

CLINICAL STUDY

Assessment of ACTH assay variability: a multicenter study

Francesca Pecori Giraldi, Andrea Saccani, Francesco Cavagnini and The Study Group on the Hypothalamo-Pituitary–Adrenal Axis of the Italian Society of Endocrinology

Chair of Endocrinology, University of Milan, Ospedale San Luca, Istituto Auxologico Italiano IRCCS, Piazzale Brescia 20, 20149 Milan, Italy

(Correspondence should be addressed to F Cavagnini; Email: cavagnini@auxologico.it)

Abstract

Objective: Measurement of plasma ACTH plays a pivotal role in the diagnosis and treatment of several endocrine disorders. Little is known, however, on the variability of ACTH assay results and its impact on clinical practice. The aim of the present study was to assess the performance of routine plasma ACTH measurements.

Design: Twenty-five fresh-frozen plasma samples collected from patients with either high, low, or normal ACTH concentrations were measured using seven different assays by 35 different laboratories. Assay precision, method agreement, and result classification were estimated.

Results: Inter- and intra-assay coefficient of variation varied considerably with some assays achieving < 10%, others consistently achieving > 20%. Overall method agreement was good (mean ratio versus target value 1.02) but subject to exceedingly large excursion (lower and upper limits of agreement at 0.13 and 1.91 respectively). Both differences between assays and between laboratories contributed to variability of method agreement. Assays correctly classified most patients with normal and high ACTH concentrations (90% (95% CI 82–97%) and 95% (95% CI 86–100%) respectively), whereas only 60% (95% CI 52–67%) of measurements from patients with low ACTH values were assigned correctly.

Conclusions: Field ACTH assays have to be interpreted with caution as they are burdened by high variability and often fail to correctly identify patients with suppressed ACTH secretion. The endocrine community has to include ACTH assays among those requiring standardization.

European Journal of Endocrinology 164 505–512

Introduction

Diagnosis of endocrine disorders relies to a large extent on hormone measurement in biological samples, be it plasma, serum, or urine. Accuracy of hormone assays is crucial for both screening procedures, e.g. vitamin D deficiency, hypothyroidism, and more specialized diagnoses, such as acromegaly, GH deficiency, adrenal insufficiency, and hypogonadism. However, concerns have been raised as to the low accuracy of some hormone assays, most notably GH, testosterone, and 25-hydroxyvitamin D (1–3). During the course of a clinical audit on Cushing's syndrome (4), we have come to appreciate that commercially available ACTH assays present considerable differences in terms of sensitivity and normal range, leading to a wide variability in ACTH measurements. Reliable ACTH assays are essential to distinguish between elevated and, more importantly, suppressed ACTH levels. The former can unmask mildly impaired adrenal function/reserve, while the latter differentiates between ACTH-dependent and ACTH-independent Cushing's syndrome and identifies patients with ACTH deficiency. In the course of the above-mentioned study (4), we were struck by the not

negligible number of patients with adrenal Cushing's syndrome in whom plasma ACTH levels were not suppressed but clearly measurable or even present within the normal range, thereby possibly misleading the diagnostic work-up.

ACTH assays have evolved considerably since the first RIAs developed by Berson & Yalow (5), and currently available ACTH assays are mostly fully automated, require minimal plasma samples, and are performed by high throughput laboratories not specialized in hormone assays. However, adequate information is as yet not available on the impact of these changes on ACTH measurements in clinical practice. We therefore decided to evaluate the performance of routine ACTH assays by comparing results from representative Italian endocrinological centers.

Methods

Endocrinological centers all over Italy participated in this prospective study, encompassing 35 laboratories from general hospitals, medical centers, and university clinics in order to reflect overall ACTH assay

performance. Participants (see below for details) had responded to the call of the Study Group on the Hypothalamo-Pituitary-Adrenal Axis of the Italian Society of Endocrinology. Plasma samples from individuals with high (patients #1–2 and #4–6 with Addison's disease, patients #3 and #7 adrenalectomized for Cushing's disease), normal (normal subjects #8–18), and low ACTH levels (patients #22–24 with autoimmune diseases on 8–16 mg methylprednisolone daily for at least 4 months; patients #19–20 and #25 with adrenal Cushing's syndrome, and patient #21 with isolated ACTH deficiency) were used. In detail, patients with adrenal Cushing presented urinary free cortisol (UFC) levels ranging from 225 to 645 $\mu\text{g}/24\text{ h}$ (normal range in our laboratory is 10–80 $\mu\text{g}/24\text{ h}$) and cortisol levels after 1 mg dexamethasone between 15.4 and 29 $\mu\text{g}/\text{dl}$. The patient with isolated ACTH deficiency (#21) presented morning serum cortisol levels around 2.3–3.8 $\mu\text{g}/\text{dl}$ (normal range 5–25) and UFC levels from 8.7 to 10 $\mu\text{g}/24\text{ h}$.

A large blood sample from each subject/patient was collected into EDTA-containing tubes, centrifuged at 4 °C, plasma-pooled, then aliquoted, and stored at –80 °C. Samples were collected in a single sitting and shipped to individual laboratories in dry ice within 2 months of collection in order to avoid peptide deterioration. Laboratories were blinded to patient status, assayed all samples in a single run using their standard procedure, and returned the results to our center for analysis together with their technical assay sheets. A single measurement or the mean of two values was used for each patient, for singlicate and duplicate measurements respectively, as per each laboratory's procedure. All samples were assayed by every laboratory. For intraassay variation, 16 additional replicates of samples from patients #2, #4, #5, and #7 for high ACTH values, patients #12, #13, #14, and #15 for normal ACTH values, and patients #19 and #20 for low ACTH values were measured with each assay in a single laboratory. The total number of samples available for analysis was 1835. The lower limit of detection reported by assay kit manufacturers as the smallest value, which can be distinguished from zero at the 95% confidence limit or mean + 2s.d. is given in Table 1. Samples yielding ACTH concentrations below or above the calibration curve of a given assay were arbitrarily assigned the lowest/highest value of said calibration curve. Producers of assays used by participants were Diagnostic Products Corporation (now Siemens Healthcare Diagnostics; DPC), Los Angeles, CA, USA; Diagnostic System Laboratories, Inc. (DSL), Webster, TX, USA; Scantibodies Laboratory, Inc., Santee, CA, USA; Diasorin, Stillwater, MN, USA; Nichols Institute Diagnostics, San Juan Capistrano, CA, USA; and CisBio International, Saclay, France. The study was approved by the Ethical Committee of the Istituto Auxologico Italiano, and subjects gave their informed consent to participate in the study.

Statistical analysis

Imprecision across methods and for individual methods was determined by Variance Function Program 5.0 (Christchurch, New Zealand). Mean, s.d., and coefficient of variation (CV) were calculated, and the median of method means (MedMM) (6) for each patient was used as a target value, as this represents a relative calibrator given that there is no accepted reference method procedure for quantification of plasma ACTH. For comparisons between standards, MedMM was calculated from assays using the same reference preparation (RP) (see Table 1). Data were analyzed by nested ANOVA for repeated measures with methods or standards as between factors and laboratories using the same method/standard as within factors. Method agreement was established according to Bland–Altman plots comparing candidate methods with MedMM both in terms of absolute differences and ratio (=method value/MedMM) with 95% limits of agreement (7). Method alignment was analyzed by weighted Deming regression using CBstat v.5.1 (Risskov, Denmark), as measurement error was inconstant over the explored range (8). Agreement in decision making, i.e. values interpreted as high, normal, or low, was assessed by Cohen's *K*-statistic (9) using prevalence/bias-adjusted kappa (PABAK) values with 95% confidence intervals since the prevalence was skewed towards the expected result (10). PABAK was calculated for agreement between each of two assays, and then it averaged over the entire series. PABAK values >0.6 defined good agreement (11). Statistical significance was accepted for $P < 0.05$.

Results

Samples were assayed by 35 centers, using seven different assays (Table 1). Figure 1 shows a wide distribution of ACTH measurements for each individual with either high, normal, or low ACTH values. For example, in subject #8, in whom normal ACTH concentrations were expected, values as low as 7.0 pg/ml and as high as 69.0 pg/ml were reported across assays and laboratories, whereas in patient #4 with Addison's disease, values ranged from 955.3 to 11 625 pg/ml. In patient #20 with adrenal Cushing's syndrome, results of ACTH assays ranged from <5 to 34.0 pg/ml.

Method precision

Inter- and intra-assay imprecision profiles are shown in Fig. 2. Intraassay CV (panel A) was <20% for DPC, Nichols IRMA, DiaSorin, and CisBio assays, while higher values were observed for chemiluminescent Nichols and DSL. Intraassay evaluation for Scantibodies could not be performed, as the assay is not distributed in

Table 1 Features of assays used in the study.

	Manufacturer						
	Diagnostic Product Corporation	Nichols	Nichols	CisBio	DiaSorin	Scantibodies	Diagnostic System Laboratories
Data reported on technical sheets							
Catalog #	LKAC (Immulite) L2KAC (Immulite 2000)	62-7004 60-4175	40-2195	ELSA-ACTH	27 065 27 130	3KG100	DSL-5100
Method	CLS	CLS	IRMA	IRMA	IRMA	IRMA	IRMA
Lower limit of detection	5 pg/ml (LKAC) 9 pg/ml (L2KAC)	1 pg/ml	1 pg/ml	2 pg/ml	1.5 pg/ml	1 pg/ml	1.5 pg/ml
Standard curve range	5-1250 pg/ml	2-1350 pg/ml	5-1400 pg/ml	18-1950 pg/ml	9.9-1330 pg/ml	5-1250 pg/ml	6-2000 pg/ml
Normal range	<46 pg/ml (LKAC) 10-60 pg/ml (L2KAC)	9-52 pg/ml (#7004) 7-51 pg/ml (#4175)	9-52 pg/ml	10-90 pg/ml	6-56.7 pg/ml	10-86 pg/ml	10-60 pg/ml
Reference preparation	Purified ACTH (NHPP)	Purified ACTH (NHPP)	Purified ACTH (WHO 74/555)	Purified ACTH (WHO 74/555)	Purified ACTH (NHPP)	Purified ACTH (WHO 74/555)	Purified ACTH (WHO 74/555)
Coefficient of variation (%)							
Intra	<10	<5	<4	<7	<6	<4	<10
Inter	<10	<9	<8	<6	<6	<7	<9
Data obtained from the present study							
Number of laboratories using the assay	12	7	7	3	3	2	1
Difference relative to MedMM							
Mean ratio	0.84	0.99	1.04	1.2	1.23	1.12	0.85
95% limits of agreement	0.40; 1.28	0.52; 1.46	0.35; 1.73	0.31; 2.07	0.61; 1.86	0.58; 1.66	0.35; 1.90
Alignment with MedMM							
Slope (95% CI)	0.95 (0.93; 0.98)*	0.98 (0.93; 1.02)	1.36 (1.29; 1.44)*	1.59 (1.34; 1.84)*	1.07 (1.04; 1.12)*	1.02 (0.75; 1.23)	0.76 (0.67; 0.87)*
Intercept (95% CI)	-0.56 (-0.69; -0.43)*	-0.07 (-0.27; 0.11)	-4.57 (-6.22; -2.92)*	-5.38 (-9.23; -1.53)*	0.72 (0.40; 1.05)	-0.19 (-3.93; 3.78)	-5.76 (-8.64; -2.89)*
Coefficient of variation (%)							
Intra	1-6	2-34	2-19	1-17	2-12	-	12-53
Inter	6-46	9-26	5-82	9-45	6-42	9-84	-
Decision making (% correctly identified values)							
Low ACTH	81	65	84	41	3	78	31
Normal ACTH	95	82	75	100	97	92	94
High ACTH	98	100	100	100	100	100	92

CLS, chemiluminescent immunometric assay; NHPP, National Hormone Pituitary Program. *P<0.01 versus MedMM.

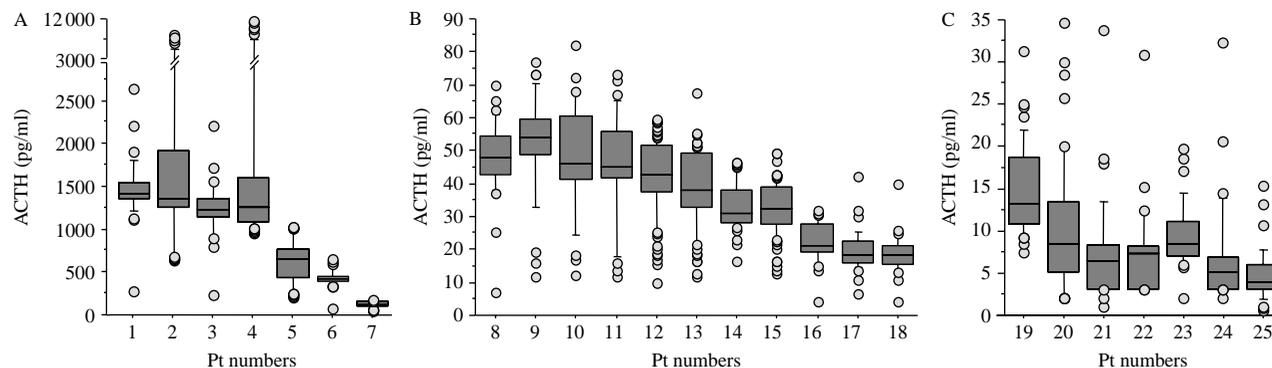


Figure 1 Centile distribution of ACTH concentrations in each patient. Box depicts median, 25° and 75° percentile; 10° and 90° percentile as s.e.m. and values outside 10° and 90° are shown as dots. Panel A: results in patients with high ACTH secretion; Panel B: results in patients with normal ACTH secretion; Panel C: results in patients with suppressed ACTH secretion.

Italy any more. Of note, maximal acceptable intraassay CV for ACTH assays is 15.5%, according to the College of American Pathologists (12); thus, the majority of assays present adequate repeatability. Interassay imprecision was greater than intraassay imprecision for all assays, as expected; precision profiles (Fig. 2, panel B) spanned 20–40% for most assays, although higher values were observed for Scantibodies and CisBio assays. DSL interassay imprecision could not be evaluated as the assay was used only in one laboratory. Overall, these precision profiles were worse than those reported in technical sheets, reflecting the difficulty in achieving manufacturer evaluation performance under routine workload conditions and considerable between-laboratory variability (Table 1).

Method agreement

The Bland–Altman plot showing the difference in individual measurements versus the target value (i.e. MedMM) reveals a proportional error, i.e. an increasing difference in ACTH values at higher ACTH concentrations (Fig. 3, panel A); therefore, agreement was re-assessed using the ratio of individual measurements versus calibration value. Overall mean ratio was 1.02, therefore close to full agreement with the calibrator, but 95% limits of agreement were quite large, 0.13 and 1.91 respectively, indicating that assay results could be nearly twofold higher or <1/5 of the expected value. Evaluation according to the source of reference preparation, e.g. National Hormone Pituitary Program or WHO 74/555, yielded near to superimposable ratios among assays using the same standard (1.0 (95% limits: 0.38; 1.61) and 1.02 (95% limits: 0.08; 1.98) respectively), although agreement between results from assays using the two standards again ranged widely (1.0 (95% limits: 0.70; 1.26)) for both low and high ACTH values (Fig. 3, panel B).

As regards individual assays, DSL and DPC appeared to mostly underestimate and CisBio and Scantibodies overestimate compared with MedMM, but drifts and

wide limits of agreement were apparent in nearly all assays (see Table 1 and Fig. 3, panel C). Analysis of method agreement by weighted Deming regression analysis revealed good alignment between Nichols chemiluminescent assay, Scantibodies, and the target value (Table 1), whereas a significant proportional bias (i.e. slope deviation from 1) was detected for all other assays, the most skewed being CisBio, Nichols radioactive assay, and DSL (Table 1). No constant error was detected for DiaSorin, Nichols chemiluminescent assay, and Scantibodies, i.e. intercept comprises 0, whereas a significant intercept deviation was detected for DSL, CisBio, Nichols IRMA assay and, to a lesser extent, for DPC (Table 1). Analysis of ACTH measurements by repeated-measures ANOVA (data were analyzed as the ratio of MedMM to achieve normal distribution) revealed significant differences in assay results between standards ($F=47.65$, $P<0.001$), between methods ($F=9.625$, $P<0.01$), and between laboratories ($F=4.65$, $P<0.01$); thus, all items contribute to the lack of agreement.

Decision making

ACTH measurements from patients were related to the reference range reported by individual laboratories and classified as high/normal/low ACTH values (Table 1). Overall, 85.1% (95% CI 83.0–87.2%) of values were interpreted correctly, with high and normal ACTH values achieving correct interpretation in 98.4% (95% CI 86.0–100%) and 89.6% (95% CI 81.8–97.4%) of measurements respectively, whereas only 59.5% (95% CI 52.3–66.8%) of measurements in patients in whom ACTH was expected to be suppressed were classified correctly. Of note, the vast majority of ACTH values in patients with suppressed ACTH (i.e. adrenal Cushing's syndrome, prolonged steroid treatment, and isolated ACTH deficiency) were deemed 'normal' by the DiaSorin assay, as were 60–70% of values measured by CisBio and DSL. In fact, even the best performing assays yielded 15–20% misclassifications (Table 1).

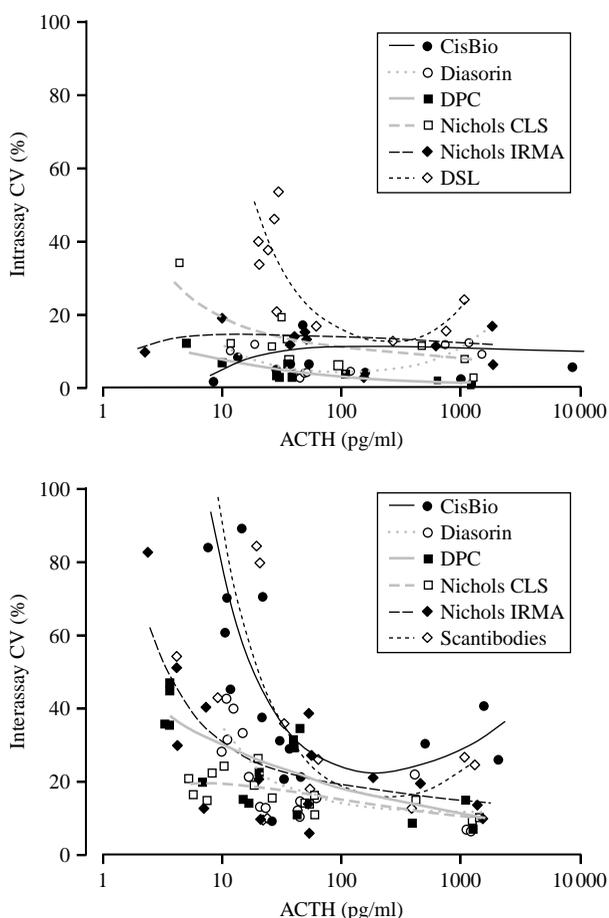


Figure 2 Imprecision profiles for each assay calculated according to Variance Function Program 5.0 (Christchurch, New Zealand). Panel A shows intraassay coefficient of variation; panel B shows interassay coefficient of variation.

On balance, overall agreement in decision making among assays was good, as mean PABAK equaled 0.88 (95% CI 0.86–0.90) due to the high agreement between values measured in patients with high ACTH concentrations (mean PABAK 0.95 (95% CI 0.89–1.00)) and good agreement for patients with normal ACTH (mean PABAK 0.79 (95% CI 0.71–0.86)). However, agreement in result assignment for patients with low ACTH values was poor (mean PABAK 0.28 (95% CI 0.17–0.38)).

Discussion

Wide variations between hormone measurements is a problem that plagued clinicians in the past, when RIA measurements were most common and nowadays, even with automated immunoassays, it still hampers reliable across-laboratory precision and, hence, the correct diagnosis. Low interlaboratory reproducibility and accuracy are being reported for several hormones,

most notably serum testosterone (1), GH (2), and 25-hydroxyvitamin D (3), and attributed to a variety of causes, including different antibody epitope specificity (3) or calibrants (2), heterogeneity of circulating isoforms (13, 14), interference with binding proteins (15) and variable recovery after extraction (16, 17). Interlaboratory variability becomes particularly relevant if specific thresholds for hormonal concentrations are used for clinical decision making, be it diagnosis of acromegaly with oral glucose load (18) or GH deficiency after provocative testing (19), adrenal insufficiency (20), or screening for hypothyroidism (21) or for hypovitaminosis D (3).

As regards the hypothalamo-pituitary-adrenal axis, validation studies for serum and urinary cortisol, as well as other steroids, yielded variable results, with some investigators reporting acceptable interassay precision (6) and others reporting large intermethod variation (16), especially for serum cortisol levels after dexamethasone suppression (22) or Synacthen stimulation (20) and salivary cortisol (23). Measurement of plasma ACTH evolved from bioassays on guinea pig or rat adrenal cells (24) through the milestone RIA of Berson & Yalow in 1968 (5) to radioactive or non-isotopic immunometric assays. The development and marketing of new ACTH assays usually led to comparison studies with existent assays and commonly revealed good correlation and comparable clinical sensitivity (25–27), although widely different results or limited agreement between two assays have also been observed (28, 29).

Our study shows a marked variability in assay performance and both constant and proportional bias among tested assays, which encompass major ACTH assay manufacturers. Most assays showed adequate repeatability, i.e. intraassay CV <15%, but assay performance in routine conditions was consistently less than precision profiles reported in assay technical sheets. Furthermore, both a systematic and proportional error was apparent, an effect due to assay RP as well as laboratory differences. These difficulties in ACTH assay agreement are well summarized by Bland-Altman statistics, which shows a good mean agreement with the expected value but an unacceptably large range of agreement for each measurement. From a clinical viewpoint, this translates into an acceptable agreement in the assignment of results obtained in patients with either normal or high ACTH values, where large excursions among measurements do not necessarily alter their clinical significance, but causes poor agreement and frequent result misclassification among patients with low ACTH levels.

Assay performance is a major concern for ACTH assay kits and, in some cases, the reason for assay withdrawal (30). Our study includes assays which are currently off the market but, given their extensive use in the past, for example the Nichols' assay in the US (26), it provides a good basis for comparison, as diagnostic cut-off values developed by Nichols' users might require

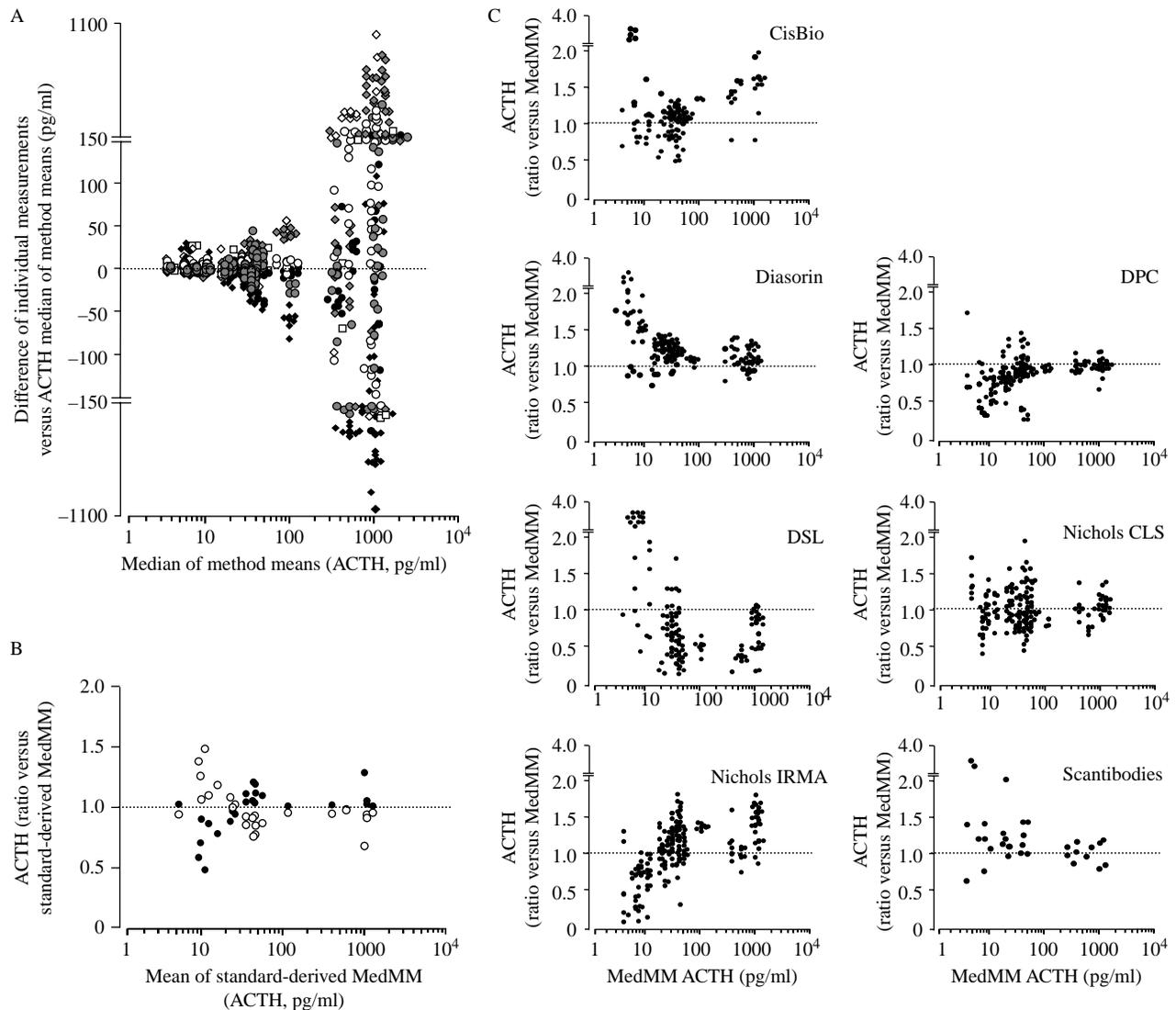


Figure 3 Bland–Altman deviance plot. (Panel A) Difference to relative target value, i.e. MedMM for each measurement (difference = 0 indicates full agreement); DPC (black dot), Nichols chemiluminescent immunometric assay (gray dot) and radioactive immunometric assay (gray diamond), CisBio (empty diamond), Scantibodies (empty square), DiaSorin (empty dot), and DSL (black diamond). (Panel B) Ratio to mean of standard-derived MedMM value. Mean ratio in each patient for assays using NHPP (black dot) or WHO (white dot) RP is shown. (Panel C) Ratio to MedMM value for each measurement graphed by assay. Ratio = 1 represents full agreement.

readjustments. Absolute ACTH values are commonly used for the distinction of ACTH-independent and ACTH-dependent Cushing's syndrome, for the differential diagnosis of adrenal insufficiency and to assess effects of treatment or during follow-up of patients with adrenal or pituitary disorders. In these contexts, interassay and interlaboratory agreement among ACTH measurements becomes mandatory.

Several causes may contribute to assay variability, starting with the different standards used by individual assays, as not all ACTH (1–39) formulations are equally recognized by assay antibodies. The different potency of ACTH standards had already been observed in the past (31) and is currently difficult to rectify, as there is no

International Standard for human ACTH. Indeed, the only International Standard for corticotropin is derived from pig pituitaries (32), and assays use synthetic or pituitary-derived purified preparations. It is also important to recall that ACTH adsorbs to unsiliconized glass tubes and degrades rapidly at room temperature (33); thus, the use of EDTA-coated plastic tubes, rapid sample processing, and cool storage prior to assay is mandatory. These factors were carefully controlled in our study but may play an additional, confounding role in everyday field ACTH measurements.

Incorrect interpretation of low ACTH values and poor agreement in value assignment in these patients are in line with results obtained on a large group of patients

with adrenal Cushing's syndrome (4) and other anecdotal observations in patients with ACTH-independent Cushing's syndrome who do not display suppressed or undetectable ACTH values (34–36), as would be expected. Indeed, diagnostic flowcharts commonly exclude adrenal Cushing's syndrome if ACTH values are above 1–2 or 5–10 pg/ml (37, 35). The fact that 40% of plasma ACTH measurements fell into the normal range in patients in whom ACTH secretion should be suppressed raises the issue of whether this pattern is due to technical problems or whether the pathophysiology of glucocorticoid negative feedback and secondary adrenal insufficiency should be revisited. Lastly, low plasma ACTH is considered a useful parameter for the diagnosis of subclinical Cushing's syndrome in patients with an adrenal incidentaloma (38), an ever increasing clinical issue, and the uncertainty in measurement of low ACTH concentrations represents an important obstacle.

In conclusion, notwithstanding the progress in assay automation, our study revealed a considerable variability in field ACTH measurements. This proved particularly relevant for low ACTH values, as up to 40% of results were classified incorrectly. On the basis of this evidence, we recommend that variability among ACTH assays should be kept in mind when dealing with different ACTH assays and low ACTH values should be interpreted with caution; this should be taken into account for both clinical and experimental studies. Endocrinologists should therefore evaluate the performance of their ACTH assays beyond standard laboratory Quality Control testing, and the Endocrinological Community should strive for harmonization of reference materials and re-evaluation of normal ranges as well as cut-off values to be used by commercial ACTH assays.

Participating laboratories

R Giorgino, S Martemucci (Bari); G Pagani, L Gualandris (Bergamo), R Pasquali, U Pagotto (Bologna); A Giustina, G Bulgari, C Iacobello (Brescia); S Loche, P Civolari (Cagliari), R Vigneri, G Parrinello, A Mirone (Catania); G Borretta, G Visconti (Cuneo); F Minuto, L Fazzuoli (Genoa); F Trimarchi, S Cannavò, A Artemisia, G Giorgianni (Messina); R Lanzi, A Soldarini (Milan); P Loli, D Valente (Milan); M Bevilacqua, T Vago, G Baldi (Milan); P Beck Peccoz, I Vaghi (Milan); C Carani, G F Baracchini, E Baraldi, M De Santis (Modena); S Mariotti, F Atzeni (Monserrato); A Colao, V Macchia, R Pivonello (Naples); A Bellastella, T Criscuolo, F Ariano, S Pepe (Naples); A Angeli, M Terzolo, F Daffara (Orbassano); C Foresta, M Plebani, D Faggian (Padua); F Mantero, C Scaroni (Padua); F Santeusano, R Pippi (Perugia); E Martino, L Manetti (Pisa); R Valcalvi, C Dotti, L Belloni (Reggio Emilia); V Toscano, A Proietti (Rome); C Zoppi, L De Marinis, S Corsello (Rome); F Sciarra, A Lenzi, G Landini, A Radicioni (Rome); G Frajese, O Porzio,

A Vassallo (Rome); B Ambrosi, C Dall'Asta (San Donato Milanese); V Trischitta, L D'Alosio (San Giovanni Rotondo); G Delitata, F Sanci (Sassari); C De Sanctis, F Altare (Turin); E Arvat, A Bertagna (Turin); M Muggeo, P Moghetti, R Castello, R Dorizzi (Verona); C Pandolfi, A Cerri (Vizzolo Predabissi).

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

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

Acknowledgements

We wish to thank Marco Besozzi, MD, Istituto Auxologico Italiano, Milan, Italy and William Sadler, MD, Christchurch, New Zealand for their software and statistical support.

References

- 1 Rosner W, Auchus RJ, Azziz R, Sluss PM & Raff H. Utility, limitations, and pitfalls in measuring testosterone: an endocrine society position statement. *Journal of Clinical Endocrinology and Metabolism* 2007 **92** 405–413. (doi:10.1210/jc.2006-1864)
- 2 Bidlingmaier M & Freda PU. Measurement of human growth hormone by immunoassays: current status, unsolved problems and clinical consequences. *Growth Hormone and IGF Research* 2010 **20** 19–25. (doi:10.1016/j.ghir.2009.09.005)
- 3 Binkley N, Krueger D, Cowgill CS, Plum L, Lake E, Hansen KE, DeLuca HF & Drezner MK. Assay variation confounds the diagnosis of hypovitaminosis D: a call for standardization. *Journal of Clinical Endocrinology and Metabolism* 2004 **89** 3152–3157. (doi:10.1210/jc.2003-031979)
- 4 Invitti C, Pecori Giraldi F, De Martin M, Cavagnini F & The Study Group of the Italian Society of Endocrinology on the Pathophysiology of the Hypothalamic–Pituitary–Adrenal Axis. Diagnosis and management of Cushing's syndrome: results of an Italian multicentre study. *Journal of Clinical Endocrinology and Metabolism* 1999 **84** 440–448. (doi:10.1210/jc.84.2.440)
- 5 Berson AS & Yalow RS. Radioimmunoassay of ACTH in plasma. *Journal of Clinical Investigation* 1968 **47** 2725–2751. (doi:10.1172/JCI105955)
- 6 Palmer-Toy DE, Wang E, Winter WE, Soldin SJ, Klee GG, Howanitz JH & Elin RJ. Comparison of pooled fresh frozen serum to proficiency testing material in College of American Pathologists Survey. *Archives of Pathology and Laboratory Medicine* 2005 **129** 305–309.
- 7 Bland JM & Altman DG. Measuring agreement in method comparison studies. *Statistical Methods in Medical Research* 1999 **8** 135–160. (doi:10.1191/096228099673819272)
- 8 Linnet K. Evaluation of regression procedures for methods comparison studies. *Clinical Chemistry* 1993 **39** 424–432.
- 9 Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 1960 **20** 37–46. (doi:10.1177/001316446002000104)
- 10 Byrt T, Bishop J & Carlin JB. Bias, prevalence and kappa. *Journal of Clinical Epidemiology* 1993 **46** 423–428. (doi:10.1016/0895-4356(93)90018-V)
- 11 Byrt T. How good is that agreement? *Epidemiology* 1996 **7** 561.

- 12 Harris EK. Proposed goals for analytical precision and accuracy in single-point diagnostic testing. *Archives of Pathology and Laboratory Medicine* 1988 **112** 416–420.
- 13 Baumann G. Growth hormone heterogeneity: genes, isohormones, variants, and binding proteins. *Endocrine Reviews* 1991 **12** 424–449. (doi:10.1210/edrv-12-4-424)
- 14 Lepage R, Roy L, Brossard JH, Rousseau L, Dorais C, Lazure C & D'Amour P. A non-(1–84) circulating parathyroid hormone (PTH) fragment interferes significantly with intact PTH commercial assay measurements in uremic samples. *Clinical Chemistry* 1998 **44** 805–809.
- 15 Ebdrup L, Fisker S, Sorensen HH, Ranke MB & Ørskov H. Variety of growth hormone determinations due to use of different immunoassays and to the interference of growth hormone-binding proteins. *Hormone Research* 1999 **51** (Supplement 1) 20–26. (doi:10.1159/000053131)
- 16 Moore A, Aitken R, Burke CW, Gaskell S, Groom G, Holder G, Selby C & Wood P. Cortisol assays: guidelines for the provision of a clinical biochemistry service. *Annals of Clinical Biochemistry* 1985 **22** 435–454.
- 17 Lips P, Chapuy MC, Dawson-Hughes B, Pols HAP & Holick MF. An international comparison of serum 25-hydroxyvitamin D measurements. *Osteoporosis International* 1999 **9** 394–397. (doi:10.1007/s001980050162)
- 18 Arafat AM, Mohlig M, Weickert MO, Perschel FH, Purschwitz J, Spranger J, Strasburger CJ, Schöfl C & Pfeiffer AFH. Growth hormone response during oral glucose tolerance test: the impact of assay method on the estimation of reference values in patients with acromegaly and in healthy controls, and the role of gender, age, and body mass index. *Journal of Clinical Endocrinology and Metabolism* 2008 **93** 1254–1262. (doi:10.1210/jc.2007-2084)
- 19 Rakover Y, Lavi I, Masalah R, Issam T, Weiner E & Ben-Shlomo I. Comparison between four immunoassays for growth hormone (GH) measurement as guides to clinical decisions following GH provocative tests. *Journal of Pediatric Endocrinology and Metabolism* 2000 **13** 637–643.
- 20 Chatha KK, Middle JG & Kilpatrick ES. National UK audit of the short Synacthen test. *Annals of Clinical Biochemistry* 2010 **47** 158–164. (doi:10.1258/acb.2009.009209)
- 21 Brabant G, Beck-Peccoz P, Jarzab B, Laurberg P, Orgiazzi J, Szabolcs I, Weetman AP & Wiersinga WM. Is there a need to redefine the upper normal limit of TSH? *European Journal of Endocrinology* 2006 **154** 633–637. (doi:10.1530/eje.1.02136)
- 22 Odagiri E, Naruse M, Terasaki K, Yamaguchi N, Jibiki K, Takagi S, Tanabe M & Takano K. The diagnostic standard of preclinical Cushing's syndrome: evaluation of the dexamethasone suppression test using various cortisol kits. *Endocrine Journal* 2004 **51** 295–302. (doi:10.1507/endocrj.51.295)
- 23 Baid SK, Sinaii N, Wade M, Rubino D & Nieman LK. Radioimmunoassay and tandem mass spectrometry measurement of bedtime salivary cortisol levels: a comparison of assays to establish hypercortisolism. *Journal of Clinical Endocrinology and Metabolism* 2007 **92** 3102–3107. (doi:10.1210/jc.2006-2861)
- 24 Sayers G. Bioassay of ACTH using isolated cortex cells. *Annals of the New York Academy of Sciences* 1977 **297** 220–241. (doi:10.1111/j.1749-6632.1977.tb41856.x)
- 25 White A, Smith H, Hoadley ME, Dobson SH & Ratcliffe JG. Clinical evaluation of a two-site immunoradiometric assay for adrenocorticotrophin in unextracted human plasma using monoclonal antibodies. *Clinical Endocrinology* 1987 **26** 41–51. (doi:10.1111/j.1365-2265.1987.tb03637.x)
- 26 Raff H & Findling JW. A new immunoradiometric assay for corticotropin evaluated in normal subjects and patients with Cushing's syndrome. *Clinical Chemistry* 1989 **35** 596–600.
- 27 Raff H. Immulite vs Scantibodies IRMA plasma ACTH assay. *Clinical Chemistry* 2008 **54** 1409–1410. (doi:10.1373/clinchem.2008.107987)
- 28 Vogeser M, Engelhardt D & Jacob K. Comparison of two automated adrenocorticotrophic hormone assays. *Clinical Chemistry* 2000 **46** 1998–2000.
- 29 Howe LA & Smeaton T. Anomalous results with the CIS-Sorin adrenocorticotrophic hormone radioimmunoassay kit. *Clinical Chemistry* 1979 **25** 816–817.
- 30 FDA Recalls and field corrections. Enforcement report accessed at <http://www.fda.gov/Safety/Recalls/EnforcementReports/2005/ucm120361.htm> on 3-8-2005.
- 31 Dobson SH, Gibson S & White A. Assessment of the potency of different standards in the immunoradiometric assay of ACTH. *Annals of Clinical Biochemistry* 1989 **26** 96–101.
- 32 Bangham DR, Musset MV & Stack-Dunne MP. The third international standard for corticotrophin. *Bulletin of the World Health Organization* 1962 **27** 395–408.
- 33 Evans MJ, Livesey JH, Ellis MJ & Yandle TG. Effect of anticoagulants and storage temperatures on stability of plasma and serum hormones. *Clinical Biochemistry* 2001 **34** 107–112. (doi:10.1016/S0009-9120(01)00196-5)
- 34 Klose M, Kofoed-Enevoldsen A & Østergaard Kristensen L. Single determination of plasma ACTH using an immunoradiometric assay with high detectability differentiates between ACTH-dependent and -independent Cushing's syndrome. *Scandinavian Journal of Clinical and Laboratory Investigation* 2002 **62** 33–38. (doi:10.1080/003655102753517181)
- 35 Lindsay JR, Shanmugam VK, Oldfield EH, Remaley AT & Nieman LK. A comparison of immunometric and radioimmunoassay measurement of ACTH for the differential diagnosis of Cushing's syndrome. *Journal of Endocrinological Investigation* 2006 **29** 983–988.
- 36 Penezic Z, Zarkovic K, Vujovic S, Beleslin B, Ciric J & Drezgic M. Value of assessing adrenocorticotrophic hormone (ACTH) in differential diagnosis of hypercorticism. *Medicinski Pregled* 2004 **57** 335–342. (doi:10.2298/MPNS0408335P)
- 37 Raff H & Findling JW. A physiologic approach to diagnosis of Cushing syndrome. *Annals of Internal Medicine* 2003 **138** 980–991.
- 38 Nieman LK. Approach to the patient with an adrenal incidentaloma. *Journal of Clinical Endocrinology and Metabolism* 2010 **95** 4106–4113. (doi:10.1210/jc.2010-0457)

Received 11 January 2011

Accepted 20 January 2011