RESEARCH

Age- and sex-adjusted reference intervals for steroid hormones measured by liquid chromatography–tandem mass spectrometry using a widely available kit

Sonja Kunz[1], Xiao Wang[1], Uta Ferrari[1], Michael Drey[1], Marily Theodoropoulou[1], Katharina Schilbach[1], Martin Reincke[1], Margit Heier[2,3], Annette Peters[2,4], Wolfgang Koenig[4,5,6], Tanja Zeller[7,8], Barbara Thorand[2] and Martin Bidlingmaier[1]

[1]Department of Medicine IV, University Hospital, Ludwig Maximilian University Munich, Munich, Germany
[2]Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
[3]KORA Study Centre, University Hospital of Augsburg, Augsburg, Augsburg, Germany
[4]German Centre for Cardiovascular Research (DZHK), Partner site Munich Heart Alliance, Munich, Germany
[5]German Heart Centre Munich, Technical University of Munich, Munich, Germany
[6]Institute of Epidemiology and Medical Biometry, University of Ulm, Ulm, Germany
[7]Department of General and Interventional Cardiology, University Heart Center Hamburg, Hamburg, Germany
[8]German Centre for Cardiovascular Research (DZHK), Partner Site Hamburg/Kiel/Lübeck, Hamburg, Germany

Correspondence should be addressed to M Bidlingmaier: martin.bidlingmaier@med.uni-muenchen.de

Abstract

Objective: Measurements utilizing commercially available sets of reagents for determination of steroid hormone profiles by liquid chromatography–tandem mass spectrometry (LC-MS/MS) have become increasingly important for routine laboratories. However, method-specific publications of reference intervals obtained from sufficiently large studies are often missing.

Methods: After validation of performance characteristics, a widely available kit for steroid analysis by LC-MS/MS was used to measure concentrations of 15 endogenous steroids (aldosterone, cortisol, cortisone, corticosterone, 11-deoxycortisol, 21-deoxycortisol, dehydroepiandrosterone sulfate, estradiol, testosterone, androstenedione, dihydrotestosterone, dehydroepiandrosterone, 17-hydroxyprogesterone, 11-deoxycorticosterone, progesterone) in more than 500 blood samples from a population-based study. While randomly selected from a larger cohort, the samples equally represented both sexes and covered a wide range of adult age groups. Age- and sex-specific reference intervals were calculated, and correlation with BMI was assessed.

Results: Performance characteristics of the assay matched expectations for 9 of 15 steroids. For most of them, reference intervals obtained from our study population were comparable to those reported by others, with age and sex being the major determinants. A sex-specific correlation with BMI was found for seven steroids. We identified limitations regarding sensitivity of the method for quantification of progesterone in males and postmenopausal females. Concentrations of aldosterone, 21-deoxycortisol, estradiol, 11-deoxycorticosterone, and dihydrotestosterone could not be quantified in a large percentage of samples.

Conclusions: The reference intervals for nine steroids will support meaningful interpretation for steroid profiles as measured by a widely used kit for LC-MS/MS-based quantification. Laboratories using such kits must be aware of potential limitations in sensitivity for some steroids included in the profile.

Significance Statement

Quantification of steroid hormones is a cornerstone for diagnosis of several diseases. Commonly used immunoassays have limitations in specificity. Liquid chromatography–tandem mass spectrometry (LC-MS/MS) is a promising alternative, particularly if methods are harmonized across laboratories. The use of kits from commercial suppliers might support this. Clinical interpretation of steroid concentrations requires availability of appropriate reference intervals (RIs), but studies on RIs reported in the literature differ in preanalytical and analytical procedures. Here, we provide RIs for steroids measured by a widely available kit under preanalytical conditions mirroring common clinical practice. Such RIs might facilitate interpretation for those using the same method and comparable conditions in clinical routine.

Introduction

Steroid hormones are signaling molecules involved in regulation of electrolyte concentrations, blood pressure, reproductive functions, and metabolism (1). A lack or an excess of steroids results from enzyme defects, adenomas, or malignant tumors affecting endocrine glands and causes diseases like Cushing syndrome, congenital adrenal hyperplasia, and primary aldosteronism (2, 3, 4). Diagnosis of such diseases requires an accurate quantification of steroid hormone concentrations (5). Several analytical methods are available to clinical laboratories, with immunoassays being most widely used today. However, specificity of many immunoassays is limited, and patient samples contain a wide spectrum of endogenous steroids with similar chemical structures and physical properties, circulating at very different concentrations (6, 7). More recently, measurements by liquid chromatography (LC) coupled to mass spectrometry (MS) became more widely available. It offers improved specificity by combining chromatographic separation with specific fragmentation of molecules and detection (1, 8). Unfortunately, despite undoubted advantages, results from different laboratories using different LC-MS/MS-based methods still can exhibit variability. Apart from differences in the preanalytical procedures, this is explained by differences in extraction protocols, LC conditions, and MS techniques but also by differences in calibration (9, 10, 11). Therefore, because robust RIs from an appropriately characterized background population are a prerequisite for clinical interpretation of measurement results, it remains important to establish method-specific RIs even for LC-MS/MS methods (12, 13).

The use of 98/79/EC (IVDD) certified calibrators, quality controls (QCs), and reagents in combination with uniform measurement procedures and preanalytical procedures might improve between-laboratory comparability (1, 9) and potentially allow for the application of common RIs. In recent years, some manufacturers have started to market such ready-to-use kits for quantification of steroid hormone profiles by LC-MS/MS. However, published RIs for such kits from well-characterized study populations are scarce or available only for selected steroids included in the respective kits.

We examined the analytical performance of a more widely used, commercially available kit for simultaneous quantification of 15 steroids by LC-MS/MS and determined limits of quantification during routine use. Using this method, we analyzed samples from the population-based Cooperative Health Research in the Region of Augsburg (KORA) studies; generated method-specific RIs; and investigated the impact of sex, age, and BMI.

Materials and methods

Subjects

Subjects were participants of the population-based KORA F4 and KORA Age studies. Study design and sampling methods are described elsewhere (14, 15, 16, 17). For the present analysis, 585 samples were randomly selected, stratified by sex and 10-year age groups. We excluded 63 subjects because of hormone intake and 3 subjects with a BMI >45 kg/m^2, resulting in a final study population of 290 male and 229 female subjects. Age ranged from 32 to 90 years (median 65 years in females, 59 years in males). The study population was selected to reflect the general population in terms of several health-related factors (antihypertensive or lipid-lowering drugs, physical activity, diabetes, smoking, cardiovascular diseases, waist circumference,

and alcohol consumption) (18, 19). The general characteristics of the study population and prevalence of common cardiovascular risk factors are shown in Table 1. Categorization for waist circumference and alcohol consumption followed published recommendations (21, 21). In females, information about the day of menstrual cycle was not available. Pre- or postmenopausal status was determined based on interview questions, age (>60 years), and measurements of estradiol.

Written informed consent was obtained from all participants, and the study was approved by the ethics committee of the Bavarian Chamber of Physicians (Munich) and conducted in accordance with the Declaration of Helsinki.

Preanalytical procedure

Samples were taken in seated position after at least 10 min rest. For participants from KORA F4 (subjects between 32 and 61 years), sampling was done following an overnight fast between 07:30 h and 10:30 h in the morning (15). Due to the advanced age of participants in the KORA Age study (subjects aged 65 to 90 years), samples were taken under nonfasting conditions, with time ranging from morning hours for most participants to early afternoon for some. Generally, EDTA plasma was used (Sarstedt Monovette K EDTA). Due to restrictions in available volumes, however, serum samples were analyzed from KORA Age (Sarstedt Monovette with gel separator). After collection, all samples were centrifuged and supernatants were immediately stored at −80°C until analysis (16).

Laboratory measurements

Steroid profiles were measured using a commercially available, 98/79/EG (IVDD) certified, kit (MassChrom® (72072/96), Chromsystems, Gräfelfing, Germany) following the manufacturer's instructions for sample preparation, LC column and conditions and validated multiple reaction monitoring (MRM) mass transitions. According to the manufacturer, all calibrators and QCs are traceable to certified reference materials, or, if not available for certain steroid, to primary standards. For cortisol, progesterone and testosterone NIST 971 reference material is used to assign values for calibration. Calibrators and QCs are based on human serum, delivered in lyophilized form and dissolved on site. 500 µL of samples, calibrator solutions and QCs were mixed with internal standards (72044, Chromsystems) and loaded to an SPE plate with extraction buffer (72005, Chromsystems) via centrifugation. Following two washing steps and elution in 500 µL elution buffer (72033, Chromsystems), the extracts were evaporated to dryness under a stream of nitrogen at 50°C and dissolved in 100 µL reconstitution buffer (72006, Chromsystems). Seven different concentrations

of calibrators (72038 and 72039, Chromsystems) and three QC samples at different concentration levels (0341–0347, Chromsystems) were measured with each batch. Calibrations were only accepted with a coefficient of determination (R^2) >0.99. For the first 15 measurements of each QC lot from Chromsystems, we used the target value and range from the manufacturer's instructions for use (<30% from the target value for QC level 1, <20% from the target value for QC levels 2 and 3). After 15 measurements of the same QC lot, we reduced the acceptable deviation to <15% from the target value (based on the recommendation for interassay precision of the European Medicines Agency (EMA) guideline (23)). In addition, we included three QC samples from an independent supplier (361–363, Liquichek, Bio-Rad Laboratories). Notably, the manufacturer of these control samples does not provide target values established by a method of higher methodological order but only method-specific targets for several commercially available immunoassay. Therefore, we only used these QC samples as an additional control for method precision. Following the acceptance criterion for inter assay precision specified in the EMA guideline (23), we accepted a deviation <15% from the overall mean of our assay runs. Accuracy of the method was assessed by participation in the national external quality assessment scheme for steroid hormones (Reference Institute for Bioanalytics, RfB, Bonn, Germany). Unfortunately, the scheme does only include a limited number of steroids (aldosterone, cortisol, estradiol, progesterone, testosterone, DHEAS, 17-OHP). We report accuracy data on the respective steroids in Supplementary Table A2 (see section on supplementary materials given at the end of this article). Measurements were performed in two panels (panel 1: aldosterone, cortisol, cortisone, corticosterone, 11-deoxycortisol, 21-deoxycortisol; panel 2: dehydroepiandrosterone sulfate (DHEAS), estradiol, testosterone, androstenedione, dihydrotestosterone (DHT), dehydroepiandrosterone (DHEA), 17-hydroxyprogesterone (17-OHP), 11-deoxycorticosterone, and progesterone). An injection volume of 20 µL and a gradient elution with mobile phase A (72011, Chromsystems) and B (72002, Chromsystems) was applied throughout.

Chromatography was performed on the 1290 Infinity II HPLC System (Agilent Technologies) coupled to a QTrap 6500+ tandem mass spectrometer (Sciex, Framingham, MA, USA). Chromatography columns were included in the commercially available kit, but the manufacturer does not provide details about dimensions or stationary phase. Only minor adaptations of the gradient profiles specified by the manufacturer were required to optimally adjust to our system. Ionization was performed by negative electrospray ionization (ESI) for aldosterone and DHEAS and positive ESI for all other steroids. Multireaction monitoring was used as measurement mode. Recommended MRM transitions were selected and decimals were tuned according to manufacturer's instructions. Chromatographic data acquisition and

Table 1 Descriptive statistics of sex, number (*n*), age (years), and BMI (kg/m^2) and general characteristics for 519 healthy subjects. Age is shown as median and minimum–maximum. BMI is shown as mean and minimum–maximum values.

	Males					Females		
	Total	30–39 years	40–49 years	50–59 years	60–90 years	Total	Premenopause	Postmenopause
n	290	34	60	53	143	229	80	149
Age	59 (32–87)	37 (32–39)	44.5 (40–49)	55 (50–59)	72 (60–87)	65 (32–90)	43 (32–54)	72 (50–90)
BMI	27.6 (18.8–40.4)	26.3 (21.4–36.6)	26.4 (19.2–34.8)	27.7 (19.0–40.3)	28.4 (18.8–40.4)	27.8 (18.2–41.5)	26.2 (18.2–40.7)	28.7 (19.2–41.5)
Lipid lowering drugs[a]	45 (15.5%)	0 (0%)	4 (6.7%)	9 (17%)	32 (22.4%)	37 (16.2%)	1 (1.3%)	36 (24.2%)
Antihypertensives[b]	107 (36.9%)	0 (0%)	8 (13.3%)	15 (28.3%)	84 (58.7%)	91 (39.7%)	4 (5%)	87 (58.4%)
Hypertension[c]	136 (46.9%)	3 (8.8%)	14 (23.3%)	15 (28.3%)	104 (72.7%)	107 (46.7%)	7 (8.8%)	100 (67.1%)
Physical active[d]	175 (60.3%)	21 (61.8%)	38 (63.3%)	27 (50.9%)	89 (62.2%)	129 (56.3%)	46 (57.5%)	83 (55.7%)
Diabetes[e]	22 (7.6%)	0 (0%)	1 (1.7%)	3 (5.7%)	18 (12.6%)	20 (8.7%)	0 (0%)	20 (13.4%)
Cardiovascular event[f]	34 (11.7%)	0 (0%)	3 (7.5%)	4 (7.5%)	27 (18.9%)	17 (3.7%)	0 (0%)	17 (11.4%)
Smoking[g]	54 (18.6%)	12 (35.3%)	17 (28.3%)	17 (28.3%)	13 (9.1%)	26 (11.4%)	14 (17.5%)	12 (8.1%)
Waist circumference > cutoff[h]	192 (66.2%)	13 (38.2%)	28 (46.7%)	33 (62.3%)	118 (82.5%)	76 (33.2%)	15 (18.8%)	61 (40.9%)
Alcohol consumption > cutoff[i]	106 (36.6%)	5 (14.7%)	22 (36.7%)	20 (37.7%)	59 (41.3%)	23 (10%)	9 (11.3%)	14 (9.4%)

[a]Excl. vegetable substances, omega-3 fatty acid; [b]Classification according to recommendations of the German League; [c]Current hypertension (≥140/90 mm Hg) or known hypertension controlled by medication according to ISH-WHO 1999; [d]Questionnaire; [e]Diabetes mellitus; [f]Myocardial infarct and stroke; [g]Regular smokers and irregular smokers; [h]Cutoff for increased risk of metabolic complications associated with obesity according to WHO report: >94 cm for males and >80 cm for females (21); [i]Recommendation for healthy diet characteristics from *European Heart Journal*: alcohol consumption >20 g/day for males and >10 g/day for females (21).

processing was performed using Analyst 1.6.3 and Sciex OS 1.6.1 (Sciex, Framingham, MA, USA). MRM transitions and retention times of all quantifier transitions are summarized in Supplementary Fig. A1, Supplementary Fig. A2 and Supplementary Table A1.

Sex hormone-binding globulin (SHBG) concentrations were measured using an automated chemiluminescence immunoassay (Cobas e14, Roche).

Calculations and statistical methods

When concentrations of a specific steroid were below the limit of quantification (LoQ) in less than 10% of the samples, we arbitrarily set the concentration to the corresponding LoQ in those samples (Tables 2 and 3), and included this value in the calculation of RIs. We did not calculate RIs for steroids where more than 10% of samples had concentrations below the LoQ as determined in our laboratory (details see results and Table 3). Six Steroids fell into this group. Among those, three steroids (aldosterone, estradiol, and progesterone in males and postmenopausal females) had clearly detectable, baseline separated signals when concentrations were above LoQ. For these steroids, we report the maximum values observed in our cohort. In contrast, chromatograms for 21-deoxycortisol, DHT, and 11-deoxycorticosterone were difficult to interpret for a variety of reasons explained below. Therefore, from our analyses, we cannot report RIs or maximum values for these steroids.

Logarithmic, square root, cube root and Box–Cox transformations were used to achieve approximate normality for steroids with less than 10% of

Table 2 LoQs (nmol/L) derived from in-house testing vs manufacturer's claims.

Steroid	LoQ (in-house)	LoQ (manufacturer)
Aldosterone	0.078	0.039
Cortisol	3.380	4.190
Cortisone	0.283	0.411
Corticosterone	0.231	0.505
11-Deoxycortisol	0.416	0.087
21-Deoxycortisol	0.121	0.078
DHEAS	39	66
Estradiol	0.015	0.092
Testosterone	0.035	0.017
Androstenedione	0.112	0.080
DHEA	0.738	0.794
11-Deoxycorticosterone	0.024	0.070
Dihydrotestosterone	0.279	0.145
17-OHP	0.045	0.121
Progesterone	0.111	0.095

In-house testing: Sciex QTrap 6500+ mass spectrometer; manufacturer: Sciex QTrap 5500 mass spectrometer.
17-OHP, 17-hydroxyprogesterone; DHEA, dehydroepiandrosterone; DHEAS, dehydroepiandrosterone sulfate; LoQ, limit of quantification.

concentrations below LoQ (cortisol, cortisone, corticosterone, 11-deoxycortisol, DHEAS, testosterone, androstenedione, DHEA, 17-OHP). The histogram and Shapiro–Wilk test were used to confirm normal distribution. Reference intervals are given as upper and lower limit of the 95% CI, which in this case corresponds to 2.5 and 97.5 percentiles. Progesterone concentrations in premenopausal females were not normally distributed, even after transformation. Accordingly, median, minimum and maximum concentrations are reported instead of mean and 95% CI. For calculating the mean and nth percentile for steroid concentrations that changed with age, male participants were grouped into the following age categories: 30–39, 40–49, 50–59, and 60–90 years; female participants were grouped into pre- and postmenopausal. The nth percentile of steroid concentrations was calculated based on the following formula: $y_n = \text{mean}_n(\text{age}) \pm 1.96 \times \text{s.d.}_n(\text{age})$. Mean$_n$(age) is the mean, s.d.$_n$(age) is the standard deviation of the corresponding group.

The graphical presentation of continuous age-adjusted RIs for steroids was based on multivariate fractional polynomial analyses (23, 24, 25), using age transformation for higher ages as recommended by Royston and Wright (23, 25). After best-fit polynomial models between normalized steroid concentrations and transformed age were estimated, the nth age-specific RIs were calculated by reversing to the original scale.

Sex-specific correlations of BMI to steroid concentrations were assessed using Spearman's rank correlation coefficients with $P < 0.05$ considered significant. Concentrations between groups of subjects classified as normal weight, overweight and obese according to the WHO classification (normal weight = 18–24.9 kg/m², overweight = 25–29.9 kg/m², obese >30 kg/m²) (26) were compared by Kruskal–Wallis test followed by pairwise Wilcoxon rank sum test with Bonferroni correction. P-values below 0.05 were considered significant. All the statistical analyses were conducted in R studio (R version 3.6.2).

The free androgen index (FAI) was calculated by the following formula (c = concentration in nmol/L): FAI = (c(testosterone)/c(SHBG)) × 100.

Quality control and performance characteristics

QC samples provided with the kit and by an independent supplier (Bio-Rad Laboratories) were included with each analytical run. Performance characteristics of the assay as implemented in our laboratory were evaluated according to the EMA guideline (23). Intra-assay reproducibility was assessed by measuring five aliquots of each QC level within one assay. One aliquot of each QC was measured on 12 different days with independent calibrations to determine inter assay reproducibility. The LoQ for each steroid was formally assessed using

Table 3 Percentage of samples with concentrations below LoQ for selected steroids.

	Males (*n* = 290)	Females, premenopausal (*n* = 80)	Females, postmenopausal (*n* = 149)
	<LoQ	<LoQ	<LoQ
Dihydrotestosterone	17%	83%	62%
Estradiol	60%	29%	50%
21-Deoxycortisol	47%	50%	38%
Aldosterone	43%	31%	36%
11-Deoxycorticosterone	34%	44%	13%
Progesterone	28%	3%	51%
Testosterone	3%	3%	1%
11-Deoxycortisol	3%	0%	4%
17-Hydroxyprogesterone	0%	0%	2%
Cortisol	0%	0%	0%
Cortisone	0%	0%	0%
Corticosterone	0%	0%	0%
DHEAS	0%	0%	0%
DHEA	0%	0%	0%
Androstenedione	0%	0%	0%

DHEA, dehydroepiandrosterone; DHEAS, dehydroepiandrosterone sulfate; LoQ, limit of quantification.

serial dilutions of QC samples in a physiological sodium chloride solution. Five aliquots of each dilution were measured, and the lowest concentration which could be measured with a coefficient of variation <20%, and a deviation of <20% from the nominal value (according to the manufacturer) was defined as LoQ.

Results

Analytical performance of the assay

In our hands, following the protocol recommended by the manufacturer, the LoQ of the assays was in agreement with the manufacturers claim for five steroids (cortisol, cortisone, androstenedione, DHEA, and progesterone). For five of the steroids, the LoQ in our laboratory was superior to the manufacturer's claims (corticosterone, DHEAS, estradiol, 11-deoxycortisosterone, and 17-OHP), while we could not reach the claimed LoQ for aldosterone, 11-deoxycortisol, 21-deoxycortisol, testosterone, and DHT (Table 2). At concentrations above the LoQ, the intra- and interassay reproducibility as determined in our laboratory was excellent and in line with the manufacturer's claims (see Supplementary Table A3).

While most steroids could be measured in all samples from our reference population, concentrations were below the LoQ in more than 10% of the samples for progesterone in males and postmenopausal females and for DHT, estradiol, 21-deoxycortisol, aldosterone, and 11-deoxycorticosterone in both sexes (Table 3).

Reference intervals

For cortisol, cortisone, corticosterone, and 11-deoxycortisol, we established RIs independent of age, even though these steroids showed a slight decrease in concentration against age. Overall, our analyses demonstrated comparable concentrations throughout, and the slight decrease seems clinically irrelevant. The impact of sex is small (Table 4, Supplementary Table A4).

Concentrations of DHEAS, DHEA, and androstenedione were associated with age (Fig. 1, Table 5, Supplementary Table A4). In males, the upper limit of the reference interval decreased from the youngest to the oldest age group by 73% for DHEA and 44% for DHEAS and androstenedione. In females, the upper normal limit was 8%, 12%, and 46% lower after menopause for DHEA, DHEAS, and androstenedione, respectively. A similar pattern was also seen for age-/menopausal status- and

Table 4 Mean concentrations (nmol/L) with 2.5 and 97.5 percentiles for glucocorticoids and maximum concentrations (nmol/L) for aldosterone.

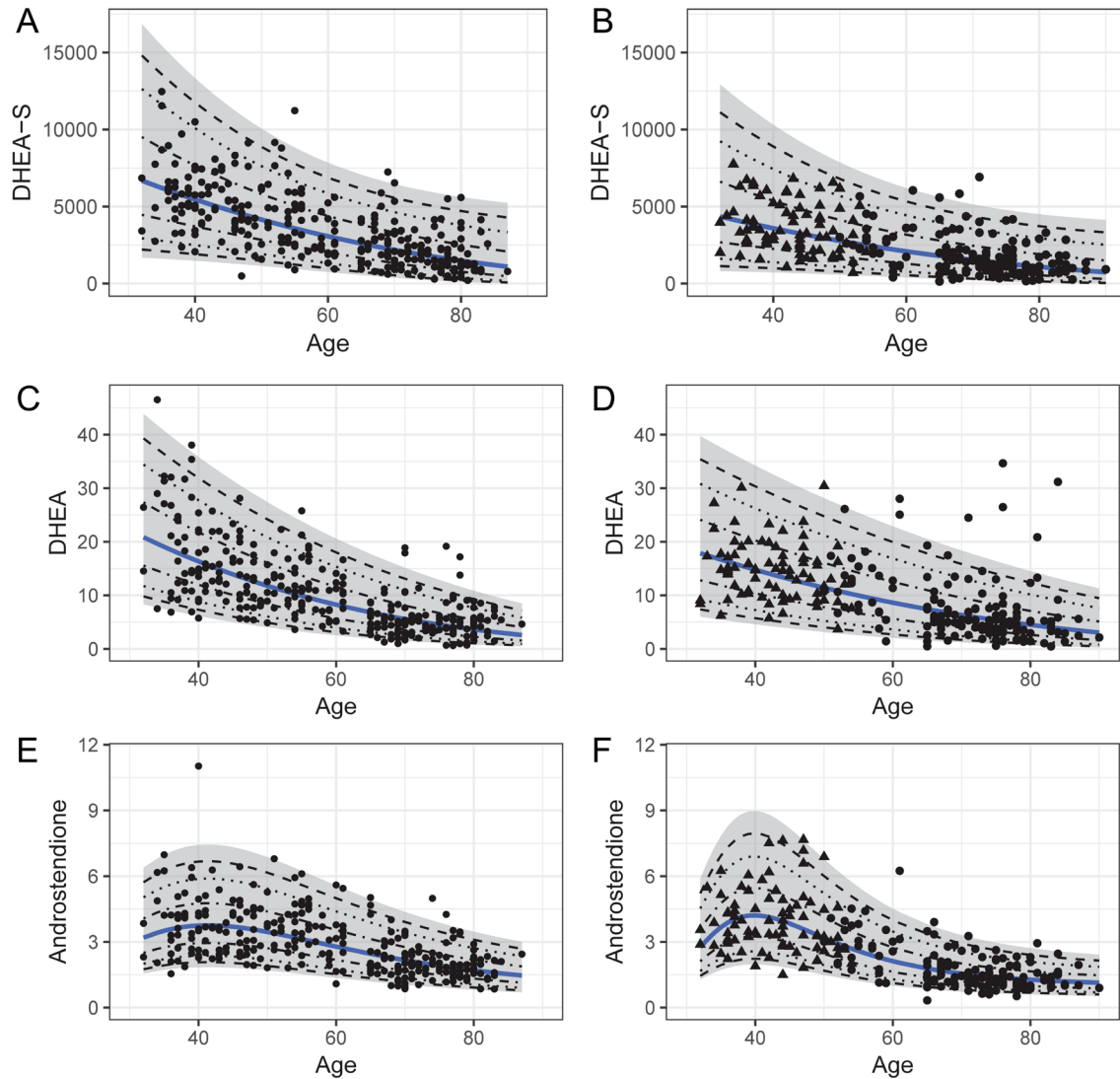|	Males (*n* = 290)		Females (*n* = 229)	
Aldosterone		<0.55		<0.47
Cortisol	350	(165–635)	323	(166–617)
Cortisone	53	(32–75)	53	(31–74)
Corticosterone	7.54	(1.18–26.65)	6.58	(1.54–24.95)
11-Deoxycortisol	0.76	(0.19–3.01)	0.62	(0.12–2.21)

Figure 1

Age dependent distribution of concentrations (nmol/L) for DHEAS (A, B), DHEA (C, D), and androstenedione (E, F) males (left) and females (right). For females, triangles represent premenopausal status and dots represent postmenopausal status. Mean concentrations estimated after transformation are represented by the blue line. The 95% CI is marked gray, whereas the dashed lines illustrate the 50%, 80%, and 90% CIs.

sex-specific RIs for 17-OHP (Table 5, Supplementary Fig. A4).

Concentrations of testosterone and SHBG were dependent on age in males (Table 5). While testosterone concentrations were highest in the youngest age group, but relatively stable thereafter, SHBG concentrations increased consistently with increasing age. Consequently, the FAI dropped with age (Fig. 2). In contrast, testosterone and SHBG concentrations or FAI did not differ by menopausal status in females (Table 5).

Next to testosterone, estradiol and progesterone exhibited the greatest difference between sexes, with concentrations in premenopausal females greatly exceeding those in men. In postmenopausal females, estradiol was mostly undetectable, and concentrations of progesterone were 99% lower than in premenopausal females (Table 5). In contrast, concentrations of estradiol and progesterone – when detectable – did not change with age in males.

In view of the obvious limitations of assay sensitivity for DHT, estradiol, 21-deoxycortisol, aldosterone, 11-deoxycorticosterone, and progesterone in males and postmenopausal females, we decided to not calculate RIs for these steroids. However, to provide at least some indication about the distribution of concentrations in a healthy population, we report maximum values for aldosterone, estradiol, and progesterone in males and postmenopausal females (Tables 4 and 5). In case of aldosterone, signal integration was possible for most samples, but concentrations often were close to LoQ. We

Table 5 Mean concentrations (nmol/L) with 2.5 and 97.5 percentiles for androgens, SHBG, and 17-OHP in males and females. For progesterone in premenopausal females (which was not normally distributed) median, minimum and maximum concentrations (nmol/L) are shown. For estradiol in all groups, and for progesterone in males and postmenopausal females, the method has insufficient sensitivity. Therefore, only the maximum value observed is listed.

	Males					Females	
	30–39 years (n=34)	40–49 years (n=60)	50–59 years (n=53)	60–90 years (n=143)	Males (n=290)	Premenopausal (n=80)	Postmenopausal (n=149)
DHEAS	5745 (2782–11864)	5108 (1230–8987)	3942 (986–8867)	1831 (502–6678)		3468 (1023–7362)	1340 (277–6488)
DHEA	18.39 (6.53–51.78)	13.35 (5.95–29.98)	10.58 (4.51–24.83)	5.40 (0.85–13.86)		13.61 (6.37–29.10)	5.53 (1.14–26.86)
Androstenedione	3.63 (1.67–7.86)	3.49 (1.80–6.79)	3.48 (1.90–6.37)	2.01 (0.93–4.38)		3.69 (1.81–7.54)	1.59 (0.62–4.06)
17-OHP	3.13 (1.02–5.24)	2.92 (1.37–6.23)	2.23 (0.86–5.80)	2.14 (0.56–4.76)		1.58 (0.29–8.73)	0.54 (0.13–2.16)
Testosterone	17.44 (9.02–33.70)	16.40 (8.95–30.06)	13.97 (6.34–30.76)	16.64 (8.18–28.07)		0.81 (0.37–1.79)	0.75 (0.25–2.23)
SHBG	43.69 (8.13–79.25)	41.18 (18.38–92.23)	48.66 (14.56–102.71)	60.82 (29.38–125.90)		73.13 (32.94–162.36)	75.22 (32.00–176.78)
FAI	46.20 (18.31–74.08)	39.83 (19.59–80.99)	30.34 (14.69–62.65)	27.29 (13.27–46.32)		1.11 (0.40–3.08)	1.00 (0.27–3.74)
Estradiol					<0.23	<3.64	<0.48
Progesterone					<0.56	11.912 (0.071–72.1)	<0.36

For progesterone in premenopausal females, median and minimum–maximum values (nmol/L) are given.
17-OHP, 17-hydroxyprogesterone; DHEA, dehydroepiandrosterone; DHEAS, dehydroepiandrosterone sulfate; FAI, free androgen index (given without unit).

provide a graphical representation of the distribution of measurable aldosterone concentrations across different ages for aldosterone and include an estimate of the 97.5th centile from these samples in Supplementary Fig. A3.

For 21-deoxycortisol, 11-deoxycorticosterone, and DHT, we observed that chromatograms had high background, preventing reliable integration of the signals, as well as calculation of RIs and graphical representation of data (example chromatograms are shown in Supplementary Fig. A5 and A6).

Impact of BMI

Mean BMI was 27.6 kg/m^2 for males and 27.8 kg/m^2 for females (Table 1). We observed a significant, sex-specific correlation ($P < 0.05$) to BMI for seven of the steroids, including testosterone (Spearman correlation coefficient (r_s) = −0.31, $P < 0.0001$), 17OHP (r_s = −0.19, $P = 0.0014$), DHEA (r_s = −0.20, $P = 0.0007$), and DHEAS (r_s = −0.22, $P = 0.0002$) for males, and androstenedione (r_s = −0.20, $P = 0.0020$), estradiol (r_s = −0.16, $P = 0.0134$), and cortisone (r_s = −0.14, $P = 0.0284$) for females. For SHBG (r_s = −0.12, $P = 0.0380$ for males, r_s = −0.34, $P < 0.0001$ for females) and progesterone (r_s = −0.22, $P = 0.0002$ for males, r_s = −0.19, $P = 0.0038$ for females), we observed a correlation to BMI in both sexes. Significant differences in concentrations between BMI groups were found for testosterone (normal weight vs overweight: $P = 0.0027$, overweight vs obesity: $P = 0.0275$, normal weight vs obesity: $P < 0.0001$), 17-OHP (normal weight vs obesity: $P = 0.0037$), DHEA (normal weight vs obesity: $P = 0.0260$), progesterone (normal weight vs obesity: $P = 0.0033$), and DHEAS (normal weight vs overweight: $P = 0.0339$, normal weight vs obesity: $P = 0.0017$) in males and androstenedione (normal weight vs obesity: $P = 0.0270$) and progesterone (normal weight vs overweight: $P = 0.032$, normal weight vs obesity: $P = 0.039$) in females. SHBG concentrations were also different between BMI groups in males (normal weight vs overweight: 0.042) and in females (normal weight vs overweight: $P = 0.0236$, overweight vs obesity: $P = 0.0004$, normal weight vs obesity: $P < 0.0001$). Sample sizes and hormone levels of each group are given in Supplementary Table A5. FAI correlated positively with BMI in females ($P < 0.0001$) and negatively in males ($P = 0.0178$).

Discussion

Our study provides RIs for 9 of the 15 steroids included in a widely used, commercially available kit for LC-MS/MS-based measurement of steroid profiles. We confirm the assay has adequate reproducibility at concentrations within the measurement range for all steroids. However, our study revealed that using the recommended protocol, sensitivity is limited for measurement of six steroids included in the kit (dihydrotestosterone, estradiol, 21-deoxycortisol,
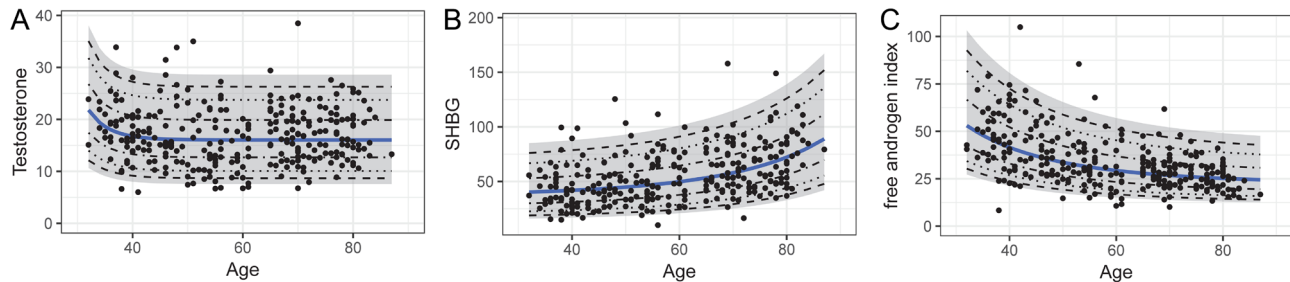
Figure 2

(A) Age-dependent distribution of testosterone concentrations (nmol/L), (B) SHBG concentrations (nmol/L), and (C) FAI in males. Mean concentrations estimated after transformation are represented by the blue line. The 95% CI is marked gray, whereas the dashed lines illustrate the 50%, 80%, and 90% CIs.

aldosterone, 11-deoxycorticosterone, and progesterone in males and postmenopausal females), preventing us from establishing respective RIs.

Differences in measurement techniques (e.g. sample preparation, chromatography, and ionization), biological characteristics of the populations studied (e.g. sample size, age distribution, ethnic background, concomitant medications), and different statistical methods applied make it difficult to directly compare RIs from different studies, even if the studies all use LC-MS/MS (13). Also preanalytical procedures can increase variation of results between laboratories and should be considered when interpreting RIs (27, 28, 29). Preanalytical procedures followed a strict procedure in our study. However, while we generally used EDTA plasma samples, volume limitations for this sample type forced us to use serum (Sarstedt Monovettes with gel separator) for samples from KORA Age. We are aware that problems with the use of gel-separator tubes have been described for some steroids. In our laboratory, we could not detect significant differences between steroids measured from serum or EDTA plasma (even in the percentage of samples with concentrations below the LoQ). According to the manufacturer, interfering signals originating from some of the gel separators are separated from quantifier MRM signals. Many routine laboratories today receive serum samples collected with separator gel, or a spectrum of other tube types. We acknowledge this must be taken into account, and more detailed studies of different sampling tubes might provide important additional information.

Overall, the majority of the steroids concentrations and RIs obtained from our study nicely compare to existing literature, including recent reports by Eisenhofer (23) and Bae (30) which had used in-house LC-MS/MS methods. This particularly applies to RIs for DHEAS, 11-deoxycortisol, testosterone, androstenedione, and 17-OHP in all groups stratified by age or menopausal status, and for cortisol and cortisone in males and postmenopausal females. For cortisol and cortisone in premenopausal women, the upper limit of our RI is about 40% lower than that published by Bae (30) but almost identical to that of Eisenhofer *et al.* (23) and Parikh *et al.*

(31). For some of the steroids included in the commercial kit, published RIs are scarce or lacking. For DHEA, our upper limits of the RIs for males up to an age of 40 and premenopausal females fit well to those from Eisenhofer (23), while the lower limits fit well to those provided by Parikh (31). For corticosterone, concentrations in our study are significantly lower than those reported by Eisenhofer (23), but significantly higher than those reported by Parikh (31). We speculate that, besides characteristics of the respective study populations, differences in calibrators or internal standards for these rarely measured steroids could have contributed to the difference.

As expected, sex was the most important determinant for concentrations of androgens (DHEAS, DHEA, androstenedione, testosterone), gestagens (17-OHP, progesterone), and estradiol. In addition, 17-OHP, progesterone and estradiol were clearly different between pre- and postmenopausal females. Direct comparison of RIs for the FAI from different studies is difficult because it combines the variability associated with measurements of both, testosterone and SHBG (32). Furthermore, there are few studies reporting such RIs in males and postmenopausal females with testosterone measured by LC-MS/MS. For the group of premenopausal females, however, our results are in good agreement with those from two previously published studies (33, 34). FAI declines with age in males, making it advisable to use age adjusted RIs.

In both sexes, we also observed the expected association for DHEAS, DHEA, and androstenedione with age. A particularly pronounced decline in DHEA concentrations in males and androstenedione concentrations in females has also been reported by others (23, 35). In contrast, RIs for cortisol, cortisone, corticosterone, and 11-deoxycortisol were largely independent of age in both sexes. Recently, a trend toward lower concentrations with increasing age was reported for cortisol, cortisone and corticosterone (23). While our data seem to confirm this trend (Fig. 3), the effect seems too small to justify the use of age-adjusted RIs (Supplementary Table A4). Of note, our study population included subjects >80 years of age, and some of those subjects could not be sampled
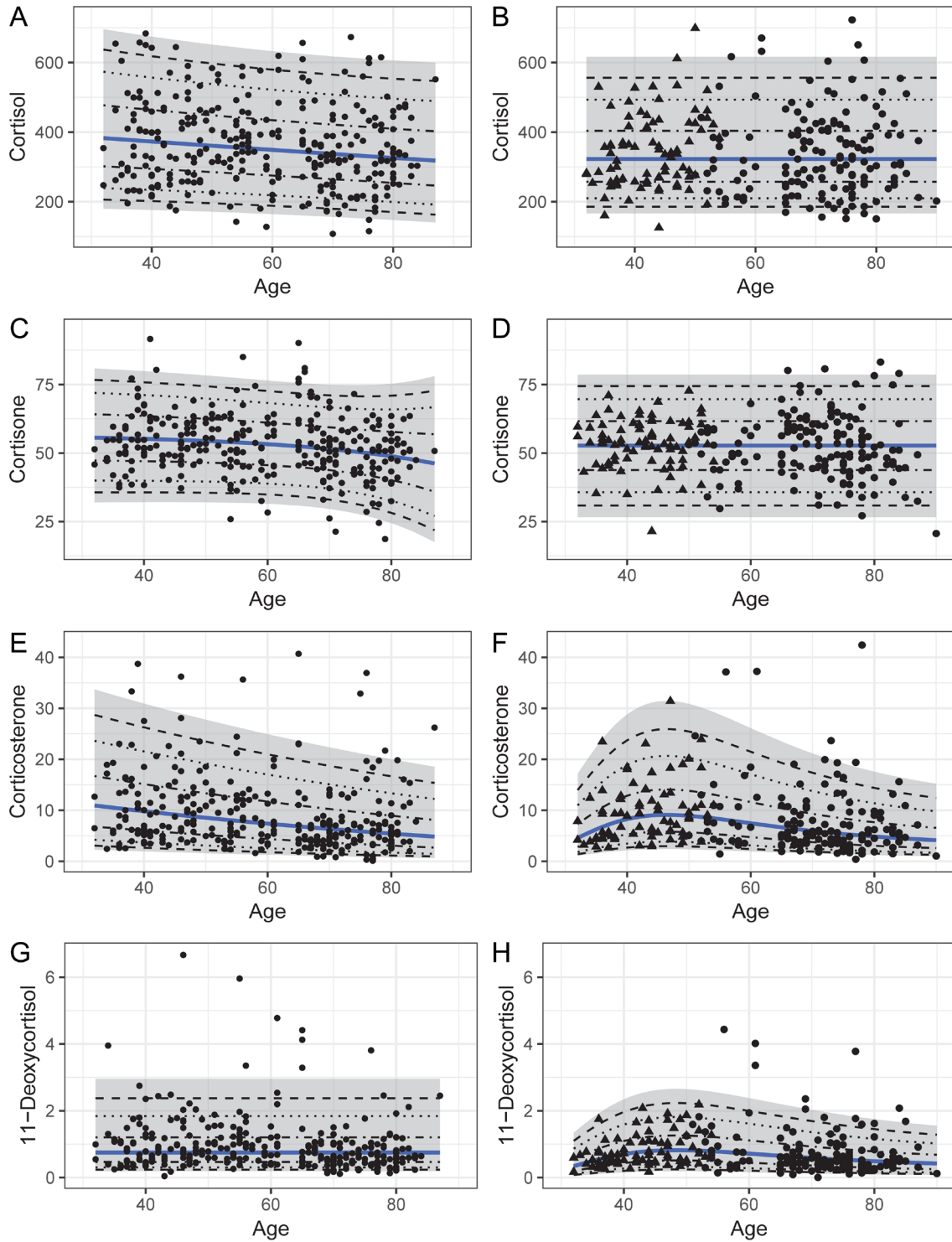
Figure 3

Concentrations (nmol/L) of cortisol (A, B), cortisone (C, D), corticosterone (E, F), and 11-deoxycortisol (G, H) plotted against age for males (left) and females (right). For females, triangles represent premenopausal status, dots represent postmenopausal status. Mean concentrations estimated after transformation are represented by the blue line. The 95% CI is marked gray, whereas the dashed lines illustrate the 50%, 80%, and 90% CIs.

in the early morning. Although this affected only a minority of the subjects, we cannot exclude the possibility that the known diurnal rhythm of cortisol secretion might have contributed to lower concentrations in this age group (31).

Circulating concentrations of five steroids in males and four steroids in females were correlated with BMI ($P < 0.05$). The negative correlation between BMI and testosterone in males and SHBG in both sexes was described before (36, 37, 38, 39). Other steroids exhibiting a negative correlation with BMI in our study population (when measurable) included DHEA, DHEAS, progesterone, and 17-OHP in males and androstenedione, cortisone, estradiol, and progesterone in females. This is an intriguing observation, but in the literature we could not find data providing an explanation. Furthermore, although we have analyzed a comparably large cohort, the numbers of subjects in each BMI group still is too small for a detailed analysis, particularly when split by sex.

Following the protocol suggested by the manufacturer, we obtained well interpretable chromatograms for the majority of the steroids included in this kit. Appropriateness of the analytical method in our hands was further confirmed by results from external quality assessment schemes whenever available and by our confirmation of the manufacturers' claims about precision and LoQs (Table 2). However, it must be mentioned that our study revealed significant shortcomings in the analytical performance of this LC-MS/MS method for some steroids. The approach to measure a panel of 15 steroids with identical chromatographic conditions and similar MS setting inevitably is associated with less optimal conditions for some of the steroids. Aldosterone showed a high percentage of samples with undetectable or extremely low concentrations in all groups of our reference population. This points to limited sensitivity, at least in the context of a routine analytical setting. For DHT, 11-deoxycorticosterone, and 21-deoxycortisol, we also found poor signal to noise ratios with artifact peaks and high background signal making interpretation of chromatograms challenging (Supplementary Fig. A5 and Supplementary Fig. A6). Integration of the 21-deoxycortisol signal is particularly difficult as it elutes right next to the more intense, isobaric corticosterone signal. These two signals are not baseline separated and, in addition, share the same quantifier transitions. We conclude that it cannot be excluded for this method that cross-reactivity might affect quantification – particularly for the smaller signal derived from 21-deoxycortisol. The generally high chromatographic background of this mass transition makes an interpretation of the low intensity 21-deoxycortisol signal even more challenging (Supplementary Fig. A5). Reasons for high background signals and artifact peaks are very likely related to a variety of other biomolecules remaining in the extract even after SPE extraction (1). This assumption is supported by the fact that – in contrast to serum extracts

– measurements of the same steroids from standard solutions demonstrate low background and no artifact signals. Comparison of signals between calibration solutions and serum samples is shown in Supplementary Fig. A5 and Supplementary Fig. A6. Apparently, lower background in comparison to serum samples is also observed for QCs supplied by the manufacturer. This indicates substantial modifications of the matrix, likely necessary to account for endogenous steroid levels in the serum used as basis for calibrations and quality controls. Such problems are known for LC-MS/MS assays in biological fluids (40, 41). In this context, native, commercially available quality control materials for all steroids from independent suppliers with predefined target values from methods of higher metrological order certainly would be helpful for quality management of LC-MS/MS methods.

We assume that optimizations of measurement conditions might be possible, but the limitations in quantification of DHT, 11-deoxycorticosterone, and 21-deoxycortisol under the present protocol prompted us not to propose RIs or maximum concentrations for these steroids. According to our results, the performance of the method is insufficient to reliably quantify these specific steroids in healthy human subjects, where their concentrations physiologically are very low. Therefore, no appropriate RI or maximum value can be calculated. However, it might be relevant to keep in mind that in a clinical context, the more common situation is the need to identify patients with high or very high concentrations of these steroids, which is possible with this method.

Sensitivity issues also affected the measurement of estradiol. We therefore do not provide RIs, but only maximum concentrations for this steroid. Since information on the day of the menstrual cycle was not recorded in our study, we can only report maximum for combined luteal and follicular phase, which is a major limitation of our study. The maximum obtained from our study is higher than the 97.5 percentile reported by some studies (30, 42) but generally fits well into the 95% RI reported in one detailed study describing RIs for each day of the menstrual cycle (36). The missing information on menstrual cycle phase also impairs interpretation of progesterone concentrations in our study. Not surprisingly, data were not normally distributed. The maximum concentrations for progesterone in our study were somewhat lower than those reported in two recent studies for samples across the menstrual cycle (23, 30), but the range of concentrations found in our study falls into the RI provided for samples from the luteal phase only (23).

Conclusion

Using a population-based study sample, we report for the first time RIs for 9 out of 15 steroids included in a

commonly used commercial kit for LC-MS/MS-based quantification. Our RIs are largely comparable to those reported by others using different LC-MS/MS methods. The method has limited sensitivity for quantification of aldosterone and estradiol, the latter particularly affecting measurements in males and postmenopausal females. We do not consider it suitable for quantitative determination of 21-deoxycortisol, 11-deoxycorticosterone, and DHT in healthy subjects, at least in a routine setting as applied in our study. For the other steroids, this method exhibits adequate performance and offers an easy and practicable option for LC-MS/MS-based quantitation. The use of harmonized sets of reagents might help to improve interlaboratory variability.

Supplementary materials

This is linked to the online version of the paper at https://doi.org/10.1530/EC-23-0225.

Declaration of interest

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

Data availability

All data used for the calculation of RIs and other data are not shown in the article or supplementary materials but can be provided by the corresponding author upon reasonable request.

References

1 Wudy SA, Schuler G, Sánchez-Guijo A & Hartmann MF. The art of measuring steroids: principles and practice of current hormonal steroid analysis. *Journal of Steroid Biochemistry and Molecular Biology* 2018 **179** 88–103. (https://doi.org/10.1016/j.jsbmb.2017.09.003)

2 Reisch N, Arlt W & Krone N. Health problems in congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Hormone Research in Paediatrics* 2011 **76** 73–85. (https://doi.org/10.1159/000327794)

3 Eisenhofer G, Masjkur J, Peitzsch M, Di Dalmazi G, Bidlingmaier M, Grüber M, Fazel J, Osswald A, Beuschlein F & Reincke M. Plasma steroid metabolome profiling for diagnosis and subtyping patients with Cushing syndrome. *Clinical Chemistry* 2018 **64** 586–596. (https://doi.org/10.1373/clinchem.2017.282582)

4 Schirpenbach C & Reincke M. Primary aldosteronism: current knowledge and controversies in Conn's syndrome. *Nature Clinical Practice. Endocrinology and Metabolism* 2007 **3** 220–227. (https://doi.org/10.1038/ncpendmet0430)

5 Honour JW. Diagnosis of diseases of steroid hormone production, metabolism and action. *Journal of Clinical Research in Pediatric Endocrinology* 2009 **1** 209–226. (https://doi.org/10.4274/jcrpe.v1i5.209)

6 Handelsman DJ & Wartofsky L. Requirement for mass spectrometry sex steroid assays in the journal of clinical endocrinology and metabolism. *Journal of Clinical Endocrinology and Metabolism* 2013 **98** 3971–3973. (https://doi.org/10.1210/jc.2013-3375)

7 Turpeinen U & Hämäläinen E. Determination of cortisol in serum, saliva and urine. *Best Practice and Research. Clinical Endocrinology and Metabolism* 2013 **27** 795–801. (https://doi.org/10.1016/j.beem.2013.10.008)

8 Vogeser M & Parhofer KG. Liquid chromatography tandem-mass spectrometry (LC-MS/MS) - Technique and applications in endocrinology. *Experimental and Clinical Endocrinology and Diabetes* 2007 **115** 559–570. (https://doi.org/10.1055/s-2007-981458)

9 Fanelli F, Cantù M, Temchenko A, Mezzullo M, Lindner JM, Peitzsch M, Hawley JM, Bruce S, Binz PA, Ackermans MT, *et al.* Report from the HarmoSter study: impact of calibration on comparability of LC-MS/MS measurement of circulating cortisol, 17OH-progesterone and aldosterone. *Clinical Chemistry and Laboratory Medicine* 2022 **60** 726–739. (https://doi.org/10.1515/cclm-2021-1028)

10 Owen LJ & Keevil BG. Testosterone measurement by liquid chromatography tandem mass spectrometry: the importance of internal standard choice. *Annals of Clinical Biochemistry* 2012 **49** 600–602. (https://doi.org/10.1258/acb.2012.012037)

11 Vesper HW, Myers GL & Greg Miller WG. Current practices and challenges in the standardization and harmonization of clinical laboratory tests1-3. *American Journal of Clinical Nutrition* 2016 **104**(Supplement 3) 907S–912S. (https://doi.org/10.3945/ajcn.115.110387)

12 NCCLS. How to define and determine reference intervals in the Clinical Laboratory; approved guideline—second edition. In *NCCLS Document C28-A2,* vol 20, pp 1–38. Wayne, PA, USA: National Committee for Clinical Laboratory Standards, 2000.

13 Tavita N & Greaves RF. Systematic review of serum steroid reference intervals developed using mass spectrometry. *Clinical Biochemistry* 2017 **50** 1260–1274. (https://doi.org/10.1016/j.clinbiochem.2017.07.002)

14 Holle R, Happich M, Löwel H, Wichmann HE & KORA. A research platform for population based health research. *Gesundheitswesen* 2005 **67**(Supplement 1) 19–25. (https://doi.org/10.1055/s-2005-858235)

15 Hannemann A, Bidlingmaier M, Friedrich N, Manolopoulou J, Spyroglou A, Völzke H, Beuschlein F, Seissler J, Rettig R, Felix SB, *et al.* Screening for primary aldosteronism in hypertensive subjects: results from two German epidemiological studies. *European Journal of Endocrinology* 2012 **167** 7–15. (https://doi.org/10.1530/EJE-11-1013)

16 Emeny RT, Bidlingmaier M, Lacruz ME, Linkohr B, Peters A, Reincke M & Ladwig KH. Mind over hormones: sex differences in associations of well-being with IGF-I, IGFBP-3 and physical activity in the KORA-Age study. *Experimental Gerontology* 2014 **59** 58–64. (https://doi.org/10.1016/j.exger.2014.08.001)

17 Peters A, Döring A, Ladwig KH, Meisinger C, Linkohr B, Autenrieth C, Baumeister SE, Behr J, Bergner A, Bickel H, *et al.* Multimorbidität und erfolgreiches Altern: ein Blick auf die

Bevölkerung im Rahmen der KORA-Age-Studie. *Zeitschrift für Gerontologie und Geriatrie* 2011 **44**(Supplement 2) 41–54. (https://doi.org/10.1007/s00391-011-0245-7)

18 Kuan V, Denaxas S, Patalay P, Nitsch D, Mathur R, Gonzalez-Izquierdo A, Sofat R, Partridge L, Roberts A, Wong ICK, *et al.* Identifying and visualising multimorbidity and comorbidity patterns in patients in the English National Health Service: a population-based study. *Lancet. Digital Health* 2023 **5** e16–e27. (https://doi.org/10.1016/S2589-7500(22)00187-X)

19 Marques A, Peralta M, Martins J, Loureiro V, Almanzar PC & de Matos MG. Few European adults are living a healthy lifestyle. *American Journal of Health Promotion* 2019 **33** 391–398. (https://doi.org/10.1177/0890117118787078)

20 World Health Organization. *Waist circumference and waist-hip ratio: report of a WHO expert consultation.* Geneva, Switzerland: WHO, 2008. (available at: https://www.who.int/publications/i/item/9789241501491)

21 Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, Cooney MT, Corrà U, Cosyns B, Deaton C, *et al.* European Guidelines on cardiovascular disease prevention in clinical practice. *European Heart Journal* 2016 **37** 2315–2381. (https://doi.org/10.1093/eurheartj/ehw106)

22 European Medicines Agency. *Guideline on bioanalytical method validation*. London, UK: EMA, 2011. (available at: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-bioanalytical-method-validation_en.pdf)

23 Eisenhofer G, Peitzsch M, Kaden D, Langton K, Pamporaki C, Masjkur J, Tsatsaronis G, Mangelis A, Williams TA, Reincke M, *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. *Clinica Chimica Acta* 2017 **470** 115–124. (https://doi.org/10.1016/j.cca.2017.05.002)

24 Franscini LC, Vazquez-Montes M, Buclin T, Perera R, Dunand M, Grouzmann E & Beck-Popovic M. Pediatric reference intervals for plasma free and total metanephrines established with a parametric approach: relevance to the diagnosis of neuroblastoma. *Pediatric Blood and Cancer* 2015 **62** 587–593. (https://doi.org/10.1002/pbc.25385)

25 Royston P & Wright EM. A method for estimating age-specific reference intervals ('normal ranges') based on fractional polynomials and exponential transformation. *Journal of the Royal Statistical Society Series A* 1998 **161** 79–101. (https://doi.org/10.1111/1467-985X.00091)

26 Seidell JC & Flegal KM. Assessing obesity: classification and epidemiology. *British Medical Bulletin* 1997 **53** 238–252. (https://doi.org/10.1093/oxfordjournals.bmb.a011611)

27 Plebani M. Quality indicators to detect pre-analytical errors in laboratory testing. *Clinical Biochemist. Reviews* 2012 **33** 85–88. (https://doi.org/10.1016/j.cca.2013.07.033)

28 Wilkinson IB, Raine T, Wiles K, Goodhart A, Hall C & O'Neill H. Reference intervals, etc. *Oxford Handbook of Clinical Medicine* 2017 **29** 750–757. (https://doi.org/10.1093/med/9780199689903.003.0017)

29 Vesper HW & Thienpont LM. Traceability in laboratory medicine. *Clinical Chemistry* 2009 **55** 1067–1075. (https://doi.org/10.1373/clinchem.2008.107052)

30 Bae YJ, Zeidler R, Baber R, Vogel M, Wirkner K, Loeffler M, Ceglarek U, Kiess W, Körner A, Thiery J, *et al.* Reference intervals of nine steroid hormones over the life-span analyzed by LC-MS/MS: effect of age, gender, puberty, and oral contraceptives. *Journal of Steroid Biochemistry and Molecular Biology* 2019 **193** 105409. (https://doi.org/10.1016/j.jsbmb.2019.105409)

31 Parikh TP, Stolze B, Ozarda Y, Jonklaas J, Welsh K, Masika L, Hill M, Decherney A & Soldin SJ. Diurnal variation of steroid hormones and their reference intervals using mass spectrometric analysis. *Endocrine Connections* 2018 **7** 1354–1361. (https://doi.org/10.1530/EC-18-0417)

32 Adaway J, Keevil B, Miller A, Monaghan PJ, Merrett N & Owen L. Ramifications of variability in sex hormone-binding globulin measurement by different immunoassays on the calculation of free testosterone. *Annals of Clinical Biochemistry* 2020 **57** 88–94. (https://doi.org/10.1177/0004563219888549)

33 Bui HN, Sluss PM, Hayes FJ, Blincko S, Knol DL, Blankenstein MA & Heijboer AC. Testosterone, free testosterone, and free androgen index in women: reference intervals, biological variation, and diagnostic value in polycystic ovary syndrome. *Clinica Chimica Acta* 2015 **450** 227–232. (https://doi.org/10.1016/j.cca.2015.08.019)

34 Yasmin E, Balen AH & Barth JH. The association of body mass index and biochemical hyperandrogenaemia in women with and without polycystic ovary syndrome. *European Journal of Obstetrics, Gynecology, and Reproductive Biology* 2013 **166** 173–177. (https://doi.org/10.1016/j.ejogrb.2012.09.025)

35 Haring R, Hannemann A, John U, Radke D, Nauck M, Wallaschofski H, Owen L, Adaway J, Keevil BG & Brabant G. Age-specific reference ranges for serum testosterone and androstenedione concentrations in women measured by liquid chromatography-tandem mass spectrometry. *Journal of Clinical Endocrinology and Metabolism* 2012 **97** 408–415. (https://doi.org/10.1210/jc.2011-2134)

36 Verdonk SJE, Vesper HW, Martens F, Sluss PM, Hillebrand JJ & Heijboer AC. Estradiol reference intervals in women during the menstrual cycle, postmenopausal women and men using an LC-MS/MS method. *Clinica Chimica Acta* 2019 **495** 198–204. (https://doi.org/10.1016/j.cca.2019.04.062)

37 Osuna JA, Gómez-Pérez R, Arata-Bellabarba G & Villaroel V. Relationship between BMI, total testosterone, sex hormone-binding-globulin, leptin, insulin and insulin resistance in obese men. *Archives of Andrology* 2006 **52** 355–361. (https://doi.org/10.1080/01485010600692017)

38 Maggio M, Lauretani F, Basaria S, Ceda GP, Bandinelli S, Metter EJ, Bos AJ, Ruggiero C, Ceresini G, Paolisso G, *et al.* Sex hormones binding globulin levels across the adult lifespan in women - the role of body mass index and fasting insulin. *Journal of Endocrinological Investigation* 2008 **31** 597–601. (https://doi.org/10.1007/BF03345608)

39 Rohrmann S, Shiels MS, Lopez DS, Rifai N, Nelson WG, Kanarek N, Guallar E, Menke A, Joshu CE, Feinleib M, *et al.* Body fatness and sex steroid hormone concentrations in US men: results from NHANES III. *Cancer Causes and Control* 2011 **22** 1141–1151. (https://doi.org/10.1007/s10552-011-9790-z)

40 Khamis MM, Adamko DJ & El-Aneed A. Strategies and challenges in method development and validation for the absolute quantification of endogenous biomarker metabolites using liquid chromatography-tandem mass spectrometry. *Mass Spectrometry Reviews* 2021 **40** 31–52. (https://doi.org/10.1002/mas.21607)

41 Eeckhaut A Van, Lanckmans K, Sarre S, Smolders I & Michotte Y. Validation of bioanalytical LC-MS/MS assays: evaluation of matrix effects. *Journal of Chromatography. B, Analytical Technologies in the Biomedical and Life Sciences* 2009 **877** 2198–2207. (https://doi.org/10.1016/j.jchromb.2009.01.003)

42 Mezzullo M, Pelusi C, Fazzini A, Repaci A, Dalmazi G Di, Gambineri A, Pagotto U & Fanelli F. Female and male serum reference intervals for challenging sex and precursor steroids by liquid chromatography - tandem mass spectrometry. *Journal of Steroid Biochemistry and Molecular Biology* 2020 **197** 105538. (https://doi.org/10.1016/j.jsbmb.2019.105538)