI. Supporting our previous findings, agreement among
8
9 314 both SST sampling times was better when primary AI was suspected [observed agreement: 97%;
10
11 315 κ : 0.846 (95%CI: 0.714-0.977)] compared with a suspicion of central AI [observed agreement:
12
13 316 89%; κ : 0.667 (95%CI: 0.548-0.785)], data being almost the same observed in Figure 4.
14
15
16
17 317

18 19 318 **DISCUSSION**

20
21 319 AI is a clinical condition associated with a high morbidity and mortality. Unstimulated
22
23 320 early morning SC values below 138 nmol/l (5 μ g/dl) show a high PPV for AI, whereas
24
25 321 concentrations over 500 nmol/l predict a normal adrenal response. However, values between 138
26
27 322 and 500 nmol/l are considered indeterminate and require adrenal stimulation to confirm or rule
28
29 323 out a diagnosis, always in consonance with the clinical picture [1–3].
30
31
32

33 324 Baseline SC concentrations showed stronger linear correlations with cosyntropin-
34
35 325 stimulated SC levels at 30 and 60 min samples of the SST, in agreement with previous reports
36
37 326 [22]. Our data also show that both 30 and 60 min SC measurements during a SST have an
38
39 327 adequate index of consistency, but the same is not true in terms of absolute agreement,
40
41 328 particularly when a central AI is suspected. Furthermore, a single determination at 60 min during
42
43 329 the SST appears to have the higher Sp and PPV for the diagnosis of subjects presenting with
44
45 330 either primary or central AI. In consonance, after evaluating retrospectively 73 subjects, Zueger
46
47 331 et al. [23] reported that sampling at 30 min of the SST did not provide any additional diagnostic
48
49 332 advantage over performing a single determination at 60 min of the test. Although similar results
50
51 333 have been also reported by others [13,14], these studies did not take into account the primary or
52
53 334 central origin of AI and did not apply sex- and assay specific cut-off values, a fact of paramount
54
55
56
57
58
59

1
2
3 335 importance because of the considerable influence that cortisol immunoassays exerts on the final
4
5 336 values observed after cosyntropin-stimulation [17,18].
6

7 337 Our results also indicate that SC measurement at 30 min during the SST, when using sex-
8
9 338 and assay-specific cut-off values, are enough to rule out clinically relevant primary AI since only
10
11 339 4% of patients in this particular situation showed a subnormal response at 30 min followed by
12
13 340 normal response at 60 min. Furthermore, these subjects presented with stimulated SC
14
15 341 concentrations which were very close to the cut-off concentrations, to the extent that the
16
17 342 differences with these normal limits may be explained by the analytical variability of these
18
19 343 commercial immunoassays used here. Even more important from a clinical point of view, none
20
21 344 of these subjects required replacement therapy during their follow-up, suffered an acute adrenal
22
23 345 crisis, nor were diagnosed with any adrenal condition during follow-up, strongly suggesting that
24
25 346 their HPA function was actually normal at the time the SST was performed. The use of sex- and
26
27 347 assay-specific cut-off values appears to be essential, since other authors have suggested that
28
29 348 some healthy individual may have a delayed response to SST using classic reference values [24].
30
31
32
33
34

35 349 On the other hand, 60 min samples appear to be more specific than 30 min measurements
36
37 350 when central AI is suspected. In such a case, 12.7% of the subjects presenting with a subnormal
38
39 351 response at 30 min actually had a normal response at 60 min, avoiding unnecessary treatments in
40
41 352 them. Although a subnormal response 30 min after cosyntropin-stimulation in patients with
42
43 353 suspicion of secondary AI may not translate into the need of adrenal replacement in a non-
44
45 354 critical scenario, it is likely that most physicians would feel more confident with not starting
46
47 355 replacement therapy after obtaining a cosyntropin-stimulated SC concentration above the LLN,
48
49 356 favoring the use of 60 min samples over 20 min determinations for this particular reason.
50
51 357 Furthermore, relying mostly on 60 min SC responses to cosyntropin when suspecting a central
52
53 358 origin of AI is also supported by the fact that, in 2 out of the 5 patients in our series who showed
54
55
56
57
58
59

1
2
3 359 a subnormal response at 60 min preceded by normal SC values at 30 min, AI was actually
4
5 360 confirmed during follow-up because of former pituitary radiotherapy.
6

7 361 Our present findings also reinforce the need of sex- and assay-specific cut-off values to
8
9 362 interpret the results of the SST, in agreement with recent clinical guidelines[3]. The use of such
10
11 363 cut-off values lead in our study to a reduction in FP results, higher Sp and PPV, less discordant
12
13 364 results among sampling times of the SST, and fewer unnecessary treatments [20 patients (5%)
14
15 365 could have been treated unnecessarily if classic cut-off values were applied for diagnosis]. The
16
17 366 reliability of sex- and assay-specific cut-off values was confirmed in our population of
18
19 367 premenopausal women with normal HPA axis, in whom these cut-offs were more appropriate
20
21 368 than relying on classic values to assess the functionality of their HPA axis. In this population, the
22
23 369 LLNs for stimulated SC at 30 min were very close to those reported for each immunoassay by
24
25 370 the manufacturers, which relied on the 2.5th percentile [17], yet reinforcing the need to establish
26
27 371 local normative data in order to improve the diagnostic accuracy of cortisol measurements during
28
29 372 SSTs [17,25].
30
31
32
33
34

35 373 Among the strengths of our study, we would highlight the large series of subjects
36
37 374 suspected of suffering AI who were evaluated with a standardized dynamic study, and the careful
38
39 375 review of subjects' medical records that followed such evaluations. However, we are aware of
40
41 376 several weaknesses derived from the observational and retrospective design of the study, making
42
43 377 impossible to rule out information bias. Our best efforts might have not been enough to avoid
44
45 378 misclassification of patients according to the suspicion of primary or central AI. Also, the
46
47 379 administration of supraphysiological doses of cosyntropin does not permit ruling out partial
48
49 380 deficiencies either, particularly in those suspected of central HPA defects. Another limitation is
50
51 381 that published assay-specific normative value used in our study derived from SC sampling at 30-
52
53 382 min [17]. Thus, the possibility exists that SC sampling at 60-min may require its own normative
54
55
56
57
58
59
60

1
2
3 383 cut-off. Also, and even considering the large sample of subjects included in our study, our
4
5 384 present results may not be extrapolable to other populations in whom SC has been measured with
6
7 385 different immunoassays that would require specific local normative data. Moreover, analysis of
8
9 386 Sp and PPV has not been challenged against a biochemical gold-standard in most cases and, as a
10
11 387 consequence, we have not been able to establish false negative rates, sensitivity and negative
12
13 388 predictive values. Nonetheless, besides those assessments had been unethical in most cases, the
14
15 389 lack of a laboratory gold-standard such as an ITT did not override our results, since from a
16
17 390 practical point of view, we are looking for patients needing replacement therapy and not for
18
19 391 those with a partial AI who do not require any treatment. Another limitation was that the
20
21 392 confirmation group is not fully representative of our main study population since was only
22
23 393 comprised of premenopausal women and stimulated SC was only available at the 30 min
24
25 394 sampling time. Lastly, we could not rule out entirely pre-treatment with progestogens in the
26
27 395 context of induction of withdrawal bleeding in our confirmative population. Because these drugs
28
29 396 might exert a mild suppressive effect on the HPA axis [19,26], their administration in a few cases
30
31 397 could have, at least in theory, lowered stimulated SC values, precluding the generation of local
32
33 398 normative data from their results. Instead, we had to rely on published assay-specific cut-off
34
35 399 values for this reason.
36
37
38
39
40
41
42
43

400

401 CONCLUSIONS

402 To assist clinical judgement, and compared with the use of classic cut-off values derived
403 from the literature, application of sex- and assay-specific cut-off values of SC responses to
404 cosyntropin results into higher Sp and PPV for establishing a diagnosis of AI, thereby avoiding
405 unnecessary treatments. Measurement of stimulated SC at 30 min after cosyntropin-stimulation

406 may suffice for supporting a clinical diagnosis of primary AI, yet 60 min measurements might be
407 preferable when central AI is suspected.

408

409 **Conflict of interest:** None.

410

411 **Funding:** This work has been supported by grant PI1400649) from Fondo de Investigación
412 Sanitaria, Instituto de Salud Carlos III, Spanish Ministry of Economy and Competitiveness.
413 M.L.-R. has a local grant for clinical research from the Instituto Ramón y Cajal de Investigación
414 Sanitaria (IRYCIS). CIBERDEM is also an initiative of Instituto de Salud Carlos III, partially
415 supported by Fondo Europeo de Desarrollo Regional FEDER. There were no other sources of
416 funding.

417

418 **Data sharing statement:** Individual participant data that underlie the results reported in this
419 article, after deidentification, so as the study protocol would be available immediately after
420 publication to anyone who wishes to access the data to achieve aims in the approved proposal
421 and for individual participant data meta-analysis. Proposals should be directed to
422 andres_ortiz_f@yahoo.com or to manuel.luque@salud.madrid.org. To gain access, data
423 requestors will need to sign a data access agreement.

424

425 **Authorship statement:** A.O.-F. y M.L.-R. designed the protocol and performed the statistical
426 analysis. A.O.-F. y E.S.-C. reviewed the clinical data using the electronic or written records if
427 necessary. A.G.-C. and L.J.-M. performed the electronic search of serum cortisol samples. A.O.-

1
2
3 428 F. y M.L.-R. wrote the first draft of the study. All the authors, including L.N.-C. and H.F.E.-M,
4
5 429 reviewed the manuscript before its submission and contributed to intellectual content. All the
6
7 430 authors have accepted responsibility for the entire content of the manuscript and approved the
8
9
10 431 final submission.
11
12

13 432
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

433 **REFERENCES**

- 434 1. Charmandari E, Nicolaidis NC, Chrousos GP. Adrenal insufficiency. *The Lancet*.
435 2014;383(9935):2152-67.
- 436 2. Neary N, Nieman L. Adrenal Insufficiency- etiology, diagnosis and treatment. *Curr Opin*
437 *Endocrinol Diabetes Obes*. 2010;17(3):217-23.
- 438 3. Bornstein SR, Allolio B, Arlt W, Barthel A, Don-Wauchope A, Hammer GD, et al.
439 *Diagnosis and Treatment of Primary Adrenal Insufficiency: An Endocrine Society Clinical*
440 *Practice Guideline*. *J Clin Endocrinol Metab*. 2016;101(2):364-89.
- 441 4. Grossman AB. The Diagnosis and Management of Central Hypoadrenalism. *J Clin*
442 *Endocrinol Metab*. 2010;95(11):4855-63.
- 443 5. Agha A, Tomlinson JW, Clark PM, Holder G, Stewart PM. The Long-Term Predictive
444 *Accuracy of the Short Synacthen (Corticotropin) Stimulation Test for Assessment of the*
445 *Hypothalamic-Pituitary-Adrenal Axis*. *J Clin Endocrinol Metab*. 2006;91(1):43-7.
- 446 6. Sprung CL, Annane D, Keh D, Moreno R, Singer M, Freivogel K, et al. Hydrocortisone
447 *Therapy for Patients with Septic Shock*. *N Engl J Med*. 2008;358(2):111-24.
- 448 7 Marik PE, Pastores SM, Annane D, *et al*. Recommendations for the diagnosis and
449 *management of corticosteroid insufficiency in critically ill adult patients: consensus*
450 *statements from an international task force by the American College of Critical Care*
451 *Medicine*. *Crit Care Med* 2008;36:1937-49.
- 452 8. Annane D, Sébille V, Troché G, Raphaël J-C, Gajdos P, Bellissant E. A 3-Level Prognostic
453 *Classification in Septic Shock Based on Cortisol Levels and Cortisol Response to*
454 *Corticotropin*. *JAMA*. 2000;283(8):1038-45.

- 1
2
3 455 9. Lindholm J, Kehlet H. Re-Evaluation of the Clinical Value of the 30 Min Acth Test in
4
5 456 Assessing the Hypothalamic-Pituitary-Adrenocortical Function. *Clin Endocrinol (Oxf)*.
6
7 457 1987;26(1):53-9.
8
9
10 458 10. Hurel SJ, Thompson CJ, Watson MJ, Harris MM, Baylis PH, Kendall-Taylor P. The short
11
12 459 Synacthen and insulin stress tests in the assessment of the hypothalamic-pituitary-adrenal
13
14 460 axis. *Clin Endocrinol (Oxf)*. 1996;44(2):141-6.
15
16
17 461 11. Gonzálbez J, Villabona C, Ramón J, Navarro MA, Giménez O, Ricart W, et al.
18
19 462 Establishment of reference values for standard dose short synacthen test (250 µg), low dose
20
21 463 short synacthen test (1 µg) and insulin tolerance test for assessment of the hypothalamo-
22
23 464 pituitary-adrenal axis in normal subjects. *Clin Endocrinol (Oxf)*. 2000;53(2):199-204.
24
25
26 465 12. Longui CA, Vottero A, Harris AG, Chrousos GP. Plasma cortisol responses after
27
28 466 intramuscular corticotropin 1-24 in healthy men. *Metab - Clin Exp*. 1998;47(11):1419-22.
29
30
31 467 13. Mansoor S, Islam N, Siddiqui I, Jabbar A. Sixty-minute post-Synacthen serum cortisol level:
32
33 468 a reliable and cost-effective screening test for excluding adrenal insufficiency compared to
34
35 469 the conventional short Synacthen test. *Singapore Med J*. 2007;48(6):519-23.
36
37
38 470 14. Chitale A, Musonda P, McGregor AM, Dhatariya KK. Determining the utility of the 60 min
39
40 471 cortisol measurement in the short synacthen test. *Clin Endocrinol (Oxf)*. 2013;79(1):14-9.
41
42 472 15. de Miguel Novoa P, Vela ET, García NP, Rodríguez MM, Guerras IS, Martínez de Salinas
43
44 473 Santamaría Mde L, et al. Guidelines for the diagnosis and treatment of adrenal insufficiency
45
46 474 in the adult. *Endocrinol Nutr Organo Soc Espanola Endocrinol Nutr*. 2014;61 (Suppl 1):1-35.
47
48
49 475 16. Turpeinen U, Hämäläinen E. Determination of cortisol in serum, saliva and urine. *Best Pract*
50
51 476 *Res Clin Endocrinol Metab*. 2013;27(6):795-801.
52
53
54 477 17. El-Farhan N, Pickett A, Ducroq D, Bailey C, Mitchem K, Morgan N, et al. Method-specific
55
56 478 serum cortisol responses to the adrenocorticotrophin test: comparison of gas
57
58
59
60

- 1
2
3 479 chromatography-mass spectrometry and five automated immunoassays. *Clin Endocrinol*
4
5 480 (Oxf). 2013;78(5):673-80.
6
7
8 481 18. Klose M, Lange M, Rasmussen AK, Skakkebaek NE, Hilsted L, Haug E, et al. Factors
9
10 482 Influencing the Adrenocorticotropin Test: Role of Contemporary Cortisol Assays, Body
11
12 483 Composition, and Oral Contraceptive Agents. *J Clin Endocrinol Metab*. 2007;92(4):1326-33.
13
14 484 19. Kontula K, Paavonen T, Luukkainen T, Andersson LC. Binding of progestins to the
15
16 485 glucocorticoid receptor: Correlation to their glucocorticoid-like effects on in vitro functions
17
18 486 of human mononuclear leukocytes. *Biochem Pharmacol*. 1983;32(9):1511-8.
19
20
21 487 20. Escobar-Morreale HF, Sanchón R, San Millán JL. A prospective study of the prevalence of
22
23 488 nonclassical congenital adrenal hyperplasia among women presenting with hyperandrogenic
24
25 489 symptoms and signs. *J Clin Endocrinol Metab*. 2008;93(2):527-33.
26
27
28 490 21. Templeton G. A Two-Step Approach for Transforming Continuous Variables to Normal:
29
30 491 Implications and Recommendations for IS Research. *Commun Assoc Inf Syst*. 2011;28(1).
31
32 492 Available at: <http://aisel.aisnet.org/cais/vol28/iss1/4>
33
34
35 493 22 Dorin RI, Qiao ZG, Bouchonville M, et al. Characterization of Cortisol Secretion Rate in
36
37 494 Secondary Adrenal Insufficiency. *J Endocr Soc* 2017;1:945-56.
38
39
40 495 23. Zueger T, Jordi M, Laimer M, et al. Utility of 30 and 60 minute cortisol samples after high-
41
42 496 dose synthetic ACTH-1-24 injection in the diagnosis of adrenal insufficiency. *Swiss Med*
43
44 497 *Wkly* 2014;144:w13987.
45
46
47 498 24. Karaca Z, Lale A, Tanriverdi F, et al. The comparison of low and standard dose ACTH and
48
49 499 glucagon stimulation tests in the evaluation of hypothalamo-pituitary-adrenal axis in healthy
50
51 500 adults. *Pituitary* 2011;14:134-40.
52
53
54
55
56
57
58
59

- 1
2
3 501 25. Clark PM, Neylon I, Raggatt PR, *et al.* Defining the normal cortisol response to the short
4
5 502 Synacthen test: implications for the investigation of hypothalamic-pituitary disorders. *Clin*
6
7 503 *Endocrinol (Oxf)* 1998;49:287–92.
8
9
10 504 26. Bornstein SR. Predisposing Factors for Adrenal Insufficiency. *N Engl J Med.*
11
12 505 2009;360(22):2328-39.
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

TABLE 1. Baseline characteristics of the main study population as a function of the clinical suspicion of primary or central adrenal disease.

	Clinical suspicion of primary AI				Clinical suspicion of central AI			
	(n = 150)		(n = 220)		(n = 220)		(n = 220)	
	Assay 1		Assay 2		Assay 1		Assay 2	
Sex	Women (n = 70)	Men (n = 36)	Women (n = 28)	Men (n = 16)	Women (n = 75)	Men (n = 46)	Women (n = 64)	Men (n = 35)
Age (years)	52 ± 19	58 ± 14	55 ± 18	51 ± 14	54 ± 14	57 ± 13	56 ± 18	54 ± 13
Weight(kg)	59 ± 13	69 ± 14	59 ± 9	78 ± 14	73 ± 14	84 ± 12	72 ± 13	83 ± 20
BMI (kg/m ²)	24 ± 5	24 ± 4	23 ± 4	26 ± 5	29 ± 6	29 ± 3	28 ± 5	29 ± 6
Na (mmol/l)	138 ± 3	137 ± 5	138 ± 4	138 ± 4	139 ± 2	139 ± 4	140 ± 2	140 ± 3
K (mmol/l)	4.3 ± 0.6	4.5 ± 0.8	4.1 ± 0.5	4.2 ± 0.3	4 ± 0.3	4.1 ± 0.4	4.2 ± 0.3	4.2 ± 0.4
Ca (mmol/l)	2.4 ± 0.1	2.3 ± 0.2	2.3 ± 0.1	2.4 ± 0.1	2.4 ± 0.1	2.3 ± 0.1	2.4 ± 0.1	2.4 ± 0.1
Cr (μmol/l)	62 (44 – 1114)	80 (44 – 1158)	71 (53 – 230)	80 (62 - 115)	62 (44 – 875)	71 (53 -150)	71 (18 – 97)	71 (44 – 141)
eGFR (MDRD) (ml/min/1.73m ²)	88 (4 – 137)	80 (4 – 183)	77 (20 – 110)	98 (57 – 125)	90 (5 - 144)	95 (43 – 154)	81 (48 – 361)	91 (44 – 163)
ACTH (pmol/l)	3 (1 – 16)	5 (1 - 21)	4 (1 – 25)	6 (1 – 230)	3 (1 – 28)	4 (1 – 17)	4 (1 – 43)	5 (1 – 19)

1
2
3 *Abbreviations, BMI, body mass index; Ca, total serum calcium; Cr, serum creatinine; eGFR, estimated glomerular filtration rate; K, serum*
4 *potassium; Na, serum sodium.* Data are presented as mean \pm SD or median (minimum-maximum) as appropriate.
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

For peer review only

TABLE 2. Basal and cosyntropin-stimulated serum cortisol concentrations as a function of the presence of a normal or abnormal result during the SST, and mean follow-up of the patients in each subgroup. From the whole sample, ACTH measurements were available for 342 samples.

	Normal responses at both times (n = 307)	Confirmed primary AI (n = 18)	Confirmed secondary AI (n = 45)
Basal ACTH (pmol/l)	4 (1 – 43)	6 (1 – 71) *	3 (1 - 11)
Basal SC (nmol/l)	386 ± 166	165 ± 110	138 ± 83
SC at 30 min (nmol/l)	662 ± 193	248 ± 110	276 ± 110
SC at 60 min (nmol/l)	745 ± 221	304 ± 138	304 ± 110
Follow-up (months)	37 ± 17	43 ± 18	36 ± 15

Data are presented as mean ± SD or median (minimum-maximum) as appropriate. To convert SC to metric units, multiply nmol/l by 0.03625 (result in µg/dl). To convert ACTH to metric units, multiply pmol/l by 4.54545 (result in pg/ml). * Despite not having any hypothalamic-pituitary condition at diagnosis or throughout their follow-up, and not having received drugs that suppress the HPA axis, seven patients with clinical suspicion of primary disease who required replacement therapy presented with normal ACTH levels. Three of them had begun glucocorticoid therapy at the time of SST. In another 4 cases, there is a strong suspicion of that was the case, although the possibility of an inadequate sample processing also existed (i.e.: sample transport at room temperature). *Abbreviations, ACTH, adrenocorticotropin hormone; AI, adrenal insufficiency; SC, serum cortisol.*

TABLE 3. Specificity and positive predictive value (PPV) of the short high-dose cosyntropin test, for the diagnosis of adrenal insufficiency (AI), according to serum cortisol cut-off concentrations (classic and sex- and assay-specific), and as a function of the suspected origin of the disease.

	Classic cut-off values						Sex- and assay-specific cut-off values					
	Global		Clinical suspicion Primary AI		Clinical suspicion Central AI		Global		Clinical suspicion Primary AI		Clinical suspicion Central AI	
	30	60	30	60	30	60	30	60	30	60	30	60
Sampling time (min)	30	60	30	60	30	60	30	60	30	60	30	60
Specificity (%)	86	99	95	100	79	98	89	98	96	100	84	97
PPV (%)	68	97	74	100	66	96	65	93	75	100	61	90

Legend to figures

Figure 1. Panel A, Serum cortisol levels at different sampling times. Data are shown as mean (95%CI) and mean differences (MD) (95%CI). Comparisons among time points were performed by a repeated-measure ANOVA addressing main effects by a Bonferroni's confidence interval adjustment. **P* value < 0.001. **Panel B**, Pearson's correlation analysis between serum cortisol values at 30 and 60 min sampling times. Solid red line represents the simple linear regression and dotted black lines represent the 95%CI of the regression line. **Panel C**, Bland-Altman plot. Solid black line represents the perfect agreement among both time points. Solid blue line is the mean of the percentage difference among both sampling times, and dashed blue lines are ± 2 standard deviation (SD) of that mean. Solid red line is the regression line of the percentage differences.

Figure 2. Baseline and stimulated serum cortisol concentrations as a function of clinical suspicion and response to cosyntropin test. Data are shown as mean and 95%CI. *Abbreviations:*

AI: Adrenal Insufficiency

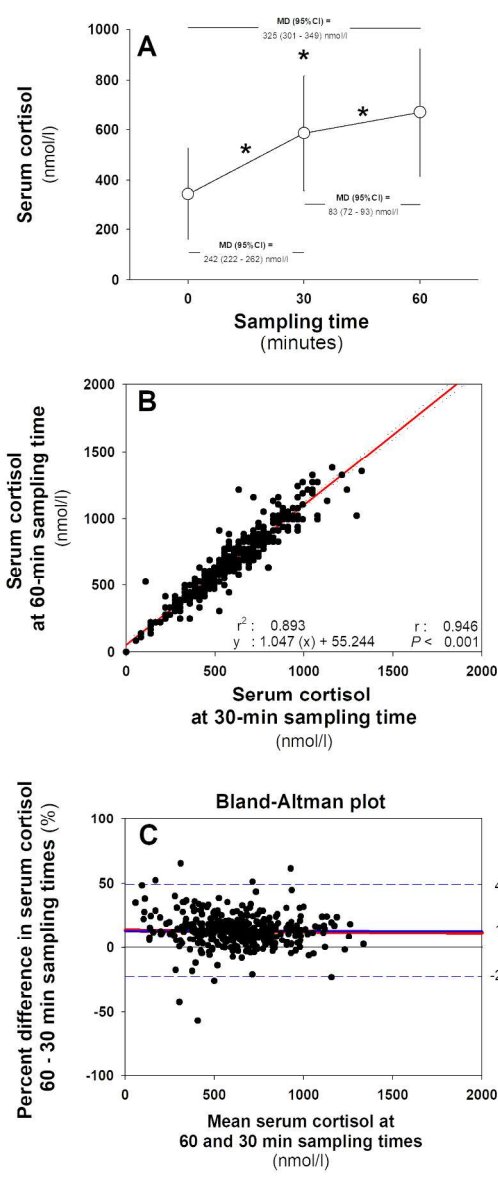
Figure 3. Subgroups of patients according to serum cortisol responses to cosyntropin-stimulation as a function of classic and sex- and assay-specific cut-offs. Figures on top of the bars indicate the number of patients included in each subgroup. Diagnostic agreement is shown as the percentage of observed agreements and kappa coefficients (95%CI).

Figure 4. Subgroups of patients according to serum cortisol responses to cosyntropin-stimulation as a function of cut-off values and clinical suspicion of primary or central AI. Figures on top of

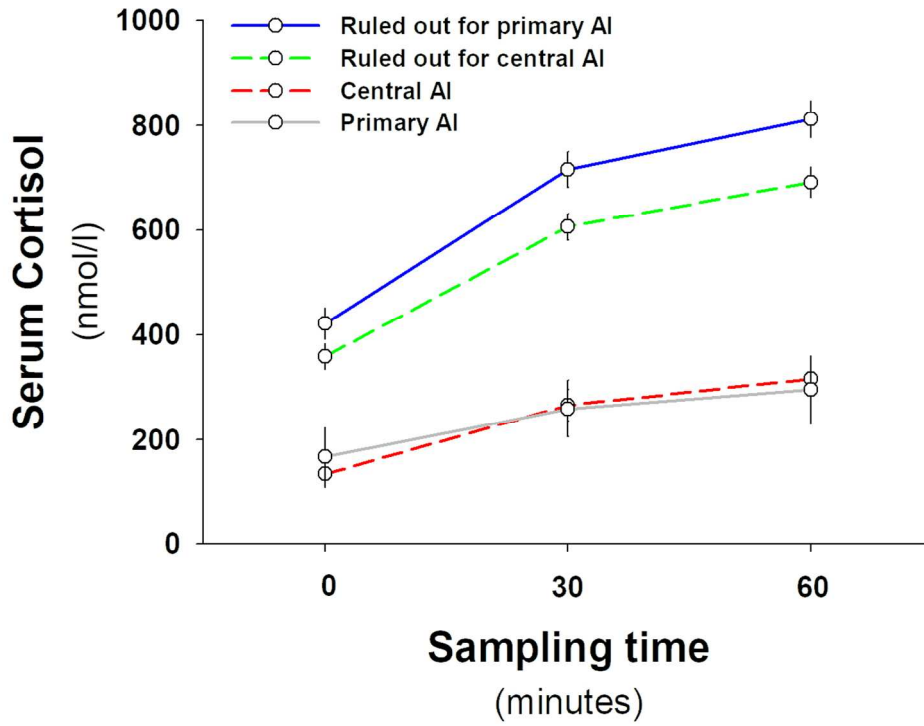
1
2
3 the bars show the number of patients included in the different subgroups. Diagnostic agreement
4
5 is shown as the percentage of observed agreements and kappa coefficients (95%CI).
6
7
8
9

10 **Figure 5.** Descriptive statistics and distribution of 30 min cosyntropin-stimulated serum cortisol
11 concentrations in a population of premenopausal healthy women with evidence of normal HPA
12 axis function. The boundary of the box closest to zero indicates the 25th percentile, the solid and
13
14 long dash lines within the box marks the median and mean, respectively, and the boundary of the
15
16 farthest from zero indicates the 75th percentile. Whiskers above and below the box indicate the
17
18 90th and 10th percentiles. Black circles represent the 5th percentile and the dashed red line
19
20 indicates the lower limit of normality (2.5th percentile) for each immunoassay.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

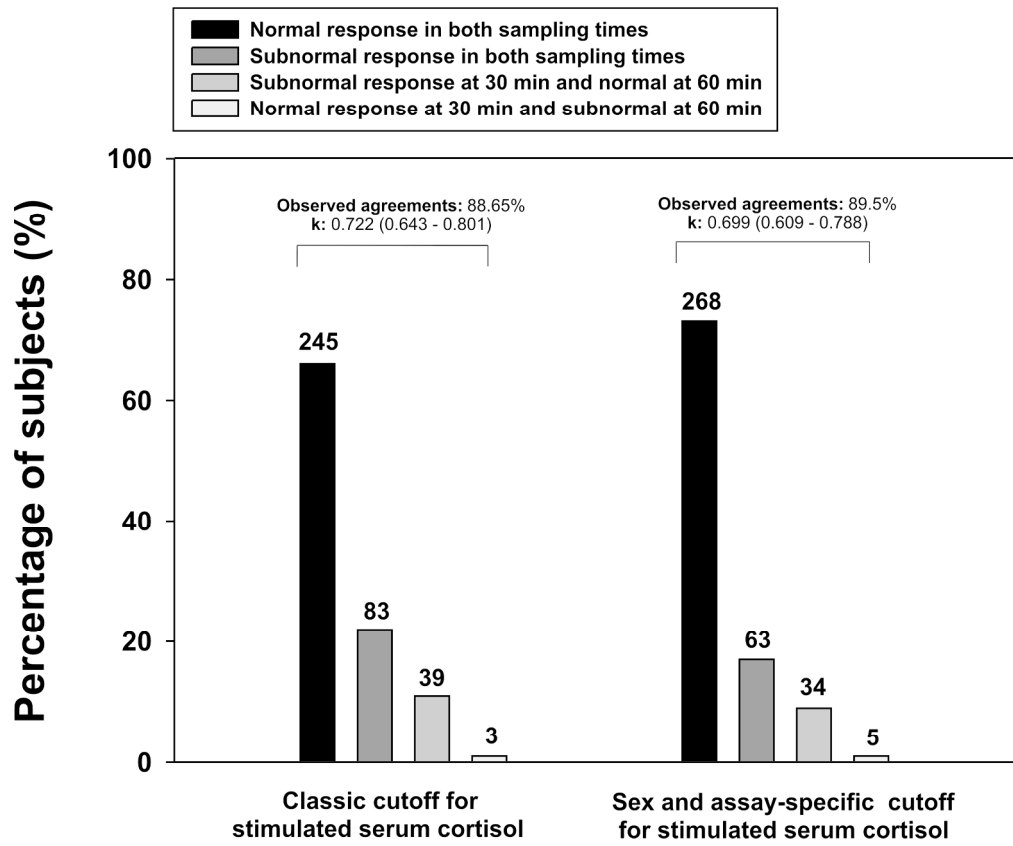
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



199x282mm (300 x 300 DPI)

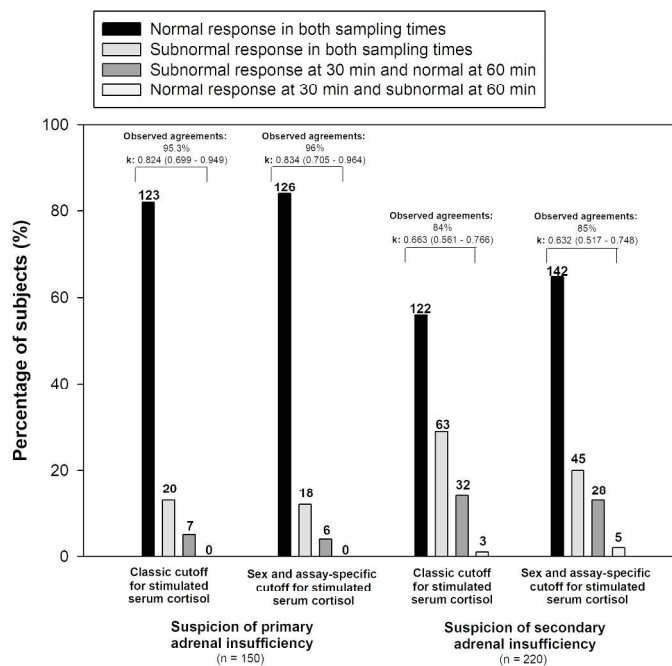


118x97mm (300 x 300 DPI)



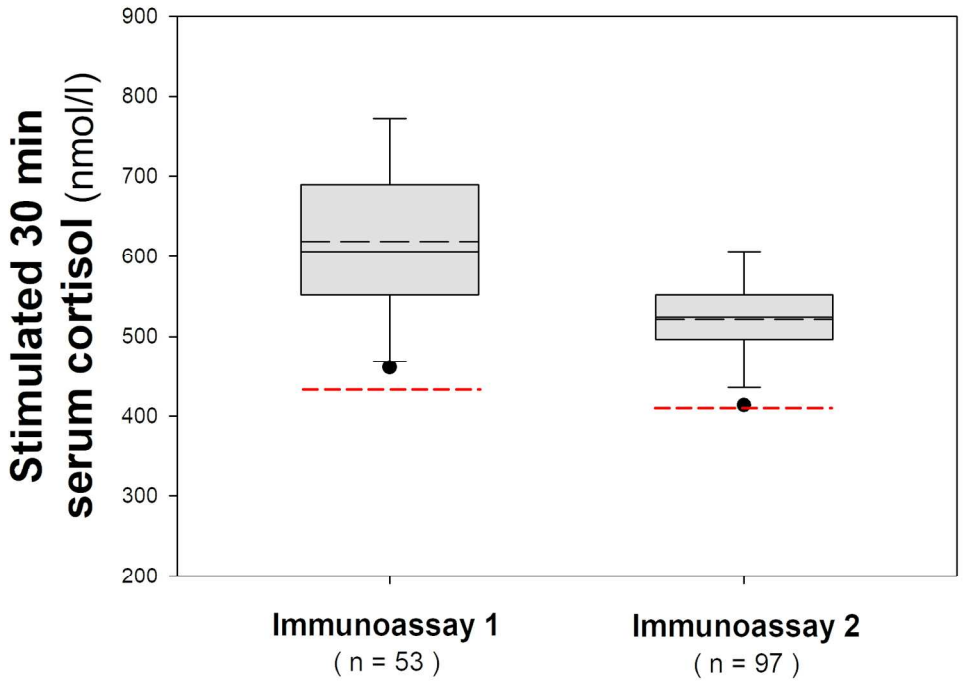
www only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



296x209mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



160x120mm (300 x 300 DPI)

View Only

The STARD 2015 list*

Section and topic	No	Item
Title or abstract:		
Page 3, Line 47 to 49	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)
Abstract:		
Page 3 to 4	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)
Introduction:		
Page 5 to 6	3	Scientific and clinical background, including the intended use and clinical role of the index test
Page 6, line 109	4	Study objectives and hypotheses
Methods		
Study design: Page 7, Line 121	5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)
Participants Page 7	6	Eligibility criteria
Page 7, Line 124-133; Page 8 Line 139 and page 9 line 168	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)

Section and topic	No	Item
Page 7, line 122	8	Where and when potentially eligible participants were identified (setting, location, and dates)
Page 7, line 124 - 133	9	Whether participants formed a consecutive, random, or convenience series
Test methods Page 9 - 10	10a	Index test, in sufficient detail to allow replication
Page 9 - 10	10b	Reference standard, in sufficient detail to allow replication
Page 9 - 10	11	Rationale for choosing the reference standard (if alternatives exist)
Page 10, line 193-196	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory
Page 10, line 193-196	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory
Page 20, line 130	13a	Whether clinical information and reference standard results were available to the performers or readers of the index test
Page 20, line 130	13b	Whether clinical information and index test results were available to the assessors of the reference standard
Analysis: Page 9, Line 187	14	Methods for estimating or comparing measures of diagnostic accuracy
Methods section, page 8, line 150-159. Therefore, no indeterminate data were avoided.	15	How indeterminate index test or reference standard results were handled
N/A (no missing data of the index text)	16	How missing data on the index test and reference standard were handled

Section and topic	No	Item
Page 10	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory
N/A	18	Intended sample size and how it was determined
Results		
Flow not performed. Description of the flow of participants at page 11	19	Flow of participants, using a diagram
Table 1, page 25. Also in results, page 11	20	Baseline demographic and clinical characteristics of participants
Figure 1, page 31 and result section, page 11	21a	Distribution of severity of disease in those with the target condition
N/A. No alternative diagnosis were studied	21b	Distribution of alternative diagnoses in those without the target condition
N/A.	22	Time interval and any clinical interventions between index test and reference standard
Test results , Figure page 31, figure 2 page 32, figure 3 page 33	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard
Table 3, page 28	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)
N/A. Our index test / reference standard has no adverse events to report.	25	Any adverse events from performing the index test or the reference standard
Discussion		
Page 17, line 376	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Section and topic	No	Item
Page 16 and page 18 (Conclusions)	27	Implications for practice, including the intended use and clinical role of the index test
Other information		
N/A, it is a retrospective chart review.	28	Registration number and name of registry
N/A. Study protocol was attached to editors as supplementary file.	29	Where the full study protocol can be accessed
Page 19 and 20	30	Sources of funding and other support; role of funders

Peer review only