Heleen I. Jansen, Antonius E. van Herwaarden, Henk J. Huijgen, Marc G. Vervloet, Jacquelien J. Hillebrand, Anita Boelen and Annemieke C. Heijboer*

Lower accuracy of testosterone, cortisol, and free T4 measurements using automated immunoassays in people undergoing hemodialysis

https://doi.org/10.1515/cclm-2022-1133
Received November 8, 2022; accepted February 16, 2023;
published online March 6, 2023

Abstract

Objectives: Hormone measurements using automated immunoassays (IAs) can be affected by the sample matrix. Liquid chromatography tandem-mass spectrometry (LC-MS/MS) is less affected by these matrix effects. In clinical laboratories, testosterone, cortisol and, free thyroxine (FT4) are often measured using IAs. Renal failure alters serum composition in blood samples from people undergoing hemodialysis (HDp) and have, therefore, a complex serum constitution compared to healthy controls (HC). The goal of this study was to investigate the accuracy of testosterone, cortisol, and FT4 measurements in samples of HDp and to get more insight in the interfering factors.

Methods: Thirty serum samples from HDp and HC were collected to measure testosterone, cortisol, and FT4 using a well standardized isotope dilution (ID)-LC-MS/MS method and 5 commercially available automated IAs (Alinity, Atellica, Cobas, Lumipulse, UniCel DXI). Method comparisons between LC-MS/MS and IAs were performed using both HDp and HC samples.

Results: Average bias from the LC-MS/MS was for testosterone, cortisol, and FT4 immunoassays respectively up to 92, 7–47 and 16–27% more in HDp than in HC samples and was IA dependent. FT4 IA results were falsely decreased in HDp samples, whereas cortisol and testosterone concentrations in females were predominantly falsely increased. Correlation coefficients between LC-MS/MS and IA results were lower in HDp compared to HC samples.

Conclusions: Several IAs for testosterone (in women), cortisol, and FT4 are less reliable in the altered serum matrix of samples of HDp than in HC. Medical and laboratory specialists should be aware of these pitfalls in this specific population.

Keywords: cortisol; free thyroxine; hemodialysis; immunoassay; mass spectrometry; testosterone.

***Corresponding author: Annemieke C. Heijboer**, Department of Clinical Chemistry, Endocrine Laboratory, Amsterdam UMC Location Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; Amsterdam Gastroenterology, Endocrinology & Metabolism, Amsterdam, The Netherlands; Department of Clinical Chemistry, Endocrine Laboratory, Amsterdam UMC Location University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands; and Amsterdam Reproduction & Development Research Institute, Amsterdam, The Netherlands, E-mail: a.heijboer@amsterdamumc.nl

Heleen I. Jansen, Department of Clinical Chemistry, Endocrine Laboratory, Amsterdam UMC Location Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; Amsterdam Gastroenterology, Endocrinology & Metabolism, Amsterdam, The Netherlands; and Department of Clinical Chemistry, Endocrine Laboratory, Amsterdam UMC Location University of Amsterdam, Amsterdam, The Netherlands. https://orcid.org/0000-0003-2359-2865

Antonius E. van Herwaarden, Department of Laboratory Medicine, Radboud University Medical Center, Nijmegen, The Netherlands

Henk J. Huijgen, Department of Clinical Chemistry, Red Cross Hospital, Beverwijk, The Netherlands

Marc G. Vervloet, Department of Nephrology, Amsterdam UMC Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Jacquelien J. Hillebrand, Amsterdam Gastroenterology, Endocrinology & Metabolism, Amsterdam, The Netherlands; and Department of Clinical Chemistry, Endocrine Laboratory, Amsterdam UMC Location University of Amsterdam, Amsterdam, The Netherlands

Anita Boelen, Amsterdam Gastroenterology, Endocrinology & Metabolism, Amsterdam, The Netherlands; Department of Clinical Chemistry, Endocrine Laboratory, Amsterdam UMC Location University of Amsterdam, Amsterdam, The Netherlands; and Amsterdam Reproduction & Development Research Institute, Amsterdam, The Netherlands

Introduction

Automated immunoassays (IAs) have replaced the manual IAs over the last decades in the measurement of hormones. Meanwhile, we have learned that these automated IAs are often functioning well in a relatively healthy group. In contrast, difficulties are seen in people with specific conditions [1–4]. Multiple factors can interfere with IAs. Differences in concentrations of these factors or differences

Open Access. © 2023 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

in composition of the serum, as can occur in some disease states, are described as alterations in the serum matrix. Therefore, an altered matrix can lead to misleading outcomes [5]. An analytical measurement method based on liquid chromatography tandem-mass spectrometry (LC-MS/MS) is more robust to matrix effects and can thus be a more accurate technique for hormone measurements in patient groups with an altered serum composition [6–8].

People undergoing hemodialysis (HDp) are known to have advanced renal failure and, therefore, an altered serum matrix based on accumulation of a variety of uremic toxins, like urea, and problems with hormone measurements using IAs are seen in these people without precisely knowing what the interfering factor is [9]. Heijboer et al. [2] and Depreter et al. [10] already proved that the measurement of vitamin D in HDp is not reliable when using automated IAs. Testosterone, cortisol, and free thyroxine (FT4) are frequently measured hormones in clinical practice. Still, it has not yet been adequately elucidated how reliable these IA hormone measurements are in samples of HDp. This may have clinical implications.

In addition, it is valuable to obtain insight into the characteristics of the interfering factors in serum samples of HDp. The accumulation of a variety of uremic toxins due to advanced renal failure in serum of HDp can partially be removed by hemodialysis. Hemodialysis can be performed using either hemodialysis (HD) or hemodiafiltration (HDF). HDF is based on a combination of clearance via diffusion and convective clearance in contrast to a technique based on only clearance via diffusion such as in HD [11]. HDF filters more and larger uremic toxins within the middle-weight molecule range compared to HD. The effect of a hemodialysis session (both HD and HDF) on the accuracy of testosterone, cortisol, and FT4 IA measurements is not yet well mapped and might help determining interfering factors. Knowledge of the impact of dialysis and the modality used can inform clinicians on the optimal timing of determining the hormones studied here.

Therefore, the main goal of this study was to assess the robustness of testosterone, cortisol, and FT4 IA measurements in samples of HDp. The second aim was to investigate potential interfering effects of hemodialysis by comparing IAs in samples taken before and after hemodialysis.

Materials and methods

Samples

Samples of people undergoing hemodialysis vs. healthy controls (experiment 1): Serum samples were obtained from 30 people undergoing hemodialysis because of advanced renal failure (mean age 57.4 years, 19 men, 11 women) and 30 healthy controls (mean age 39.5 years, 11 men, 19 women) in March and April 2021 after written consent was provided. The local Medical Ethical Committee of the Amsterdam UMC, location Academic Medical Centre confirmed that ethical approval was not required since this research was set up for laboratory analyses quality improvement. Additional blood samples were collected from HDp during their regular monthly blood control. Healthy controls were recruited among employees at Amsterdam UMC. HDp were excluded if they used thyroid medication (e.g., levothyroxine, thiamazol); healthy controls were excluded if they used thyroid medication, took oral contraceptives or were pregnant. In HDp, blood was sampled between 8 and 10 a.m. or between 2 and 4 p.m. before the start of their regular hemodialysis session at Amsterdam UMC (The Netherlands), whereas blood was sampled between 8 a.m. and 5 p.m. in healthy controls. All samples were handled identically; the samples were aliquoted after centrifugation (5 min at 1,900*g*) and kept frozen at −20 °C until analysis. Storage time did not exceed 6 months. All samples were thawed, vortexed and centrifuged after which analyses took place.

Samples before and after hemodialysis (experiment 2): To provide better insight into the interfering factors of serum of HDp, serum samples were obtained from 17 HDp before and immediately after hemodialysis (mean age 58.6 years, 10 men, 7 women) in December 2021 after written consent was provided. Patients received either regular hemodialysis (HD; n=7) or hemodiafiltration (HDF; n=10). The local Medical Ethical Committee of the Amsterdam UMC, location Academic Medical Centre confirmed that ethical approval was not required since this research was set up for laboratory analyses quality improvement. Additional blood samples were collected between 8 and 10 a.m. or between 2 and 4 p.m. before the start of their regular hemodialysis session and between 12 a.m. and 2 p.m. or 6 and 8 p.m. after their regular hemodialysis session at Amsterdam UMC (The Netherlands). The same exclusion criteria accounted for these patients as described in experiment 1. All samples were handled as in experiment 1. Storage time did not exceed 2 months.

Methods

Experiment 1 (immunoassays): Serum testosterone and FT4 concentrations were measured using five different commercially available automated immunoassays: Alinity (Abbott), Atellica (Siemens), Cobas (Roche), Lumipulse (Fuijrebio) and UniCel DXI (Beckman Coulter). Serum cortisol concentrations were measured using four different commercially available automated immunoassays: Alinity (Abbott), Atellica (Siemens), Cobas (Roche) and UniCel DXI (Beckman Coulter). SHBG was measured using the Alinity (Abbott) immunoassay, TBG using radioimmunoassay (RIA; Thermo Fischer Scientific) and albumin using the Cobas (Roche) with a colorimetric assay (bromocresol purple [BCP] method). Analyses using Alinity, Atellica, Cobas, Lumipulse and RIA were performed at the Endocrine Laboratory of Amsterdam UMC. Analyses using UniCel DXI were performed at the laboratory of Red Cross Hospital (RKZ).

Experiment 2 (immunoassays): Serum testosterone, cortisol, and FT4 concentrations were measured at the Endocrine Laboratory of Amsterdam UMC using the two automated IAs that showed the least agreement with the LC-MS/MS in experiment 1 (Alinity [Abbott] and Atellica [Siemens]). Albumin concentrations were measured using the

Cobas colorimetric assay (Roche; bromocresol purple [BCP] method) at the department of Laboratory Medicine of Radboud UMC.

Experiment 1 and 2 (LC-MS/MS): Serum cortisol and testosterone measurements using LC-MS/MS were performed at the Endocrine Laboratory of Amsterdam UMC. The testosterone and cortisol method is described briefly. A Waters Acquity UPLC system coupled to a Waters Xevo TQ-XS mass spectrometer (Wilmslow, UK) was used. Chromatographic separation was performed using a Waters® HSS-T3 column 2.1 × 50 mm, 1.8 µm. The column temperature was set to 60 °C. Mobile phases A and B consisted of 0.1% formic acid and 2 mM ammonium acetate in water and 0.1% formic acid and 2 mM ammonium acetate in methanol, respectively. An internal standard mixture of ca. 1 mg/mL of d_3-testosterone and d_4-cortisol (CDN, cat.nr. D-3793 and Cambridge Isotope Laboratories, Inc. DLM-2218) in absolute ethanol was used. Internal standard was diluted in acetonitrile and added to the samples, acetonitrile caused protein precipitation. Next, samples were filtered and injected into the LC-MS/MS. At a flow rate of 0.4 mL/min, a gradient was employed ranging from 20% of mobile phase B to 98% of mobile phase B in 6 min. The mass spectrometer was operated in positive mode using multiple reaction monitoring (MRM), where the transitions 289.17 > 97.17 for testosterone, 292.23 > 97.11 for d_3-testosterone IS, 363.20 > 121.16 for cortisol and 367.20 > 121.16 for d_4-cortisol IS were used. The settings of the mass spectrometer were as follows: capillary voltage 1 kV; cone voltage 44 V; source temperature 150 °C; desolvation temperature 650 °C; desolvation gas 1200 L/h; cone gas 50 L/h; collision energy: 20 eV. The lower limit of quantification (LLOQ) was 0.10 nmol/L for testosterone and 1.0 nmol/L for cortisol. The mean intra-assay and inter-assay variation were <5% for both testosterone and cortisol. This LC-MS/MS method performed well in measuring cortisol and testosterone set against other LC-MS/MS methods [12–14]. Acquisition and data processing were performed using Waters MassLynx Software. All testosterone and cortisol measurements were performed simultaneously. Serum FT4 measurements using ID-LC-MS/MS preceded by equilibrium dialysis were performed in duplicate by the candidate conventional reference measurement procedure at the department of Laboratory Medicine of Radboud UMC according to the conventions proposed by the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) [15] (Jansen et al. under review).

Statistics

Experiment 1: Mean testosterone, cortisol, TBG, and albumin concentrations in both groups were compared using independent t-tests and median FT4 and SHBG concentrations in both groups were compared using a Mann-Whitney U-test. Passing and Bablok regression analysis (including Cusum test for linearity), Bland-Altman plots and Pearson correlation coefficients (r^2) were used for method comparisons. To provide better insight in the absolute effect of the matrix of HDp, we recalculated the IA results towards the LC-MS/MS results based on the method comparison for healthy controls. The testosterone, cortisol, and FT4 concentrations of HDp were recalculated by the following formula: (initial analyte measured using IA – intercept)/slope. Intercept and slope were derived from the Passing and Bablok regression analyses between the respective IA and LC-MS/MS analyte results in healthy controls. The recalculated concentrations from HDp were used to make Bland-Altman plots.

Experiment 2: Passing and Bablok regression analysis (including Cusum test for linearity), Bland-Altman plots and Pearson correlation coefficients (r^2) were performed to assess the performance of the immunoassays compared to the LC-MS/MS method for all three hormones. A distinction was made between regular HD and HDF patients. Last, Mann-Whitney U-tests and independent samples t-tests were accomplished to assess testosterone, cortisol, FT4, and albumin concentration differences between patients receiving HD and HDF.

All statistical analyses were performed using MedCalc (version 18.5, MedCalc Software). A value of $p \leq 0.05$ was considered statistically significant.

Results

Experiment 1

Testosterone, SHBG, cortisol, FT4, TBG, and albumin concentrations in HDp and healthy controls (testosterone, cortisol, and FT4 measured using LC-MS/MS) were displayed in Table 1.

Figures 1–3 show the Passing and Bablok regression analyses including slopes, intercepts, and correlation coefficients of the testosterone, cortisol, and FT4 IAs compared to the LC-MS/MS in both HDp and healthy controls. Additional information regarding these parameters is presented in Supplementary Tables 1–3. All Passing and Bablok regressions could be linearly assessed. With one exception, all testosterone, cortisol, and FT4 correlation coefficients were lower in HDp compared with healthy controls, meaning that the agreement between the IAs and the LC-MS/MS was worse in HDp for all hormones. Furthermore, the figures show that the regression lines of HDp from both the testosterone and cortisol measurements exceeded the regression lines of healthy controls. This accounted for almost all testosterone and cortisol IAs, although for some to a greater extent than others. Only the regression line of HDp from the testosterone Atellica IA was lower than the regression line of the healthy controls. In contrast, the regression lines of HDp from the FT4 measurement were lower than the regression lines of healthy controls in all IAs. These results indicated a positive bias of testosterone and cortisol and a negative bias of FT4 using IAs in samples of HDp. Figures 1–3 show that IAs had a bias compared to the LC-MS/MS in healthy controls as well. Therefore to specifically highlight the bias in HDp, we controlled for the bias between the IA and LC-MS/MS in healthy controls and used the recalculated testosterone, cortisol, and FT4 results to make Bland-Altman plots as described in the method section and presented in Supplementary Figures 1–3. Indeed, a positive bias of testosterone and cortisol and a negative bias of FT4 using IAs in samples

Table 1: Independent t-test or Mann-Whitney U-test[a]; comparison HDp and healthy controls. Testosterone, cortisol, and FT4 were measured using LC-MS/MS.

Analyte		Mean/median[a] HDp	Mean/median[a] healthy controls	95% CI difference	p-Value
Testosterone	Men	8.85 nmol/L	14.2 nmol/L	1.61–9.16 nmol/L	0.0068
	Women	0.39 nmol/L	0.70 nmol/L	0.13–0.49 nmol/L	0.0013
SHBG[a]		41.9 nmol/L[a]	45.2 nmol/L[a]	0.1–26.0 nmol/L	0.0468
Cortisol		259 nmol/L	246 nmol/L	−67.8 to 40.36 nmol/L	0.6139
FT4[a]		25.1 pmol/L[a]	19.3 pmol/L[a]	−7.2 to −3.7 pmol/L	<0.0001
TBG		289 nmol/L	302 nmol/L	−7.66 to 32.7 nmol/L	0.2195
Albumin		35.7 g/L	40.5 g/L	2.20–7.31 g/L	0.0004

[a]Mann-Whitney U-test was performed for these analytes. For these analytes a median concentration was reported. This Table presents the 95% confidence interval (CI) from the difference between HDp and healthy controls.

of HDp was found. The positive bias for testosterone varied from 1.4% for the Atellica IA up to 91.7% for the UniCel DXI IA. As could be expected, testosterone concentrations in women were lower compared to men and the bias in women primarily determined the total bias (Supplementary Figure 4). The positive bias for cortisol varied from 6.6% for the Alinity IA up to 47.1% for the Atellica IA and was statistically significant for all IAs as can be derived from the 95% CIs (Supplementary Figure 2). The negative bias for FT4 varied from 16.1% for the UniCel DXI IA up to 27.3% for the Lumipulse IA which was also statistically significant (Supplementary Figure 3).

In total, three female HDp samples showed testosterone concentrations below the LLOQ for one or two of the IAs (2 for Alinity, 2 for Atellica, 1 for Cobas); results of these samples were set at half of the LLOQ and included. Two cortisol measurements from HDp were excluded for the Bland-Altman plots as these influenced the mean bias extraordinarily. The cortisol concentrations in these samples measured using LC-MS/MS were below 50 nmol/L (and were between 13 and 31 nmol/L measured using the IAs) and therefore we suspected sensitivity problems in the IAs, not especially influenced by the matrix. One person undergoing hemodialysis showed a FT4 value with a CV of 40% without clear explanation measured using the LC-MS/MS, meaning this result could not be interpreted reliably. No additional material was available to repeat this analysis, so this sample was excluded for FT4 analysis.

Experiment 2

To search for interfering factors present in serum samples of HDp testosterone, cortisol, and FT4 concentrations were measured before and after hemodialysis, Passing and Bablok regression analyses are shown in Supplementary Figure 5. Additional information regarding these parameters is presented in Supplementary Table 4. All Passing and Bablok regressions could be linearly assessed. Correlation coefficients fitting the Passing and Bablok regression analyses somewhat differed before and after hemodialysis, although the overlap between 95% confidence intervals was large. Slopes derived from the Passing and Bablok regression analyses of the testosterone, cortisol, and FT4 Atellica IA were closer to 1 after hemodialysis. These differences were however minor and 95% confidence intervals largely overlapped. The HD and HDF groups were separately assessed and Passing and Bablok regression analyses and correlation coefficients (r^2) did not show a consistent difference between these two groups. No differences in testosterone, cortisol, and FT4 concentrations between HD and HDF patients were found. Only albumin concentrations differed significantly after hemodialysis between HD and HDF patients (mean albumin HD 39.6 g/L; mean albumin HDF 33.9 g/L; 95% CI −10.9 to −0.4 g/L; p=0.036). One female HDp showed a testosterone concentration below the LLOQ for the Atellica IA both before and after hemodialysis, so these results were set at half of the LLOQ and included in the statistical analysis.

Discussion

In this study, we performed a method comparison between five frequently used automated commercially available immunoassays (IAs) and LC-MS/MS to study the effects of advanced renal failure and hemodialysis on the accuracy of testosterone, cortisol, and FT4 IA measurements. Testosterone and cortisol concentrations were mainly overestimated, whereas FT4 concentrations were mainly underestimated in samples of HDp using IAs compared to healthy controls.

A bias of testosterone, cortisol, and FT4 concentrations was found when measured using IAs compared to LC-MS/MS

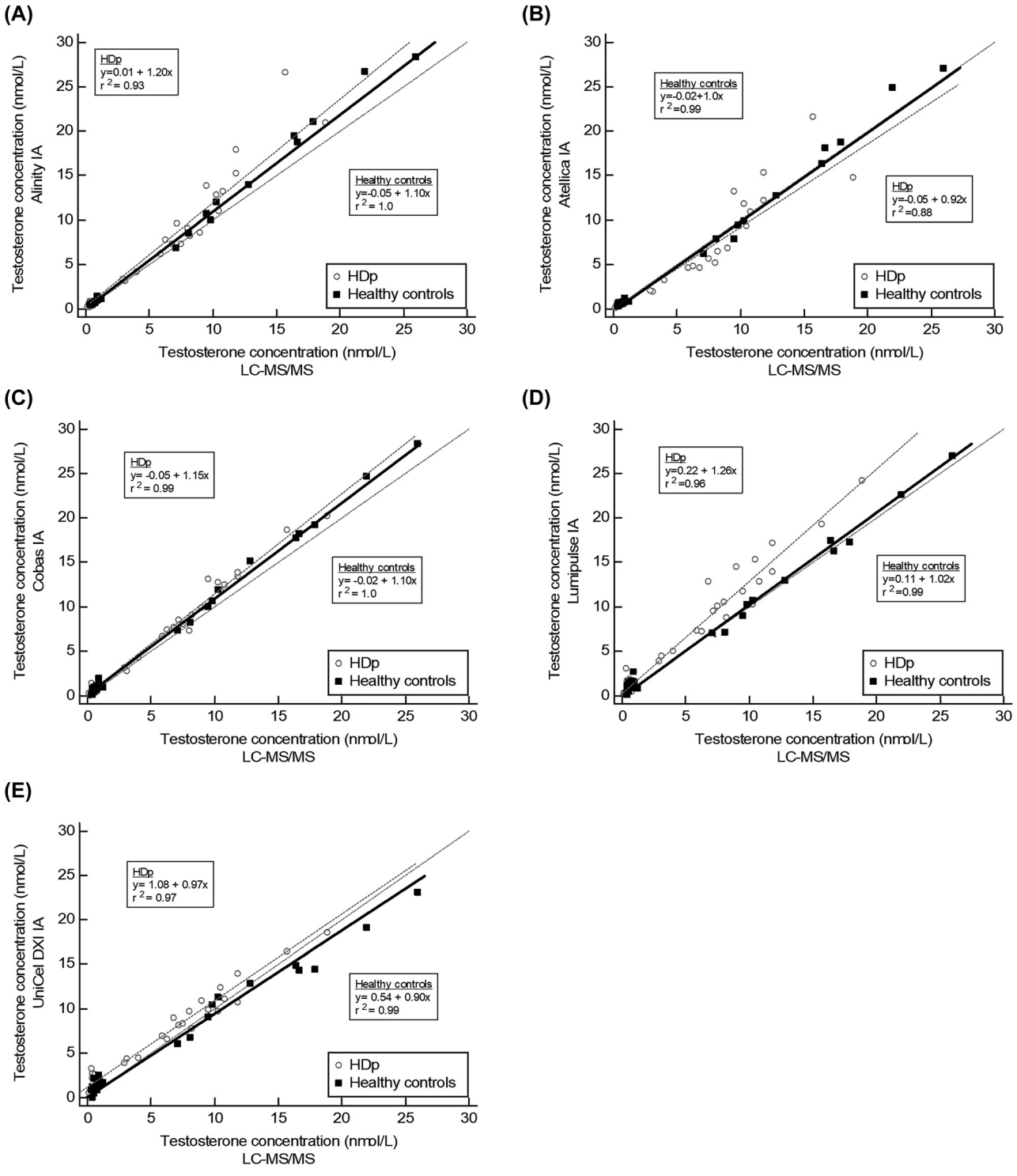

Figure 1: Passing and Bablok regression analyses for the 5 automated testosterone immunoassays in both healthy controls and HDp (experiment 1). On the x-axis, the testosterone concentrations were measured using LC-MS/MS and, on the y-axis, the testosterone concentrations measured using the respective immunoassays are shown. (A) Alinity; (B) Atellica; (C) Cobas; (D) Lumipulse; (E) UniCel DXI.

in both HDp and healthy controls. This has been acknowledged and consensus has reached that standardization of these hormones towards the LC-MS/MS in healthy controls is necessary [16–18]. Therefore, the emphasis of the current study is on the observed IA accuracy difference between samples of HDp and healthy controls. The lower correlation

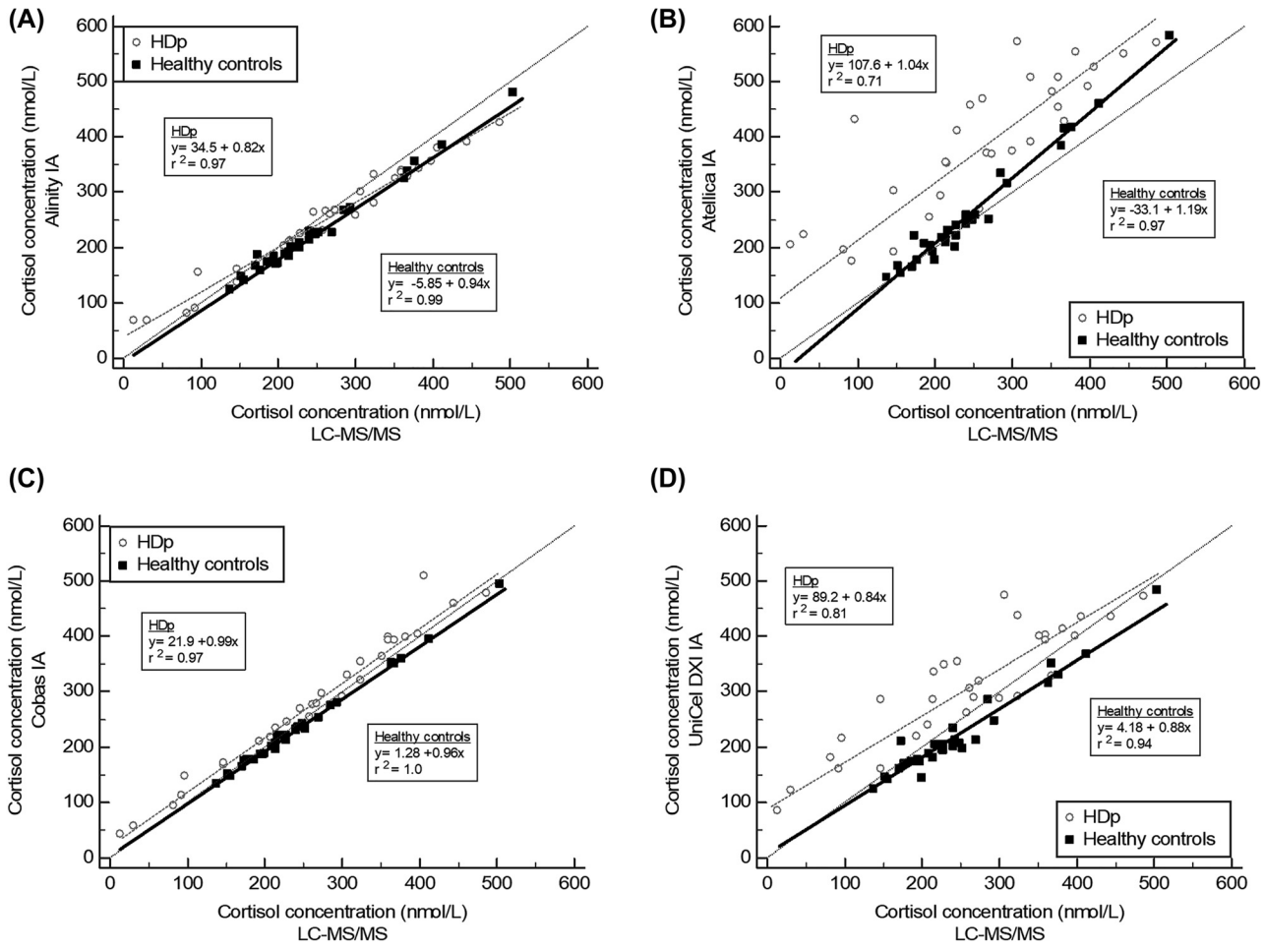

Figure 2: Passing and Bablok regression analyses for the 4 automated cortisol immunoassays in both healthy controls and HDp (experiment 1). On the x-axis, the cortisol concentrations were measured using LC-MS/MS and, on the y-axis, the cortisol concentrations measured using the respective immunoassays are shown. (A) Alinity; (B) Atellica; (C) Cobas; (D) UniCel DXI.

coefficients of nearly all IAs in HDp compared with healthy controls emphasizes the less reliable performance of IAs in measuring testosterone, cortisol, and FT4 in HDp. After controlling for the bias between the IA and LC-MS/MS in healthy controls, testosterone and cortisol concentrations showed a positive bias from the LC-MS/MS (+1.4–91.7 and +6.6–47.1%, respectively) in samples of HDp measured using all tested IAs. This positive bias assessed using Bland-Altman plots was statistically significant for all cortisol IAs, indicating falsely increased cortisol levels in HDp when measured using IAs. This matched a previous study on cortisol IA accuracy in patients with renal failure [4]. It must, however, be stated that not every IA showed an overestimation of cortisol of the same magnitude and the bias of results of some IAs may not be relevant in clinical practice. The positive bias of all testosterone IAs indicated also an overestimation of testosterone levels in HDp when measured using IAs. Even though all testosterone IAs showed some positive bias in HDp, this was not the case to the same extent for each IA and this effect was mainly observed in female samples with low concentrations of testosterone. The clinical relevance of this finding in women may be small and the number of female samples are too low to draw definite conclusions. These results mainly demonstrated that attention may be needed for testosterone measurements in samples of HDp. To the best of our knowledge, this is the first study elucidating how reliable testosterone measurements using IAs are in people with advanced renal failure undergoing hemodialysis. FT4 concentrations in samples of HDp showed a negative bias from the LC-MS/MS (–16.1 to –27.3%) which was statistically significant for all IAs. Thus, FT4 levels were mainly falsely decreased in HDp, which is in concordance with a previous study that found an underestimation of the FT4 IA measurement in patients with renal failure [1], even though that study did not use a (candidate) reference method

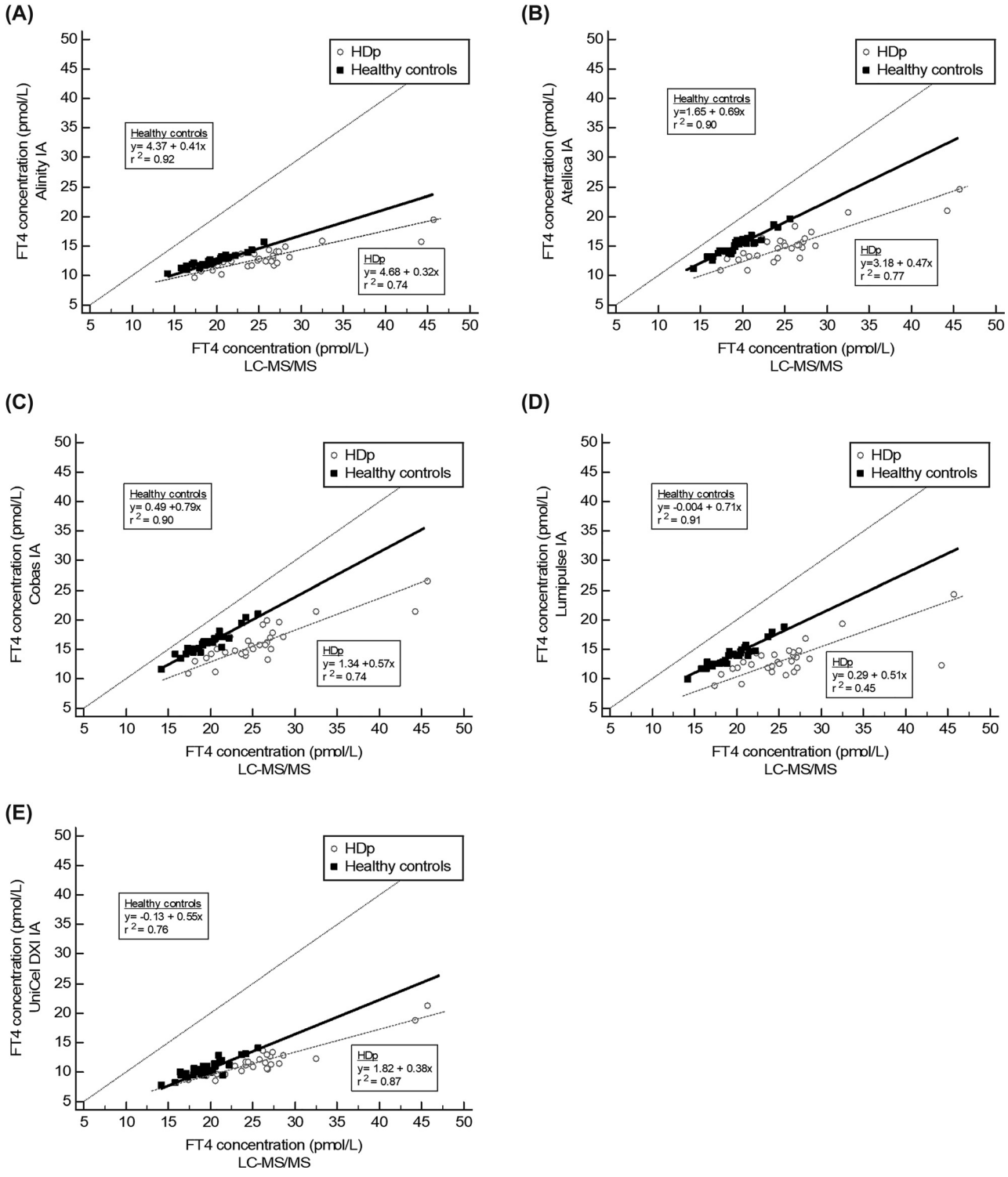

Figure 3: Passing and Bablok regression analyses for the 5 automated FT4 immunoassays in both healthy controls and HDp (experiment 1). On the x-axis, the FT4 concentrations were measured using LC-MS/MS and, on the y-axis, the FT4 concentrations measured using the respective immunoassays are shown. (A) Alinity; (B) Atellica; (C) Cobas; (D) Lumipulse; (E) UniCel DXI.

and they tested different IAs that are not on the market anymore. Previous literature reported an increased prevalence of thyroid disorders in HDp, especially the prevalence of (subclinical) hypothyroidism is higher compared to healthy, age-matched, controls [19–22]. As these recent studies predominantly use automated immunoassays which

may suffer from an underestimation of FT4 concentrations, prevalence might be different than reported in these studies. More research however is needed to get insight into the implication of this finding for clinical practice.

It is not yet known what interfering factor(s) caused the inaccuracies in testosterone, cortisol, and FT4 IA measurements in samples from people undergoing hemodialysis. Previous research showed that IAs have difficulties in measuring hormones when levels of binding globulins are deviant [3, 23, 24], although the influence of aberrant binding globulin concentrations on the IA reliability in samples of HDp is less clear. Cortisol-binding globulin (CBG) is not likely to affect the reliability of the cortisol IA measurement in HDp [25]. Dodd et al. [4] hypothesized that their observed positive deviation in HDp may be due to decreased cortisol clearance, therefore accumulation of cortisol and other steroids as well, which may lead to falsely high cortisol concentrations in IAs due to cross-reactivity. However, this hypothesis cannot be rejected nor confirmed based on our study. We did not find a significant difference in TBG concentration between HDp and healthy controls and the significant difference for SHBG concentrations was only small. Therefore, it can be questioned whether this small difference in SHBG concentration could explain testosterone IA inaccuracies. On the other hand, the binding protein albumin was significantly lower in HDp compared to healthy controls and might thus contribute to the IA inaccuracies of testosterone, cortisol, and FT4 in this group, which is corroborated by literature for cortisol and FT4 [9, 24, 26–30]. However, since the absolute difference in albumin concentration is limited, it should be kept in mind that this may not explain the observed IA bias completely.

Abovementioned associations to explain IA inaccuracies in HDp are not conclusive yet and all measurements were performed before the persons' hemodialysis session. The serum matrix effects of advanced renal failure in HDp changes after hemodialysis and the interfering factor(s) might be cleared by hemodialysis. We hypothesized the effect of hemodialysis might be beneficial on the accuracy of the IA measurements and we performed a second experiment where serum samples of HDp were collected and assessed both before and after their hemodialysis session. Our results showed no clear overall improvement of the IA reliability after hemodialysis. Matrix components can have a different impact on several IAs due to the variable working mechanism of IAs, explaining variation in IA accuracy in an altered matrix composition as HDp have. Last, we found no clear indications that either hemodialysis (HD) or hemodiafiltration (HDF) separately could improve the reliability of the tested testosterone, cortisol, or FT4 IAs. However, the HD and HDF groups were too small (HD, n=7; HDF, n=10) to make definite statements about some of the slight differences found. Therefore, we can conclude that the clearance of accumulated uremic toxins by hemodialysis is not the main solution to overcome IA inaccuracies in HDp and that other interfering factor(s) might be more important. This means that we cannot conclude that blood withdrawal prior to or after hemodialysis is the optimal timing of determining testosterone, cortisol, and FT4 in HDp, yet studies with larger patient populations should confirm this.

Some limitations should be mentioned as well. First, the ED-ID-LC-MS/MS that measured FT4 was not designed using uremic serum. However, this method adhered to the conventions of the IFCC, is known to be robust for matrix effects and FT4 concentrations remained stable upon hemodialysis (Jansen et al. under review). Therefore, we can conclude that it is justified to compare the commonly used FT4 IAs with this method. Second, it should be mentioned that the used LC-MS/MS for measuring both cortisol and testosterone, although extensively validated and comparable to other (inter)national LC-MS/MS methods, is not a reference method. It is, however, robust to matrix effects and can thus be reliably used for this method comparison. Third, experiment 2 only included the Alinity and Atellica IA, which could be seen as a limitation. However, this experiment was a proof of principle which was not set up to draw definite conclusions.

Conclusions

This study revealed that cortisol concentrations and testosterone concentrations in females can be overestimated and FT4 concentrations can be underestimated when measured using currently available immunoassays in serum samples of people with advanced renal failure undergoing hemodialysis. The observed inaccuracies in serum of HDp varied between all tested IAs and are clinically relevant for some immunoassays. It is not clear yet what interfering factor(s) cause these inaccuracies. Hemodialysis itself is clearly not the only solution in improving the accuracy. Therefore, we encourage manufacturers to improve their immunoassays in a way that not only these three hormone measurements, but all hormone measurements are less influenced by matrix effects. Furthermore, the influence of an altered matrix on IA performance should be studied in other (patient) groups characterized by an altered serum matrix to assess if caution in interpretation of their hormone measurements is necessary as well. Meanwhile, physicians and laboratory specialists should be aware that several IAs are less reliable

in the measurement of testosterone, cortisol, and FT4 in this specific group, since diagnostic or treatment decisions may depend on it.

Research funding: This research project was funded by the Landelijke Werkgroep Bindingsanalyse (LWBA). The LWBA played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

Author contributions: 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 local Institutional Review Board deemed the study exempt from review.

References

1. d'Herbomez M, Forzy G, Gasser F, Massart C, Beaudonnet A, Sapin R. Clinical evaluation of nine free thyroxine assays: persistent problems in particular populations. Clin Chem Lab Med 2003;41:942–7.
2. Heijboer AC, Blankenstein MA, Kema IP, Buijs MM. Accuracy of 6 routine 25-hydroxyvitamin D assays: influence of vitamin D binding protein concentration. Clin Chem 2012;58:543–8.
3. Heijboer AC, Savelkoul E, Kruit A, Endert E, Blankenstein MA. Inaccurate first-generation testosterone assays are influenced by sex hormone–binding globulin concentrations. J Appl Lab Med 2016;1:194–201.
4. Dodd A, Ducroq D, Neale S, Wise M, Mitchem K, Armston A, et al. The effect of serum matrix and gender on cortisol measurement by commonly used immunoassays. Ann Clin Biochem 2014;51:379–85.
5. Ward G, Simpson A, Boscato L, Hickman PE. The investigation of interferences in immunoassay. Clin Biochem 2017;50:1306–11.
6. Soldin SJ, Soldin OP. Steroid hormone analysis by tandem mass spectrometry. Clin Chem 2009;55:1061–6.
7. Stanczyk FZ, Clarke NJ. Advantages and challenges of mass spectrometry assays for steroid hormones. J Steroid Biochem Mol Biol 2010;121:491–5.
8. Heijboer AC, Zimmerman Y, Boer T, Bennink HC, Blankenstein MA. Peculiar observations in measuring testosterone in women treated with oral contraceptives supplemented with dehydroepiandrosterone (DHEA). Clin Chim Acta 2014;430:92–5.
9. Iitaka M, Kawasaki S, Sakurai S, Hara Y, Kuriyama R, Yamanaka K, et al. Serum substances that interfere with thyroid hormone assays in patients with chronic renal failure. Clin Endocrinol 1998;48:739–46.
10. Depreter B, Heijboer AC, Langlois MR. Accuracy of three automated 25-hydroxyvitamin D assays in hemodialysis patients. Clin Chim Acta 2013;415:255–60.
11. Tattersall JE, Ward RA, Canaud B, Blankestijn PJ, Bots M, Covic A, et al. Online haemodiafiltration: definition, dose quantification and safety revisited. Nephrol Dial Transplant 2013;28:542–50.
12. 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, $_{17}$OH-progesterone and aldosterone. Clin Chem Lab Med 2022;60:726–39.
13. Büttler R, Martens F, Ackermans M, Davison A, van Herwaarden A, Kortz L, et al. Comparison of eight routine unpublished LC-MS/MS methods for the simultaneous measurement of testosterone and androstenedione in serum. Clin Chim Acta 2016;454:112–8.
14. Collet S, Gieles NC, Wiepjes CM, Heijboer AC, Reyns T, Fiers T, et al. Changes in serum testosterone and adrenal androgen levels in transgender women with and without gonadectomy. J Clin Endocrinol Metab 2023;108:331–8.
15. Van Houcke SK, Van Uytfanghe K, Shimizu E, Tani W, Umemoto M, Thienpont LM. IFCC international conventional reference procedure for the measurement of free thyroxine in serum: International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) Working Group for Standardization of Thyroid Function Tests (WG-STFT). Clin Chem Lab Med 2011;49:1275–81.
16. Committee for Standardization of Thyroid Function Tests (C-STFT). IFCC–Scientific Division (SD). SD Committees [Online]. Available from: http://www.ifcc.org/ifcc-scientific-division/sdcommittees/c-stft/ [Accessed 2 Jun 2022].
17. Rosner W, Vesper H. Toward excellence in testosterone testing: a consensus statement. J Clin Endocrinol Metab 2010;95:4542–8.
18. Hawley JM, Owen LJ, Lockhart SJ, Monaghan PJ, Armston A, Chadwick CA, et al. Serum cortisol: an up-to-date assessment of routine assay performance. Clin Chem 2016;62:1220–9.
19. Da Costa ABBA, Pellizzari C, Carvalho GA, Sant'Anna BC, Montenegro RL, Zammar Filho RG, et al. High prevalence of subclinical hypothyroidism and nodular thyroid disease in patients on hemodialysis. Hemodial Int 2016;20:31–7.
20. Shantha GPS, Kumar AA, Bhise V, Khanna R, Sivagnanam K, Subramanian KK. Prevalence of subclinical hypothyroidism in patients with end-stage renal disease and the role of serum albumin: a cross-sectional study from south India. Cardiorenal Med 2011;1:255–60.
21. Kutlay S, Atli T, Koseogullari O, Nergizoglu G, Duman N, Gullu S. Thyroid disorders in hemodialysis patients in an iodine-deficient community. Artif Organs 2005;29:329–32.
22. Lo JC, Chertow GM, Go AS, Hsu C-Y. Increased prevalence of subclinical and clinical hypothyroidism in persons with chronic kidney disease. Kidney Int 2005;67:1047–52.
23. Vos MJ, Bisschop PH, Deckers MML, Endert E. The cortisol-CBG ratio affects cortisol immunoassay bias at elevated CBG concentrations. Clin Chem Lab Med 2017;55:e262–4.
24. Gant Kanegusuku A, Araque KA, Nguyen H, Wei B, Hosseini S, Soldin SJ. The effect of specific binding proteins on immunoassay measurements of total and free thyroid hormones and cortisol. Ther Adv Endocrinol Metab 2021;12:1–9.
25. Rosman PM, Benn R, Kay M, Tito J, Wallace EZ. Cortisol binding in uremic plasma. I. Absence of abnormal cortisol binding to cortisol binding to corticosteroid-binding globulin. Nephron 1984;37:160–5.
26. Ooi DS, Mahadevan MS, Greenway DC, Gertler SZ. Evaluation of four commercially available assays for free thyroxin. Clin Chem 1988;34:2302–7.
27. Bayer MF. Free thyroxine results are affected by albumin concentration and nonthyroidal illness. Clin Chim Acta 1983;130:391–6.

28. Welsh KJ, Stolze BR, Yu X, Podsiadlo TR, Kim LS, Soldin SJ. Assessment of thyroid function in intensive care unit patients by liquid chromatography tandem mass spectrometry methods. Clin Biochem 2017;50:318–22.
29. Van Deventer HE, Mendu DR, Remaley AT, Soldin SJ. Inverse log-linear relationship between thyroid-stimulating hormone and free thyroxine measured by direct analog immunoassay and tandem mass spectrometry. Clin Chem 2011;57:122–7.
30. Rosman PM, Benn R, Kay M, Wallace EZ. Cortisol binding in uremic plasma. II. Decreased cortisol binding to albumin. Nephron 1984;37: 229–31.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/cclm-2022-1133).