Flaminia Fanelli*, Stephen Bruce, Marco Cantù, Anastasia Temchenko, Marco Mezzullo, Johanna M. Lindner, Mirko Peitzsch, Pierre-Alain Binz, Mariette T. Ackermans, Annemieke C. Heijboer, Jody Van den Ouweland, Daniel Koeppl, Elena Nardi, Manfred Rauh, Michael Vogeser, Graeme Eisenhofer and Uberto Pagotto

Report from the HarmoSter study: inter-laboratory comparison of LC-MS/MS measurements of corticosterone, 11-deoxycortisol and cortisone

https://doi.org/10.1515/cclm-2022-0242
Received March 15, 2022; accepted September 27, 2022; published online October 27, 2022

Abstract

Objectives: Liquid chromatography-tandem mass spectrometry (LC-MS/MS) panels that include glucocorticoid-related steroids are increasingly used to characterize and diagnose adrenal cortical diseases. Limited information is currently available about reproducibility of these measurements among laboratories. The aim of the study was to compare LC-MS/MS measurements of corticosterone, 11-deoxycortisol and cortisone at eight European centers and assess the performance after unification of calibration.

Methods: Seventy-eight patient samples and commercial calibrators were measured twice by laboratory-specific procedures. Results were obtained according to in-house and external calibration. We evaluated intra-laboratory and inter-laboratory imprecision, regression and agreement against performance specifications derived from 11-deoxycortisol biological variation.

Results: Intra-laboratory CVs ranged between 3.3 and 7.7%, 3.3 and 11.8% and 2.7 and 12.8% for corticosterone, 11-deoxycortisol and cortisone, with 1, 4 and 3 laboratories often exceeding the maximum allowable imprecision (MAI), respectively. Median inter-laboratory CVs were 10.0, 10.7 and 6.2%, with 38.5, 50.7 and 2.6% cases exceeding the MAI for corticosterone, 11-deoxycortisol and cortisone, respectively. Median laboratory bias vs. all laboratory-medians ranged from −5.6 to 12.3% for corticosterone, −14.6 to 12.4% for 11-deoxycortisol and −4.0 to 6.5% for cortisone, with few cases exceeding the total allowable error. Modest deviations were found in regression equations among most laboratories. External calibration did not improve 11-deoxycortisol and worsened corticosterone and cortisone inter-laboratory comparability.

Conclusions: Method imprecision was variable. Inter-laboratory performance was reasonably good. However,

***Corresponding author: Flaminia Fanelli**, Senior Assistant Professor, Department of Medical and Surgical Sciences, Endocrinology Research Group, Center for Applied Biomedical Research, University of Bologna, S.Orsola Policlinic, Via Massarenti 9, 40138 Bologna, Italy, Phone: +39 051 2143902, Fax: +39 051 2143902, E-mail: flaminia.fanelli2@unibo.it. https://orcid.org/0000-0003-2601-2694
Stephen Bruce and Pierre-Alain Binz, Clinical Chemistry Laboratory, University Hospital of Lausanne (CHUV), Lausanne, Switzerland. https://orcid.org/0000-0002-0045-7698 (P.-A. Binz)
Marco Cantù, Laboratory of Clinical Biochemistry and Pharmacology, Institute of Laboratory Medicine EOLAB, Ente Ospedaliero Cantonale, Bellinzona, Switzerland
Anastasia Temchenko and Marco Mezzullo, Department of Medical and Surgical Sciences, Endocrinology Research Group, Center for Applied Biomedical Research, University of Bologna, Bologna, Italy
Johanna M. Lindner and Michael Vogeser, Institute of Laboratory Medicine, Hospital of the University of Munich (LMU), Munich, Germany
Mirko Peitzsch and Graeme Eisenhofer Institute of Clinical Chemistry and Laboratory Medicine, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany. https://orcid.org/0000-0002-8601-9903 (G. Eisenhofer)
Mariette T. Ackermans, Department of Clinical Chemistry, Endocrine Laboratory, Amsterdam UMC, Amsterdam, Netherlands; and University of Amsterdam, Amsterdam, Netherlands
Annemieke C. Heijboer, Department of Clinical Chemistry, Endocrine Laboratory, Amsterdam UMC, Amsterdam, Netherlands; University of Amsterdam, Amsterdam, Netherlands; and Vrije Universiteit Amsterdam, Amsterdam, Netherlands
Jody Van den Ouweland, Department of Clinical Chemistry, Canisius-Wilhelmina Hospital, Nijmegen, Netherlands
Daniel Koeppl and Manfred Rauh, Department of Pediatrics and Adolescent Medicine, University Hospital, Erlangen, Germany
Elena Nardi, Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Bologna, Italy
Uberto Pagotto, Department of Medical and Surgical Sciences, Endocrinology Research Group, Center for Applied Biomedical Research, University of Bologna, Bologna, Italy; and Endocrinology and Prevention and Care of Diabetes Unit, IRCCS Azienda Ospedaliero-Universitaria Policlinico S.Orsola di Bologna, Bologna, Italy

Open Access. © 2022 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

cases with imprecision and total error above the acceptable limits were apparent for corticosterone and 11-deoxycortisol. Variability did not depend on calibration but apparently on imprecision, accuracy and specificity of individual methods. Tools for improving selectivity and accuracy are required to improve harmonization.

Keywords: 11-deoxycortisol; calibration; corticosterone; cortisone; harmonization; inter-laboratory performance; liquid chromatography – tandem mass spectrometry; method comparison; steroid hormones.

Introduction

The laboratory diagnosis of adrenal diseases is primarily based on the measurement of selective steroid hormones, such as cortisol, aldosterone and 17OH-progesterone. Increasing popularity of multi-analyte panels measured by liquid chromatography-tandem mass spectrometry (LC-MS/MS) allows for the investigation of steroids of the glucocorticoid pathway as potentially relevant players in the management of endocrine disorders [1–3]. Measurements of corticosterone, 11-deoxycortisol and cortisone in conjunction with steroids with established diagnostic relevance are in particular gaining relevance for identification and characterization of subclinical and overt hypercortisolism [4–8], hyperaldosteronism [9–11] and adrenal cortex tumors [12–14]. Overall, recent studies highlight the preeminent utility of 11-deoxycortisol as a biomarker for detecting hypercortisolism and adrenocortical carcinoma, and of corticosterone in predicting the cardiovascular risk in patients affected by mild autonomous cortisol secretion [2–8, 12–14]. Moreover, monitoring adrenal steroids in the setting of congenital deficits in steroidogenesis facilitates identification of specific enzymes that carry mutations [15–20]. This thereby complements genetic investigations and assists the identification of non-classical or heterozygous forms of congenital adrenal hyperplasia [19, 20].

Despite experimental evidence on the importance of glucocorticoid pathway steroids as disease-specific biomarkers or as components of multi-steroidal predictive algorithms, corticosterone, 11-deoxycortisol and cortisone are rarely measured in clinical practice. There is also a lack of tools and procedures to assess the quality of LC-MS/MS determinations. In fact, currently, there are no reference measurement procedures, nor traceable matrix-based quality controls [21], and a few quality assessment (EQA) programs are only available for 11-deoxycortisol and corticosterone. As a consequence, there is a lack of information about the reproducibility and accuracy of LC-MS/MS methods for the aforementioned corticosteroids.

To address this shortcoming, the HarmoSter consortium recently launched a study on the harmonization of the LC-MS/MS measurement of 10 circulating steroids [22]. In our first report, we showed that LC-MS/MS methods characterized by heterogeneous pre-analytical and analytical methodologies can provide a satisfactory, albeit improvable, level of harmonization for measurements of cortisol, 17OH-progesterone and aldosterone. Moreover, we highlighted advantages for the adoption of a common calibration system [22]. This second report focuses on evaluation of intra- and inter-laboratory performance of eight LC-MS/MS methods for measurements of corticosterone, 11-deoxycortisol and cortisone, and whether adoption of common calibration materials might improve harmonization.

Materials and methods

Consortium and methods

The study design and Consortium are detailed elsewhere [22]. The study received approval by the local Ethics Committee (no. 141/2017/U/Tess). Seven laboratories used laboratory developed tests (LDTs) (Laboratories B, C, D, E, G, H and I) and one used a commercial kit (MassChrom®, Chromsystems, Munich, Germany; https://chromsystems.com) (Laboratory L) for multi-steroid analysis including corticosterone, 11-deoxycortisol and cortisone (Supplementary Table 1). Among LDTs, Laboratories D and E used the 6PLUS1® Multilevel Serum Calibrator set (Chromsystems) for in house calibration. Technical details and in-house measurement ranges for each laboratory are shown in Tables 1 and 2.

Study samples

Briefly, the sample set included 78 sets of plasma and serum samples from 26 volunteers (women/men: 13/13; age: 20–69 years). As previously detailed [22], each volunteer donated blood via three different vacuum tubes: 1. gel-separator serum, 2. bead clot activator serum and 3. lithium-heparin plasma. Additionally, the 6PLUS® Multilevel Serum Calibrator set was provided from Chromsystems (lot.5016, different from lots used for in house calibration by Laboratories D, E and L) [22]. This external calibration set included corticosterone, 11-deoxycortisol and cortisone, with respective measurement ranges of 1.33–139.0, 0.24–40.7 and 2.59–108.0 nmol/L, and was traceable to methanol certified reference materials (CRM) from an ISO 17025 and 17034 certified supplier.

Running scheme and quantitation

Samples and external calibrators were measured twice in two separate runs, each including an independent in house calibration set, according to protocols ordinarily used by each laboratory.

Table 1: Calibration, sample preparation, liquid chromatographic and mass spectrometry features of assays.

Lab	Ref	In house calibration	Sample volume, µL	Sample preparation	Instrument	Run time, min	Purification column	LC column	Mobile phases (A; B)	Source	Corticosterone			11-Deoxycortisol			Cortisone		
											Analyte (ion mode) mass transition; retention time, min	IS	IS (ion mode) mass transition; retention time, min	Analyte (ion mode) mass transition; retention time, min	IS	IS (ion mode) mass transition; retention time, min	Analyte (ion mode) mass transition; retention time, min	IS	IS (ion mode) mass transition; retention time, min
B	[23, 24]	Cerilliant CRM[a] in 4% BSA	600	PP: $ZnSO_4$ in MeOH; SPE: C18	Series 200, Perkin Elmer; API4000 QTrap, Sciex	21	POROS R1/20	Luna RP-C8 100 × 4.6 mm, 5 µm	20% MeOH in H_2O; MeOH	APCI	(+) 347/121; 9.19	Corticosterone-[2,2,4,6,6,17,21,21-D8]	(+) 355/125; 9.04	(+) 347/109; 9.55	11-Deoxycortisol-[21,21-D2]	(+) 349/97; 9.51	(+) 361/163; 6.62	Cortisone-[2,2,4,6,6,12,12-D7]	(+) 369/169; 6.57
C	[25]	Gravimetric in charcoal stripped serum	50	PP: ZnTFA in MeOH	Acquity UPLC; Xevo TQ-S, Waters	10		Kinetex Biphenyl 150 × 2.1 mm, 1.7 µm	0.2 mM NH_4F in H_2O; MeOH	ESI	(+) 347/121; 4.21	Corticosterone-[2,2,4,6,6,17,21,21-D8]	(+) 355/337; 4.13	(+) 347/97; 3.86	11-Deoxycortisol-[D7[b]]	(+) 354/100; 3.81	(+) 361/163; 2.74	Cortisone-[2,2,4,6,6,9,21,21-D8]	(+) 369/168; 2.71
D	[26]	6PLUS1®[c], Chromsystems	500	SPE: Oasis HLB	Acquity UPLC, Waters;API5500 Qtrap, Sciex	14		Kinetex C18 100 × 2.1 mm, 2.5 µm	5 mM NH_4F in H_2O; MeOH	ESI[a]	(+) 347/329; 5.88	Corticosterone-[2,2,4,6,6,17α,21,21-D8]	(+) 355/337; 5.85	(+) 347/97; 6.12	11-Deoxycortisol-[2,2,4,6,6-D5]	(+) 352/100; 6.07	(+) 361/163; 3.38	Cortisone-[1,2-D2]	(+) 363/165; 3.32
E	[27]	6PLUS1®[a,c], Chromsystems	100	PP: $ZnSO_4$ in MeOH	1260 Infinity, Agilent; API6500 Qtrap, Sciex[a]	7	Chromolith SpeedRod C18 4.6 × 50 mm	Chromolith Performance C18, 100 × 2.1 mm	5 mM NH_4HCO_2 in H_2O; 5 mM NH_4HCO_2 in MeOH[a]	ESI[a]	(+) 347/121; 3.86	Corticosterone-[2,2,4,6,6,17,21,21-D8]	(+) 355/125; 3.80	(+) 347/109; 4.00	Corticosterone-[2,2,4,6,6,17,21,21-D8]	(+) 355/125; 3.80	(+) 361/163; 2.86	Cortisone-[2,2,4,6,6,12,12-D7]	(+) 368/169; 2.84
G	n.a.	Cerilliant CRM in steroid free plasma	25	PP: ACN	Acquity-Xevo TQ-S, Waters	9.9		HSS T3 2.1 × 100 mm, 1.8 µm	0.1% FA in H2O; 0.1% FA in ACN	ESI	(+) 347/121; 5.60	Corticosterone-[2,2,4,6,6,17,21,21-D8]	(+) 355/125; 5.55	(+) 347/97; 5.75	11-Deoxycortisol-[2.2.4.6.6-D5]	(+) 352/100; 5.72	(+) 361/163; 4.31	Cortisone [2.2.4.6.6.12.12-D7]	(+) 369/169; 4.26
H	n.a.	Cerilliant CRM in 10% MeOH	1,000	PP: $ZnSO_4$ in MeOH; SPE: C18	Acquity-Xevo TQ-S, Waters	12		BEH C18 2.1 × 100 mm, 1.7 µm	0.05% FA in H_2O; 0.05% FA in MeOH	ESI	(+) 347/121; 5.78	Corticosterone-[9,11,12,12-D4]	(+) 351/121; 5.74	(+) 347/97; 6.01	11-Deoxycortisol-[2,2,4,6,6-D5]	(+) 352/100; 5.96	(+) 361/121; 3.17	Cortisone-[2,2,4,6,6,12,12-D7]	(+) 369/169; 3.11
I	n.a	Gravimetric in charcoal stripped serum	100	PP: H_3PO_4; SPE: Oasis MCX	I-class Acquity; TQS, Waters[a]	15		C18 Zorbax Eclipse Plus 2.1 × 50 mm, 1.8µm[a]	H_2O; MeOH	ESI	(+) 347/121; 4.60	Corticosterone-[2,2,4,6,6,17α,21,21–D8]	(+) 355/125; 4.50	(+) 347/97; 4.80	11-Deoxycortisol-[2,2,4,6,6-D5]	(+) 352/100; 4.75	(+) 361/163; 3.00	Cortisone-[2,2,4,6,6,9,12,12-D8]	(+) 369/169; 2.95
L[d]		6PLUS1®[c], Chromsystems	500	Chromsystems' intellectual property	1290 series HPLC 6490, Agilent	11.5		Chromsystems' intellectual property	Chromsystems' intellectual property	ESI	(+) 347/121; 8.68	Chromsystems' intellectual property	(+) 355/125; 8.53	(+) 347/97; 9.05	Chromsystems' intellectual property	(+) 352/100; 8.94	(+) 361/163; 6.41	Chromsystems' intellectual property	(+) 369/168; 6.31

[a]Modified from original publication. [b]deuterium positions not available. [c]Lots different from the one distributed in the study. [d]https://chromsystems.com/en/products/steroids/masschromr-steroids-in-serum-plasma-with-96-spe-well-plate-lc-ms-ms-72072-96%20&%2072072-480.html. N.a., not applicable; CRM, certified reference material; IS, internal standard; PP, protein precipitation; ZnSO4, zinc sulfate; MeOH, methanol; SPE, solid phase extraction; RP, reversed phase; APCI, atmospheric pressure chemical ionization; ZnTFA, zinc trifluoracetate; NH_4F, ammonium fluoride; ESI, electrospray ionization; HLB, hydrophilic-lipophilic balance; NH_4HCO_2, ammonium formate; FA, formic acid; ACN, acetonitrile; PBS, phosphate buffered saline; HSS, high strength silica; BEH, ethylene bridged hybrid; H_3PO_4, phosphoric acid; MCX, mixed-mode strong cation-exchange.

Table 2: Measurement ranges, imprecision and impact of the calibration systems within assay methods.

Analyte	Lab	In house					In house	External	%	p-Value[b]
		LLOQ-ULOQ, nmol/L	n	Mean (min–max), nmol/L	Intra-lab CV% (min–max)	n	Median (min–max), nmol/L	Median (min–max), nmol/L	Deviation (external vs. in house)	
Corticosterone	B	0.903–577.3	78	19.8 (3.7–96.5)	3.3 (0.1–10.2)	78	11.2 (3.7–96.5)	11.5 (3.8–99.3)	3.0	<0.001
	C	0.577–331.9	78	18.8 (2.7–94.3)	5.9 (0.2–17.6)	78	10.4 (2.7–94.3)	10.0 (2.6–89.5)	−3.4	<0.001
	D[a]	1.46–137.4	78	20.1 (4.3–86.7)	3.8 (0.0–11.0)	78	12.2 (4.3–86.7)	10.8 (3.5–79.9)	−11.5	<0.001
	E[a]	1.328–144.3	78	21.4 (3.9–107.7)	4.6 (0.0–14.1)	78	11.5 (3.9–107.7)	11.7 (4.0–109.8)	2.0	<0.001
	G	0.5–100	78	19.5 (3.7–91.0)	4.3 (0.0–11.0)	78	11.2 (3.8–91.0)	10.6 (3.3–86.3)	−5.7	<0.001
	H	0.4–115	78	19.2 (3.1–102.5)	3.6 (0.1–9.6)	78	11.0 (3.1–102.5)	12.3 (3.3–115.5)	11.6	<0.001
	I	0.5–240	75	19.2 (2.9–92.0)	4.9 (0.0–28.4)	75	10.6 (2.9–92.0)	11.6 (3.2–101.9)	10.0	<0.001
	L[a]	0.506–289	78	22.8 (4.0–112.8)	7.7 (0.1–16.5)	78	13.4 (4.0–112.8)	14.5 (4.2–122.9)	8.1	<0.001
11-Deoxycortisol	B	0.225–28.86	75	0.96 (0.23–3.90)	3.7 (0.0–10.7)	72	0.75 (0.29–3.90)	0.82 (0.29–4.42)	9.6	<0.001
	C	0.289–170.87	63	0.99 (0.33–3.95)	7.3 (0.3–27.5)	63	0.77 (0.33–3.95)	0.74 (0.38–3.54)	−4.9	0.023
	D[a]	0.130–28.86	78	0.81 (0.18–3.66)	9.4 (0.0–26.6)	69	0.67 (0.29–3.66)	0.68 (0.26–4.05)	2.2	<0.001
	E[a]	0.144–40.41	74	0.82 (0.17–3.73)	11.8 (0.0–35.1)	67	0.61 (0.24–3.73)	0.66 (0.25–4.01)	6.9	<0.001
	G	1–400	13	2.38 (1.05–4.07)	7.1 (1.2–16.5)	13	2.15 (1.05–4.07)	2.18 (1.16–3.95)	1.4	0.455
	H	0.05–29	78	0.83 (0.20–3.56)	3.3 (0.1–11.9)	72	0.65 (0.27–3.56)	0.72 (0.28–4.07)	10.5	<0.001
	I	0.2–160	70	0.99 (0.24–4.12)	4.3 (0.0–11.1)	70	0.70 (0.24–4.12)	0.70 (0.25–4.01)	0.8	0.811
	L[a]	0.268–39.54	72	1.05 (0.32–4.61)	11.1 (0.0–28.7)	72	0.79 (0.32–4.61)	0.80 (0.33–4.74)	2.3	<0.001
Cortisone	B	0.338–277.4	78	53.4 (35.3–74.1)	2.7 (0.0–8.0)	78	54.3 (35.3–74.1)	58.4 (38.0–79.7)	7.6	<0.001
	C	2.886–693.6	78	55.0 (35.6–73.6)	7.6 (0.1–18.6)	78	56.6 (35.6–73.6)	62.0 (39.4–80.7)	9.9	<0.001
	D[a]	0.130–277.4	78	54.6 (37.4–74.4)	3.2 (0.0–18.7)	78	55.9 (37.4–74.4)	55.4 (36.8–73.8)	−0.9	<0.001
	E[a]	0.289–111.0	78	54.4 (36.4–74.5)	4.2 (0.7–8.0)	78	55.5 (36.4–74.5)	56.1 (36.9–75.3)	1.1	<0.001
	G	1–400	78	59.1 (38.2–78.7)	3.7 (0.0–12.5)	78	60.6 (38.2–78.7)	59.6 (37.6–77.2)	−1.8	<0.001
	H	2–1,387	78	56.6 (32.6–86.5)	8.3 (0.1–23.4)	78	56.4 (32.6–86.5)	47.5 (27.6–72.6)	−15.8	<0.001
	I	0.2–320	75	57.9 (33.8–81.9)	4.3 (0.0–25.3)	75	59.9 (33.8–82.0)	58.2 (32.9–79.6)	−2.8	<0.001
	L[a]	3.00–107.6	78	57.6 (38.5–92.3)	12.8 (0.1–36.4)	78	57.1 (38.5–92.3)	59.6 (40.1–96.5)	4.4	<0.001

[a]In-house calibration material from Chromsystems. [b]Values obtained by in house and external calibration were compared by the Wilcoxon test.

Quantification was based on in house and external calibration sets [22]. Use of external calibrators was restricted to those with concentrations within in house measurement ranges. Calibration curves displayed R^2>0.97.

Data analysis and statistics

All measurements are reported in nmol/L. To convert to ng/mL, multiply by 0.346 for corticosterone and 11-deoxycortisol, and by 0.360 for cortisone. Results were excluded from the study if they were below the in house lower limit of quantification (LLOQ) or above the upper LOQ (ULOQ). Data were also excluded from the external calibration dataset when outside of measurement range. Analyses were performed on the overall sample set and included the three types of blood-derived specimens. Means and CVs of duplicate measurements were calculated (Supplementary Tables 2–4). Since measurements of glucocorticoid pathway steroids could contribute to diagnosis and management of patients with certain adrenal diseases, it is important to establish the allowable performance of measurement methods. Here, we have defined the maximum allowable imprecision (MAI) and the total allowable error (TAE) based on available data for within-subject (CVi) and between-subject (CVg) biological variation [28]. The following equations were applied accordingly:

$$MAI=0.5*CVi;$$

$$TAE=0.25*(CVi^2+CVg^2)^{0.5}+1.65*(0.5*CVi) \text{ [29].}$$

Data on biological variability were available for 11-deoxycortisol (CVi: 21.3% and CVg: 31.5%) but not for corticosterone and cortisone [28, 30]. In absence of specific data, for the purpose of the present study, MAI (10.7%) and TAE (27.1%) calculated for 11-deoxycortisol were also used for corticosterone and cortisone evaluations.

Intra-laboratory performance: The CV of each duplicate measurement was determined as the % of duplicate standard deviation over duplicate mean. The overall intra-laboratory imprecision was determined as the duplicate measurement CV calculated according to the root mean square method [31] as follows: $CV=\sqrt{\frac{1}{n}\sum_{i=1}^{n} D_i^2}*100$, where $D_i=\frac{(x_{i1}-x_{i2})/\sqrt{2}}{\overline{X}_i}$, where x_{i1} and x_{i2} are duplicate measures of each sample, $\overline{X}_i$ is the arithmetic mean of the duplicates and n is the total number of duplicates. Intra-laboratory CVs from each laboratory were then compared with the MAI. Within-method impact of calibration was evaluated by the Wilcoxon test and by Passing-Bablok regression. Slope and intercept coefficients were considered not to significantly deviate from the line of best fit when their 95% confidence interval (95CI) contained 1 and 0, respectively.

Inter-laboratory performance: Inter-laboratory analyses were only performed on duplicate means with concentrations measurable within measurement ranges of both calibration systems. Moreover, samples that displayed a CV of duplicate measurements >30% were excluded (Supplementary Tables 2–4). For each sample, the inter-laboratory CV among the eight laboratories was calculated. Thereafter, the median of all inter-laboratory CVs was used to assess between-method reproducibility and compared with the MAI. Between-method regression was assessed by Passing-Bablok analysis of laboratory values compared to all laboratory medians.

The laboratory %-bias vs. all laboratory median and Bland-Altman analyses were computed to determine between-method agreement; results were compared with the TAE. The Wilcoxon test was used to compare median inter-laboratory CV and median bias, whereas the F test was used to compare the variance of laboratory %-bias by in house and external calibration. Statistics were performed by SPSS (v.20, IBM Co., Somers, NY) and MedCalc (v.18.2.1; Mariakerke, Belgium).

Results

Corticosterone

Concentrations in plasma and serum samples ranged from 3.7 to 97.4 nmol/L (median of all laboratories by in house calibration). The intra-laboratory duplicate measurement CVs ranged from 3.3 to 7.7% (Table 2). The CV of individual duplicates listed in Supplementary Table 2 were mostly within MAI except for Laboratory L, which exceeded the threshold in 15.4% of samples. The external calibration range was within or slightly above the in house measurement range of all laboratories. As reported in Table 2, the two calibration sets determined significantly different results within all laboratories ($p<0.001$), with results from external calibration that deviated from −11.5 to 11.6% compared to in house values. Passing-Bablok comparison of results obtained by the two calibration systems within laboratories that used Chromsystems calibrators for in house calibration detected substantial consistency in Laboratory E and modest deviations in Laboratory D and L (Supplementary Figure 1).

With results from in house calibration, the median inter-laboratory CV was 10.0%, though for 38.5% the inter-laboratory CV was above MAI. Use of external calibration significantly increased the median CV% to 12.6% ($p>0.001$) and numbers of results >MAI to 66.7% (Figure 1).

Passing-Bablok analysis of laboratory results vs. the median of all laboratories is shown in Figure 2. Slopes were not statistically different from 1 in four laboratories by in house and in one laboratory by external calibration; coefficients (95CI) ranged from 0.957 (0.944–0.969) to 1.153 (1.131–1.180) with the former, and from 0.867 (0.852–0.882) to 1.239 (1.203–1.262) with the latter (Supplementary Table 5).

Figure 3 shows the plot of laboratory bias vs. the median of all laboratories. In addition, Supplementary Figure 2 shows the Bland-Altman plots of bias from individual laboratories as functions of hormone concentration. The median laboratory bias ranged from −5.6 to 12.3% with in house, and from −10.1 to 19.5% with external calibration.

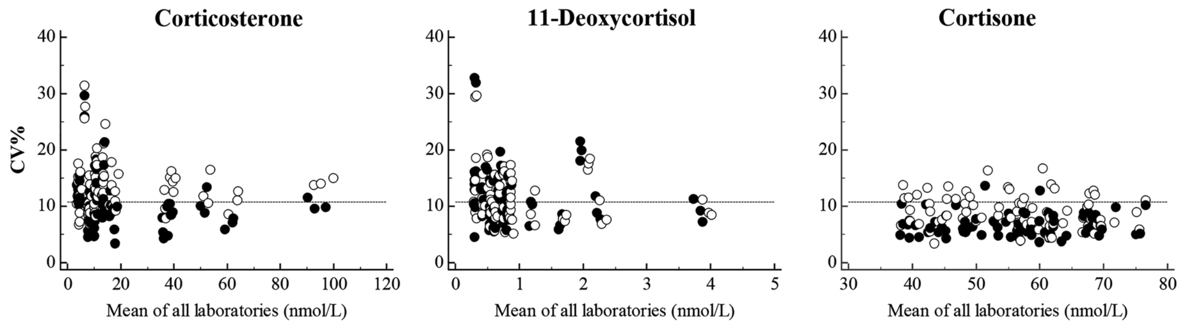

Figure 1: Inter-laboratory coefficients of variation as function of the calibration system. Black dots: in house calibration; white dots: external calibration. Dashed lines: maximal allowable imprecision (10.7%). The inter-laboratory CV (min–max) was overall lower when using in house compared to external calibration for corticosterone (10.0% (3.4–29.6%) vs. 12.6% (6.7–31.3%), respectively; p<0.001) and cortisone (6.2% (3.5–13.5%) vs. 8.1% (3.3–16.6%), respectively; p<0.001). No significant differences were found for 11-deoxycortisol (10.7% (4.5–32.7%) vs. 11.0% (5.1–29.7%), respectively).

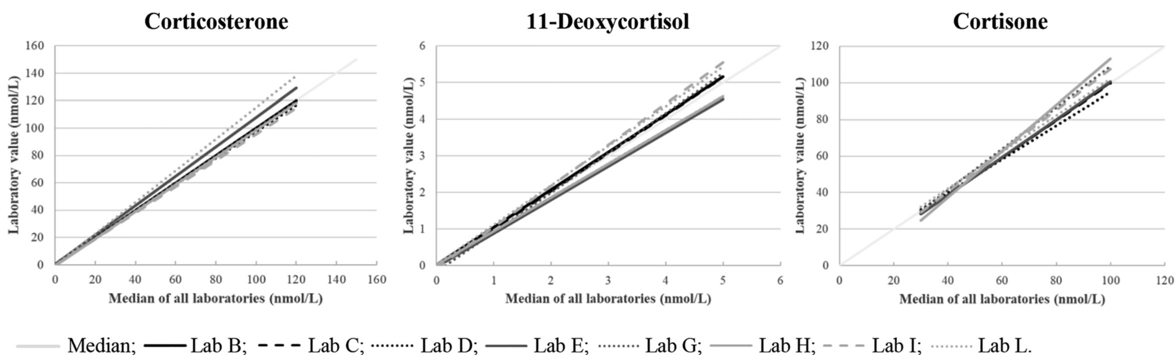

Figure 2: Passing-Bablok regression of laboratory steroid measures vs. median of all laboratories when using in house calibration. Slopes were not statistically different from 1 in four laboratories for corticosterone and cortisone and in two laboratories for 11-deoxycortisol.

Laboratory L exhibited the largest median bias, while Laboratory C exhibited the largest variance, with both calibrations. Cases above TAE were detected in Laboratory C and L with in house and in Laboratory C, H and L with external calibration (Figure 3). Compared to in house, the external calibration significantly increased the median bias in five laboratories, but reduced the variance in one laboratory (Supplementary Table 6).

11-Deoxycortisol

Concentrations in plasma and serum samples ranged from 0.20 to 3.91 nmol/L (median of all laboratories by in house calibration). Values from one or more laboratories were below the in house and/or the external calibration ranges in 15 samples, except Laboratory G, for which 65 samples were below the LLOQ (Supplementary Table 3). The intra-laboratory duplicate measurement CV ranged from 3.3 to 11.8% (Table 2). The CV of individual duplicates, listed in Supplementary Table 3, substantially exceeded the MAI in Laboratory C (11.1% of cases), D (26.9% of cases), E (32.4% of cases) and L (36.1% of cases). As shown in Table 2, compared to the in house, the external calibration determined modest but significant deviations within six laboratories (−4.9 to 10.5%; p-values from <0.001 to 0.023). Passing-Bablok comparison of results from the two calibrations within laboratories that used Chromsystems calibrators for in house calibration detected strong consistency in Laboratory L and modest deviations in Laboratory D and E (Supplementary Figure 1).

With use of the in house set, the median inter-laboratory CV was 10.7% and cases with interlaboratory CV >MAI were 50.7%. Use of external calibration did not significantly modify the inter-laboratory CV (Figure 1).

As shown in Figure 2, Passing-Bablok slopes of laboratory values vs. medians of all laboratories were similar to 1 in two laboratories with in house and in three with external calibration; coefficients (95CI) ranged from

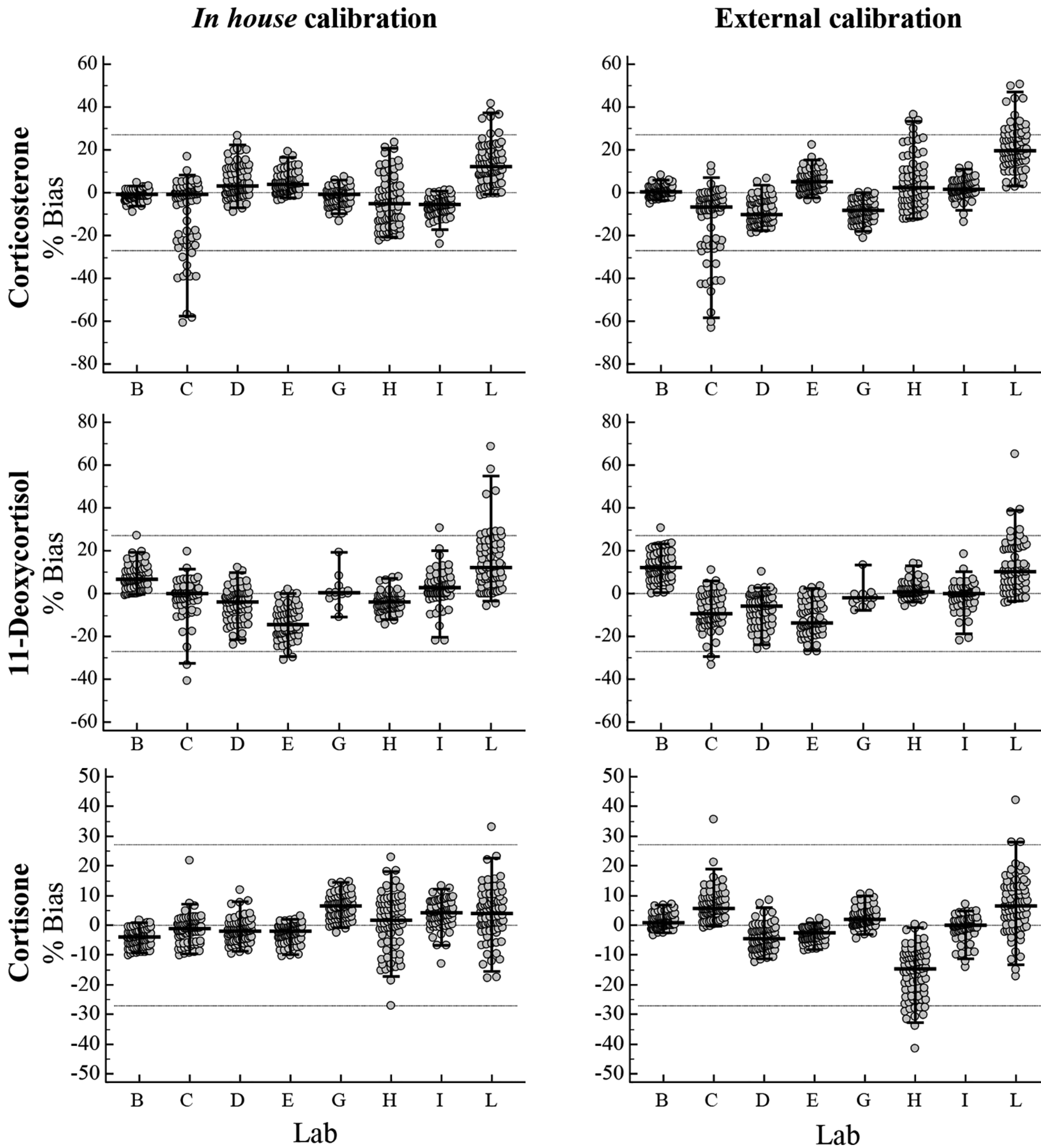

Figure 3: % Bias vs. median of all laboratories as function of the calibration system. *% Bias = ((laboratory value – median of all laboratories)/median of all laboratories) × 100*. Segments: mean; error bars: 95% agreement limits; dashed lines: zero ± total allowable error (27.1%). Median bias were all within allowable limits. However, bias variances were variable among laboratories, on occasion exceeding the allowable limits. Unifying the calibration system did not result in improving the inter-laboratory agreement.

0.895 (0.834–0.951) to 1.125 (1.101–1.158) with the former, and from 0.892 (0.857–0.931) to 1.137 (1.113–1.171) with the latter (Supplementary Table 5).

The median laboratory bias vs. median of all laboratories ranged from −14.6 to 12.4% with in house and −13.9 to 12.0% with external calibration (Figure 3; Bland-Altman plots of bias from individual laboratories as functions of hormone concentration are reported in Supplementary Figure 3). Laboratory E showed the largest bias, whereas Laboratory L showed the largest variance with both calibrations. A few samples exceeded the TAE in four laboratories with both calibrations (Figure 3). Compared to in house, use of external calibration significantly reduced the median bias in four laboratories, but significantly

increased bias in the others; variance was improved in one laboratory (Supplementary Table 6).

Cortisone

Concentrations in plasma samples ranged from 38.1 to 75.6 nmol/L (median of all laboratories by in house calibration). The intra-laboratory duplicate measurement CV ranged from 2.7 to 12.8% (Table 2). A substantial number of samples showed a CV of individual duplicates above the MAI in Laboratory C (16.7% of cases), H (19.2% of cases) and L (23.1% of cases) (Supplementary Table 4). As shown in Table 2, change of calibration determined significantly different results within all laboratories ($p<0.001$), with external calibration determining −15.8 to 9.9% deviation compared with in house calibration. Passing-Bablok comparisons of results from two calibrations within laboratories that used Chromsystems calibrators for in house calibration were strongly consistent (Supplementary Figure 1).

The median inter-laboratory CV of 6.2% observed with in house significantly increased to 8.1% with external calibration ($p>0.001$), while cases >MAI increased from 2.6 to 26.9%, respectively (Figure 1).

As shown in Figure 2, Passing-Bablok slopes of laboratory values vs. medians of all laboratories were similar to 1 in four laboratories with in house and in four laboratories with external calibration; coefficients (95CI) ranged from 0.916 (0.875–0.958) to 1.264 (1.149–1.395) with the former, and from 0.869 (0.829–0.920) to 1.047 (0.998–1.104) with the latter (Supplementary Table 5).

The median laboratory bias vs. medians of all laboratories ranged from −4.0 to 6.5% with in house and from −14.8 to 6.6% with external calibration (Figure 3; Bland-Altman plots of bias from individual laboratories as functions of the hormone concentration are reported in Supplementary Figure 4). Laboratory H and L showed the largest variance with both calibrations. Bias were mostly within TAE, except for Laboratory H with the external calibration (Figure 3). Use of external instead of in house calibration significantly worsened the median bias in four, but improved the variance in three laboratories (Supplementary Table 6).

Discussion

Our study evaluated the harmonization status for measurements of corticosterone, 11-deoxycortisol and cortisone according to different LC-MS/MS methods and use of plasma and serum specimens. MAI and TAE were used to interpret within- and between-method performance, with the restriction that these parameters were calculated from data about biological variability available only for 11-deoxycortisol, and not corticosterone and cortisone, and obtained from just two studies that used immunoassays [28]. Unfortunately, the availability of information on within- and between-subject variability for steroids generated by techniques with adequate specificity remains limited. Nevertheless, there are some studies from which between-subject variability can be derived [32–34]. Therefore, we recommend the provision of such information within the European Federation of Clinical Chemistry and Laboratory Medicine database for biological variation [30], so that appropriate performance limits can be established.

Methods investigated here are heterogeneous according to many procedural and technical aspects. Moreover, measurement ranges and analytical sensitivity differed among methods. The reasons for these differences may reflect several factors: 1. variable reporting purposes, such as a focus on the whole pathophysiological range or only high pathological levels; 2. the MS equipment available within each laboratory; and 3. the complexity of sample processing employed before MS analyses. For example, sensitivity for 11-deoxycortisol measurement is limited for Laboratory G, according to use of a simple protein precipitation and associated need for low sample volume.

Intra-assay imprecision was overall satisfactory for LDTs, even though suboptimal performance was displayed by some laboratories for 11-deoxycortisol and cortisone. As discussed in Supplementary Discussion, peak definition plays a role in intra-assay imprecision [35]. As previously recognized, the imprecision of Laboratory L, the only commercial kit user, was larger than expected due to initial inexperience, during the harmonization study, with maintenance of the MS ionization source [22]. Therefore, we recommend that laboratories that employ LC-MS/MS should have skilled staff and robust standard operating procedures for instrument maintenance and method monitoring.

Changing in house with external calibration determined a negligible deviation within most of the methods, with a few exceptions up to 15%. For laboratories that use Chromsystems' as in house calibrators, this may reflect variability in handling of materials or lot-to-lot variability. At variance, for laboratories using gravimetrically determined standards or CRM diluted in surrogated matrices, deviations may depend on the accuracy of the dilution and on commutability of matrices. Due to the endogenous nature of steroids, none of the calibration matrices used in this study consisted of native serum or plasma. Even though participating laboratories and Chromsystems

verified the adequacy of their calibrators by in house procedures, we cannot exclude that a certain degree of intra- and inter-laboratory variability is related to suboptimal commutability of surrogate matrices.

The overall inter-laboratory performance was mostly within acceptable specifications. However, imprecision above the estimated allowable limit was found in a consistent number of samples for corticosterone and 11-deoxycortisol. The bias vs. all laboratory median exceeded the estimated allowable limit only in a few cases. Nonetheless, large variance of bias was shown by some laboratories.

Unified calibration did not substantially improve the comparability of 11-deoxycortisol, and even worsened the consistency of corticosterone and cortisone measurements among laboratories. Such a finding indicates that, in this study, calibration apparently has a minor impact on overall inter-laboratory agreement compared to other factors, such as sample preparation, matrix interference, LC and IS, which can influence not only sample measurements but also the exactitude and commutability of calibrators. This contrasts with what we previously reported for cortisol, 17OH-progesterone and, to a minor extent, aldosterone [22], but is in line with the study by Owen et al. [36] focusing on testosterone.

The reasons for good consistency in some samples and suboptimal in others may reflect the features of individual methods. Here, baseline LC separation among the isobars corticosterone, 11-deoxycortisol and 21-deoxycortisol was ensured by all methods. Nonetheless, we could speculate that the diverse nature of LC methods employed among laboratories may confer susceptibility of some to other interfering compounds, which may explain the disagreement occasionally found among measurements [37].

Though speculative, a causative relationship may be suspected for 11-deoxycortisol since this analyte showed the poorest inter-laboratory performance and is also that for which four different isotopically-labeled ISs were used among the eight laboratories; this compares to two and three different IS respectively used for corticosterone and cortisone. As covered in Supplementary Discussion, IS labeling can influence analyte quantitation by multiple mechanisms [38]. With a focus on testosterone, Owen et al. [39] described how use of D2-, D5- or ^{13}C3-isotiopically-labeled IS could influence results up to 15%. However, Loh et al. [40] found that changing D2- with ^{13}C3-isotope did not influence measurements of 17OH-progesterone. Such contrasting findings reinforce the concept that the ability of the IS to correct for variation in analyte detection largely depend on the complex interaction among multiple method factors.

As for corticosterone, Laboratory C displayed the largest variance of bias. Intriguingly, this laboratory used a water loss (−18 amu) Q1/Q3 transition for the MRM detection of d8-corticosterone, which may be prone to specificity issues. The same transition was used by Laboratory D, which, conversely, did not show the same problem. All laboratories used the same Q1/Q3 transition for cortisone quantitation except Laboratory H, which exhibited a large bias.

Finally, another possible source of variability is the vacuum tube used to collect blood. Indeed, gel containing tubes have been previously shown to affect steroid concentrations [41]. A strength of our project is that we obtained paired blood samples collected with three different tubes, whose impact on steroid measurement will be investigated in a dedicated study.

All the above-mentioned potential reasons for differences in measurements among laboratories are speculative. Confirmation can only be attained by studies purposely designed to test the influence of various ISs and of Q1/Q3 transitions on analyte quantitation, as well as the efficacy of novel technologies empowering LC resolution in the separation of interfering species. Such studies could help to identify and correct causes of inaccuracy and would result in improved measurement harmonization. Ideally, technical issues should be resolved before entry of assays into the clinical routine. Until analyte- or panel-specific guidance is provided to optimize LC-MS/MS performance, recommendations can only be directed toward careful method validation and EQA scheme participation. Over the time period of this study, only two laboratories participated in EQA programs for serum corticosterone and 11-deoxycortisol. To date, only a few EQA schemes are available for these two analytes, but not for serum cortisone, and materials distributed are not traceable. EQA schemes based on traceable matrix-based quality controls should be implemented for these and other steroids that enter the clinical arena [21, 42, 43]. In such a frame, our study represented aunique opportunity for laboratories to verify their performance.

All centers are ISO accredited, however, only three laboratories have the LC-MS/MS measurement of serum corticosterone, 11-deoxycortisol and cortisone specifically included in the accreditation list. Of note, subdividing the consortium according to the clinical or research function did not help elucidating reasons for the observed inter-laboratory agreement (data not shown).

As previously recognized, a limitation of our study is that no reference method was involved. Therefore, we cannot establish the exact accuracy of individual methods. Furthermore, the sample size was not entirely compliant with guidelines [42]. In addition, our study focused on

basal blood collections. Future studies could be enlarged to dynamic tests to yield low and high levels of adrenal steroids.

In conclusion, our investigations of interlaboratory comparability of LC-MS/MS measurements of corticosterone, 11-deoxycortisol and cortisone indicated reasonable agreement among methods. Failure to improve harmonization with a common calibration system indicates that identification of other sources of disagreement, such as method imprecision, accuracy and specificity, may provide solutions to improve inter-laboratory performance. Our results suggest that harmonization of steroid measurements by LC-MS/MS is feasible. Such harmonization may facilitate correction of current shortcomings in assay performance that hamper the reproducibility of scientific findings, confound diagnoses or therapeutic monitoring, and prevent the establishment of consensus reference intervals.

Acknowledgments: We thank Dr. Oliver Midasch from Chromsystems for donating calibration materials used in the study.
Research funding: This study was supported by the Emilia-Romagna Region, Alessandro Liberati Young Researcher Grants (project number PRUA 1-2012-004, granted to FF) and by Deutsche Forschungsgemeinschaft (DFG), German Research Foundation (project number 314061271-CRC/TRR 205, granted to MP and GE).
Author contributions: FF conceived, designed and coordinated the study, performed the statistical analysis and wrote the manuscript; MC, MV and UP contributed to the study design; MC, MM, JML, MP, SB, MTA, JVDO and DK carried out sample measurement and data exports; AT coordinated subject recruitment and sample management; EN performed the statistical analysis; UP conceived the study; MC, MP, SB, PAB, MTA, ACH, JVDO, MR, GE, MV and UP contributed in result interpretation and in writing the manuscript. All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: Authors state no conflict of interest.
Informed consent: Informed consent was obtained from all individuals included in this study.
Ethical approval: The study received approval by the local Ethics Committee (no. 141/2017/U/Tess).

References

1. Olesti E, Boccard J, Visconti G, González-Ruiz V, Rudaz S. From a single steroid to the steroidome: trends and analytical challenges. J Steroid Biochem Mol Biol 2021;206:105797.
2. Fanelli F, Di Dalmazi G. Serum steroid profiling by mass spectrometry in adrenocortical tumors: diagnostic implications. Curr Opin Endocrinol Diabetes Obes 2019;26:160–5.
3. Athimulam S, Grebe S, Bancos I. Steroid profiling in the diagnosis of mild and overt Cushing's syndrome. Best Pract Res Clin Endocrinol Metabol 2021;35:101488.
4. Eisenhofer G, Masjkur J, Peitzsch M, Di Dalmazi G, Bidlingmaier M, Grüber M, et al. Plasma steroid metabolome profiling for diagnosis and subtyping patients with cushing syndrome. Clin Chem 2018;64:586–96.
5. Masjkur J, Gruber M, Peitzsch M, Kaden D, Di Dalmazi G, Bidlingmaier M, et al. Plasma steroid profiles in subclinical compared with overt adrenal cushing syndrome. J Clin Endocrinol Metab 2019;104:4331–40.
6. Hannah-Shmouni F, Berthon A, Faucz FR, Briceno JM, Maria AG, Demidowich A, et al. Mass spectrometry-based steroid profiling in primary bilateral macronodular adrenocortical hyperplasia. Endocr Relat Cancer 2020;27:403–13.
7. Di Dalmazi G, Fanelli F, Mezzullo M, Casadio E, Rinaldi E, Garelli S, et al. Steroid profiling by LC-MS/MS in nonsecreting and subclinical cortisol-secreting adrenocortical adenomas. J Clin Endocrinol Metab 2015;100:3529–38.
8. Di Dalmazi G, Fanelli F, Zavatta G, Ricci Bitti S, Mezzullo M, Repaci A, et al. The steroid profile of adrenal incidentalomas: subtyping subjects with high cardiovascular risk. J Clin Endocrinol Metab 2019;104:5519–28.
9. Eisenhofer G, Dekkers T, Peitzsch M, Dietz AS, Bidlingmaier M, Treitl M, et al. Mass spectrometry-based adrenal and peripheral venous steroid profiling for subtyping primary aldosteronism. Clin Chem 2016;62:514–24.
10. Williams TA, Peitzsch M, Dietz AS, Dekkers T, Bidlingmaier M, Riester A, et al. Genotype-specific steroid profiles associated with aldosterone-producing adenomas. Hypertension 2016;67: 139–45.
11. Wannachalee T, Turcu AF. Developments in primary aldosteronism subtyping using steroid profiling. Horm Metab Res 2020;52:373–8.
12. Taylor DR, Ghataore L, Couchman L, Vincent RP, Whitelaw B, Lewis D, et al. A 13-steroid serum panel based on LC-MS/MS: use in detection of adrenocortical carcinoma. Clin Chem 2017;63: 1836–46.
13. Schweitzer S, Kunz M, Kurlbaum M, Vey J, Kendl S, Deutschbein T, et al. Plasma steroid metabolome profiling for the diagnosis of adrenocortical carcinoma. Eur J Endocrinol 2019;180:117–25.
14. Berke K, Constantinescu G, Masjkur J, Kimpel O, Dischinger U, Peitzsch M, et al. Plasma steroid profiling in patients with adrenal incidentaloma. J Clin Endocrinol Metab 2022;107:e1181–92.
15. Costa-Barbosa FA, Carvalho VM, Nakamura OH, Bachega TA, Vieira JG, Kater CE. Zona fasciculata 21-hydroxysteroids and precursor-to-product ratios in 21-hydroxylase deficiency: further characterization of classic and non-classic patients and heterozygote carriers. J Endocrinol Invest 2011;34:587–92.
16. Costa-Barbosa FA, Carvalho VM, Oliveira KC, Vieira JGH, Kater CE. Reassessment of predictive values of ACTH-stimulated serum 21-deoxycortisol and 17-hydroxyprogesterone to identify CYP21A2 heterozygote carriers and nonclassic subjects. Clin Endocrinol (Oxf) 2021;95:677–85.
17. Kulle AE, Riepe FG, Hedderich J, Sippell WG, Schmitz J, Niermeyer L, et al. LC-MS/MS based determination of basal- and ACTH-stimulated plasma concentrations of 11 steroid hormones:

implications for detecting heterozygote CYP21A2 mutation carriers. Eur J Endocrinol 2015;173:517–24.
18. Yildiz M, Isik E, Abali ZY, Keskin M, Ozbek MN, Bas F, et al. Clinical and hormonal profiles correlate with molecular characteristics in patients with 11β-hydroxylase deficiency. J Clin Endocrinol Metab 2021;106:e3714–24.
19. Turcu AF, El-Maouche D, Zhao L, Nanba AT, Gaynor A, Veeraraghavan P, et al. Androgen excess and diagnostic steroid biomarkers for nonclassic 21-hydroxylase deficiency without cosyntropin stimulation. Eur J Endocrinol 2020;183: 63–71.
20. Oriolo C, Fanelli F, Castelli S, Mezzullo M, Altieri P, Corzani F, et al. Steroid biomarkers for identifying non-classic adrenal hyperplasia due to 21-hydroxylase deficiency in a population of PCOS with suspicious levels of 17OH-progesterone. J Endocrinol Invest 2020;43:1499–509.
21. Joint Committee for Traceability in Laboratory Medicine (JCTLM). Database of higher-order reference materials, measurement methods/procedures and services. Available from: https://www.bipm.org/jctlm/ [Accessed 10 Jan 2022].
22. Fanelli F, Cantù M, Temchenko A, Mezzullo M, Lindner JM, Peitzsch M, et al. Report from the HarmoSter study: impact of calibration on comparability of LC-MS/MS measurement of circulating cortisol, 17OH-progesterone and aldosterone. Clin Chem Lab Med 2022;60:726–39.
23. Fanelli F, Belluomo I, Di Lallo VD, Cuomo G, De Iasio R, Baccini M, et al. Serum steroid profiling by isotopic dilution-liquid chromatography-mass spectrometry: comparison with current immunoassays and reference intervals in healthy adults. Steroids 2011;76:244–53.
24. Mezzullo M, Fanelli F, Fazzini A, Gambineri A, Vicennati V, Di Dalmazi G, et al. Validation of an LC-MS/MS salivary assay for glucocorticoid status assessment: evaluation of the diurnal fluctuation of cortisol and cortisone and of their association within and between serum and saliva. J Steroid Biochem Mol Biol 2016;163:103–12.
25. Lindner JM, Vogeser M, Grimm SH. Biphenyl based stationary phases for improved selectivity in complex steroid assays. J Pharm Biomed Anal 2017;142:66–73.
26. Peitzsch M, Dekkers T, Haase M, Sweep FC, Quack I, Antoch G, et al. An LC-MS/MS method for steroid profiling during adrenal venous sampling for investigation of primary aldosteronism. J Steroid Biochem Mol Biol 2015;145:75–84.
27. Fahlbusch FB, Heussner K, Schmid M, Schild R, Ruebner M, Huebner H, et al. Measurement of amniotic fluid steroids of midgestation via LC-MS/MS. J Steroid Biochem Mol Biol 2015; 152:155–60.
28. Westgard QC. Desirable biological variation database specifications. Available from: https://www.westgard.com/biodatabase1.htm [Accessed 10 Jan 2022].
29. Oosterhuis WP, Bayat H, Armbruster D, Coskun A, Freeman KP, Kallner A, et al. The use of error and uncertainty methods in the medical laboratory. Clin Chem Lab Med 2018;56:209–19.
30. The European Federation of Clinical Chemistry and Laboratory Medicine database. Available from: https://biologicalvariation.eu/meta_calculations [Accessed 10 Jan 2022].
31. Hyslop NP, White WH. Estimating precision using duplicate measurements. J Air Waste Manag Assoc 2009;59:1032–9.
32. Eisenhofer G, Peitzsch M, Kaden D, Langton K, Pamporaki C, Masjkur J, et al. Reference intervals for plasma concentrations of adrenal steroids measured by LC-MS/MS: impact of gender, age, oral contraceptives, body mass index and blood pressure status. Clin Chim Acta 2017;470:115–24.
33. Mezzullo M, Di Dalmazi G, Fazzini A, Baccini M, Repaci A, Gambineri A, et al. Impact of age, body weight and metabolic risk factors on steroid reference intervals in men. Eur J Endocrinol 2020;182:459–71.
34. Mezzullo M, Gambineri A, Di Dalmazi G, Fazzini A, Magagnoli M, Baccini M, et al. Steroid reference intervals in women: influence of menopause, age and metabolism. Eur J Endocrinol 2021;184: 395–407.
35. Wahab MF, Dasgupta PK, Kadjo AF, Armstrong DW. Sampling frequency, response times and embedded signal filtration in fast, high efficiency liquid chromatography: a tutorial. Anal Chim Acta 2016;907:31–44.
36. Owen LJ, MacDonald PR, Keevil BG. Is calibration the cause of variation in liquid chromatography tandem mass spectrometry testosterone measurement? Ann Clin Biochem 2013;50:368–70.
37. Settlage J, Oglesby T, Rajasekaran A, Williard C, Scott G. The importance of chromatographic resolution when analyzing steroid biomarkers. Steroids 2015;99:45–8.
38. Vogeser M, Seger C. Pitfalls associated with the use of liquid chromatography–tandem mass spectrometry in the clinical laboratory. Clin Chem 2010;56:1234–44.
39. Owen LJ, Keevil BG. Testosterone measurement by liquid chromatography tandem mass spectrometry: the importance of internal standard choice. Ann Clin Biochem 2012;49:600–2.
40. Loh TP, Ho CS, Hartmann MF, Zakaria R, Lo CWS, van den Berg S, et al. Influence of isotopically labeled internal standards on quantification of serum/plasma 17α-hydroxyprogesterone (17OHP) by liquid chromatography mass spectrometry. Clin Chem Lab Med 2020;58:1731–9.
41. Hepburn S, Wright MJP, Boyder C, Sahertian RC, Lu B, Zhang R, et al. Sex steroid hormone stability in serum tubes with and without separator gels. Clin Chem Lab Med 2016;54:1451–9.
42. International Consortium for Harmonization of Clinical Laboratory Results. Toolbox of technical procedures to be considered when developing a process to achieve harmonization for a measurand; 2013. Available from: www.harmonization.net.
43. Greaves RF. The central role of external quality assurance in harmonisation and standardisation for laboratory medicine. Clin Chem Lab Med 2017;55:471–3.

Supplementary Material: The online version of this article offers supplementary material (https://doi.org/10.1515/cclm-2022-0242).