Clinica Chimica Acta 415 (2013) 109–117

Contents lists available at SciVerse ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/clinchim

Development and validation of a serum total testosterone liquid chromatography–tandem mass spectrometry (LC–MS/MS) assay calibrated to NIST SRM 971

Deborah French *

Department of Laboratory Medicine, University of California-San Francisco, San Francisco, CA 94107, United States

ARTICLE INFO

Article history:
Received 28 July 2012
Received in revised form 4 October 2012
Accepted 5 October 2012
Available online 16 October 2012

Keywords:
Total testosterone
LC–MS/MS
Immunoassay
Standardization

ABSTRACT

Background: At our institution, serum testosterone in adult males is measured by immunoassay while female and pediatric specimens are sent to a reference laboratory for liquid chromatography–tandem mass spectrometry (LC–MS/MS) analysis due to low concentrations. As this is of significant cost, a testosterone LC–MS/MS assay was developed in-house.
Methods: A 5500 QTRAP® using electrospray ionization and a Shimadzu Prominence with a C18 column were used. Gradient elution with formic acid, water and methanol:acetonitrile at 0.5 ml/min had a 7-min run-time. A liquid–liquid extraction with hexane:ethyl acetate was carried out on 200 μl of serum. Multiple reaction monitoring was employed.
Results: Sample preparation took ~80 min for 21 samples. Six calibrators were used (0–1263 ng/dl; concentration assigned by NIST SRM 971) with 3 quality controls (9, 168 and 532 ng/dl). The limits of detection and quantitation were 1 and 2 ng/dl respectively. Extraction recovery was ~90% and ion suppression ~5%. Within-run and total precision studies yielded <15% CV at the limit of quantitation and <7% CV through the rest of the linear range. Isobaric interferences were baseline separated from testosterone. Method comparisons between this assay, an immunoassay, and another LC–MS/MS assay were completed.
Conclusions: An accurate and sensitive LC–MS/MS assay for total testosterone was developed. Bringing this assay in-house reduces turnaround time for clinicians and patients and saves our institution funds.

1. Introduction

Testosterone is one of the most commonly measured steroid hormones and is present in very low concentrations of around 2–45 ng/dl in the serum of females, and pediatric males of certain age groups [1–3]. Reasons for measuring total testosterone clinically include evaluation of hypogonadism in males, and in females to investigate infertility, amenorrhea, hirsutism or for diagnosis of polycystic ovarian syndrome [4–7]. In children, testosterone is measured in cases of delayed or precocious puberty and is used in the laboratory investigation of infants with ambiguous genitalia [7,8].

Serum total testosterone is most often measured clinically by immunoassay, as documented in the College of American Pathologists (CAP) proficiency testing participants (98.6% as of September 2011). However, several publications have shown that although immunoassays have sufficient accuracy and precision for measuring total testosterone in males, they are lacking in these parameters when it comes to females and pediatrics due to the low concentrations normally found in these patients [9–11]. These findings prompted The Endocrine Society to release a position statement in 2007 stating that they believe the best way of measuring total testosterone is using a method that involves extraction, chromatography and mass spectrometry [12].

A number of LC–MS/MS assays used to measure testosterone have been reported in the literature, each of which is developed in a different way, using different extraction procedures and indeed different types of ionization, although every paper reviewed herein used MS analysis in the positive mode [1,13–17]. The method of Kushnir et al. used liquid–liquid extraction with methyl t-butyl ether, hydroxylamine derivatization, 2-dimensional and then analytic chromatographic separation followed by MS analysis using electrospray ionization (ESI) [1]. The authors report a limit of quantitation (LOQ) of 1 ng/dl, an upper limit of quantitation (ULOQ) of 2500 ng/dl with no interference from 50 steroids and steroid metabolites. The assay published by Salameh et al. used online extraction by high-turbulence flow liquid chromatography followed by chromatographic separation and MS analysis using atmospheric pressure chemical ionization (APCI) with

Abbreviations: NIST SRM 971, National Institute of Standards and Technology Standard Reference Material 971.

* University of California-San Francisco, 185 Berry Street, Suite 290, San Francisco, CA 94107, United States. Tel.: +1 415 514 6627; fax: +1 415 353 1178.
E-mail address: deborah.french@ucsf.edu.

0009-8981/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cca.2012.10.007

a total run time of 1.15 min per sample [13]. The limit of quantitation was determined to be 0.3 ng/dl and the ULOQ was 2000 ng/dl. Savolainen et al. also used online extraction in their assay, but in this case it was a solid phase extraction method [14]. The extraction was followed by LC–MS/MS analysis using ESI and a run time of 6.7 min. The LOQ was determined to be 7 ng/dl, and the ULOQ was 1000 ng/dl. The LC–MS/MS assay of Shiraishi et al. for testosterone used a liquid liquid extraction with ethyl acetate and hexane that was performed twice and then removal of acidic contaminants was achieved by the addition of sodium hydroxide [15]. The extracted samples were then injected onto the LC system and MS analysis using ESI was carried out, with a run time of 18.5 min per sample. The LOQ and ULOQ were 2 and 2000 ng/dl respectively. Cawood et al. used protein precipitation, followed by LC–MS/MS analysis using ESI with a run time of 4.75 min [16]. The reported LOQ was 9 ng/dl with a ULOQ of 2882 ng/dl. Finally, the LC–MS/MS assay of Guo et al. used protein precipitation, online clean-up of the injected sample followed by chromatographic separation and MS analysis using atmospheric pressure photoionization (APPI) [17]. The limit of detection was determined to be 0.15 ng/dl. These authors have also reported that using the APPI source, a soft ionization source, instead of ESI or APCI for analysis of testosterone can lead to a cleaner baseline and thereby increase the signal to noise ratio of the chromatographic peaks which is a definite advantage when measuring potentially low level analytes such as testosterone [18].

Although mass spectrometry has high specificity, it is not without its limitations for testosterone testing. Due to other endogenous compounds that are the same (e.g., dehydroepiandrosterone; DHEA), or very similar in structure and molecular weight, testosterone must be chromatographically separated from these in order to obtain specific and accurate quantification by liquid chromatography–tandem mass spectrometry (LC–MS/MS), especially since some of these compounds are present at much higher concentrations than testosterone. Additionally, the gel in serum separator tubes in which patient blood samples are routinely collected causes interference in one of the main testosterone transitions and so clinical laboratories are required to ask clinicians to collect the sample in an alternative tube type unless they run a long chromatographic method or refrain from using these specific transitions (mass to charge ratios of 289/97 and 289/109) [19]. Matrix effects from phospholipids are another concern and they may cause a high chromatographic baseline, retention time alterations and suppression of the testosterone signal [20,21].

Another issue in total testosterone testing is the lack of standardization between different assay platforms and indeed between users of the same platform, specifically LC–MS/MS methods [12,22]. The National Institute of Standards and Technology aimed to help with this issue by releasing a Standard Reference Material 971 that contains certified concentrations of testosterone in serum [23]. This can be used by immunoassay manufacturers, or indeed mass spectrometry assay developers as a way of standardizing total testosterone assays to one measurement. The Centers for Disease Control are very aware of the lack of standardization of steroid hormone assays, and in 2010, implemented the hormone standardization program (HoST) [24]. Laboratories using self-developed methods, or assay manufacturers, can enroll in this program with an aim to improve and monitor accuracy compared to a reference method for total testosterone; however, there is a substantial cost to participate in this program which may be prohibitive for smaller laboratories.

At our institution, total testosterone in male patients is measured by chemiluminescent immunoassay, while for female and pediatric patients, samples are sent to a reference laboratory for LC–MS/MS analysis. As this practice is of significant cost, our aim was to develop a LC–MS/MS assay in-house that is suitable for the measurement of testosterone in all samples and has the calibration verified by the NIST SRM 971.

2. Materials

2.1. Experimental

Mass spectrometry-grade solvents and water were from VWR International (Brisbane, CA) or from Thermo Fisher Scientific (Pittsburgh, PA). Formic acid was from Sigma-Aldrich (St Louis, MO). Testosterone, 16, 16, 17-d3 labeled testosterone, dehydroepiandrosterone, estradiol and epitestosterone were from Cerilliant Corporation (Round Rock, TX). Hormones in frozen human serum standard reference material (NIST SRM 971) was from The National Institute of Standards and Technology (Gaithersburg, MD). Human serum calibrators containing testosterone were from UTAK Laboratories Inc. (Valencia, CA) and the concentration of these was verified by running the male and female serum samples of the NIST SRM 971 that have assigned testosterone concentrations on the LC–MS/MS assay. Lyphochek® Immunoassay Plus quality control material Levels 1 and 2 were from BioRad Laboratories (Irvine, CA). Double charcoal stripped human serum was obtained from Golden West Biologicals Inc. (Temecula, CA).

2.2. Patient samples

Institutional review board approval was obtained from the University of California, San Francisco Committee on Human Research for this study. Sixty-six patients who had ultra-sensitive total testosterone (LOQ = 1 ng/dl), performed by LC–MS/MS at a reference laboratory, ordered on their serum (from serum separator tubes) in 2011 and had sufficient sample volume remaining were included in this study. Four of the 66 patient specimens were from pediatric patients (two 4 years, one 7 years and one 13 years). Further, 30 patients who had total testosterone ordered on their serum (from serum separator tubes) in December 2011 performed on the Siemens Centaur® XP immunoassay by UCSF Clinical Laboratories were retrospectively included in this study. One of the 30 patient specimens was from a pediatric patient (13 years). Of these 30 patient specimens, 17 had sufficient sample volume remaining and were extracted using the liquid–liquid extraction from the LC–MS/MS method. These 17 specimens, along with the 6 calibrators from the LC–MS/MS assay were then run on the Siemens Centaur® XP immunoassay. The total testosterone concentration for each of these 96 patients using the methodologies mentioned above (and described in the method comparison study section below) was recorded and compared to the total testosterone concentration obtained from the newly developed total testosterone LC–MS/MS assay. Demographic information for the patients who the samples were obtained from is documented in Table 1.

2.3. Liquid–liquid extraction

Twenty-five microliters of 10 ng/ml internal standard (D3-testosterone in methanol) was added to 200 µl of serum and vortexed. One milliliter of 90:10 hexane:ethyl acetate was added to each tube, vortexed for 2 min, left at room temperature for 5 min and then centrifuged at 3000 rpm for 10 min. The tubes were then placed in dry ice until the aqueous bottom layer froze and the organic top layer was poured into a fresh tube. The solvent was evaporated to dryness under nitrogen at 45 °C and the sample reconstituted in 125 µl of methanol and water (60:40) as during method

Table 1
Demographic information for the patient samples utilized in the method comparison studies.

Method	Age (years)			Gender	
	Mean	Median	Range	Male	Female
LC–MS/MS	45.3	43.5	4–89	19	47
Centaur	55.5	59.5	13–76	29	1

development, this ratio gave higher peak areas for testosterone while retaining good peak shape than when the samples were injected in the initial mobile phase starting conditions (which are documented in Section 2.4).

2.4. Liquid chromatography

A Shimadzu Prominence UFLC system (Shimadzu Scientific Instruments, Pleasanton, CA) and a Kinetex® C18, 3.0×100 mm, 2.6 μm column (Phenomenex, Torrance, CA) maintained at 40 °C were used. Mobile phases consisted of 0.1% formic acid in water (mobile phase A) and 0.1% formic acid in 70:30 methanol:acetonitrile (mobile phase B), and gradient elution starting at 10% mobile phase B ramping with a constant slope to 95% mobile phase B at 5.7 min at 0.5 ml/min was used with a run time of 7 min including re-equilibration to the initial conditions. The injection volume was 50 μl.

2.5. Mass spectrometry

An ABSCIEX 5500 QTRAP® was utilized with an electrospray ionization source (ESI) and positive polarity. The source parameters included curtain gas of 30 psi, ion spray voltage of 5000 V, temperature of 650 °C and medium collision gas. Multiple reaction monitoring was used, with two transitions (quantifier and qualifier) for both testosterone and the internal standard (D3-testosterone; Table 2). The raw signal for testosterone was normalized to that of the internal standard and concentrations calculated from a six-point calibration curve (0–1263 ng/dl). Three levels of quality control were used with established means of 9, 168 and 532 ng/dl respectively (the lowest control was made in-house by spiking double charcoal stripped serum with testosterone and the other 2 controls were from BioRad Laboratories). Ion ratios (peak area of the qualifier ion divided by peak area of the quantifier ion) were calculated for both testosterone and the internal standard in 5 calibrators (excluding the 0 calibrator) to give a mean and an allowable range of +/− 20%. None of the controls or patient samples in the method comparison studies had ion ratios >20% from the mean. A calibration curve and the three quality control samples were run with each batch of patient samples.

2.6. LC–MS/MS assay validation

Ion suppression and extraction recovery were assessed by adapting a previously published method [25]. Briefly, 100 ng/dl of testosterone was spiked into 3 different aliquots of double charcoal stripped serum and 3 different aliquots of water before the liquid–liquid extraction was carried out. Three further aliquots of double charcoal stripped serum were taken through the extraction at the same time as the spiked serum and water, and before the dry-down step, 100 ng/dl of testosterone was added to the serum only samples. All 9 samples were then run on the LC–MS/MS assay. To calculate the % ion suppression, the peak areas of the serum samples spiked before extraction were divided by the peak areas of the water samples spiked before extraction. To calculate the % extraction recovery, the peak areas of the serum samples spiked before extraction were divided by the peak areas of the serum samples spiked after extraction. The linearity of the assay was determined by running the 6 calibrators on the LC–MS/MS assay in triplicate, noting the assigned concentration and the concentration back calculated using linear regression analysis in the LC–MS/MS software. The results were then entered into the linearity function in EP Evaluator for determination of clinical linearity in which a best-fit line is plotted through the specimen points, with a user-defined allowable systematic error for each testosterone concentration; in this case 2 ng/dl or 10%. In order to determine an appropriate dilution strategy for patient samples that have a testosterone concentration that exceeds the linear range of the assay (>1263 ng/dl), testosterone was spiked into double charcoal stripped serum at 1500, 2000, 2500 and 3000 ng/dl. These samples were then diluted 1:5 using water and were extracted by LLE and run on the LC–MS/MS assay as would be the case for any patient sample. The concentration of each of these diluted high concentration samples was determined by the calibration curve to ensure that the correct concentration would be obtained using this dilution strategy should the testosterone concentration of a patient sample fall above 1263 ng/dl. To determine the limit of detection (LOD) and limit of quantitation (LOQ), testosterone was spiked into double charcoal stripped human serum at 5, 4, 3, 2 and 1 ng/dl and run on 20 different days on the LC–MS/MS assay. The criteria for the LOQ was total precision of ≤20% %CV and for the LOD was a signal-to-noise ratio of greater than 3:1. Within-run precision experiments were performed by extracting 10 independent aliquots of each of the 3 quality controls within 1 run and comparing the concentration assigned based upon the standard curve. The mean, SD and %CV were then calculated. Total precision experiments were performed by extracting the 3 quality controls on 35 different days and comparing the concentration assigned based upon the standard curve. The mean, SD and %CV were then calculated. Interference studies were carried out using hemolysed, lipemic, icteric and cholesterolemic samples. Additionally, the isobaric compounds dehydroepiandrosterone and epitestosterone were investigated as potential interferences, as well as the structurally similar compounds androstenedione and estradiol. Five patients were identified that had serum collected in both serum separator tubes (gold top) and in clot activator tubes (red top) during the same blood collection in order to test whether the long chromatographic run was sufficient to separate the interference in the testosterone transition seen in serum collected from the gold top tubes, from the testosterone peak. Deming regression and Bland–Altman analyses were used to compare the three different analytical methods currently in clinical use to the developed LC–MS/MS assay described here. Carryover studies were undertaken by extracting six low concentration (10 ng/dl) testosterone spiked double charcoal stripped human serum samples and one high concentration (5000 ng/dl) testosterone spiked double charcoal stripped human serum sample. Three low concentration samples were injected, followed by the high concentration sample, and then the other three low concentration samples. The concentrations of the low concentration samples calculated from the standard curve injected after the high concentration sample were compared to those injected before the high sample to determine if there was any carryover. Testosterone stability was determined under a number of different conditions that may be encountered in a clinical setting: serum left on red cells for 3 days at 4 °C, serum separated from red cells for 2 weeks at 4 °C, after liquid–liquid extraction in 60:40 methanol:water for 2 weeks at 4 °C. It is already reported in the literature that testosterone is stable

Table 2
Multiple reaction monitoring transitions and voltages for testosterone and the deuterated internal standard (D3-testosterone).

Analyte name	Q1 (*m/z*)	Q3 (*m/z*)	DP (V)	EP (V)	CXP (V)	CE (eV)
Testosterone 1 (quantifier transition)	289.3	97.0	100	10	12	30
Testosterone 2 (qualifier transition)	289.3	109.1	100	10	18	30
D3-testosterone 1 (quantifier transition)	292.3	97.1	106	10	16	37
D3-testosterone 2 (qualifier transition)	292.3	109.2	106	10	16	37

Q1: quadrupole 1; *m/z*: mass to charge ratio; Q3: quadrupole 3; DP: declustering potential; V: volts; EP: exit potential; CXP: collision cell exit potential; CE: collision energy; eV: electronvolts.

in serum for at least 40 years at −20 °C and so this test was not carried out during our method validation [26].

2.7. Method comparison study

Testosterone was quantified in 66 patient samples by a clinically validated LC–MS/MS assay using electrospray ionization with positive polarity at a reference laboratory [1]. This method utilized liquid–liquid extraction with methyl t-butyl ether followed by hydroxylamine derivatization of testosterone. The reportable range of this assay is 1–2500 ng/dl, has a LOQ of 1 ng/dl and has no cross-reactivity with more than 50 steroid and steroid-metabolites. This LC–MS/MS assay is certified by the CDC HoST program as of May 2012.

Testosterone was quantified in 30 patient samples by the ADVIA Centaur® XP Immunoassay system (Siemens Healthcare Diagnostics, Deerfield, IL) at the University of California San Francisco Clinical Laboratories. This assay is a competitive immunoassay using direct chemiluminescent technology. Testosterone in the patient sample competes with acridinium ester-labeled testosterone in the reagent for a limited amount of polyclonal rabbit anti-testosterone antibody bound to monoclonal mouse anti-rabbit antibody, coupled to paramagnetic particles in the solid phase. Testosterone releasing agent releases bound testosterone from the endogenous binding proteins in the sample. The assay is linear from 10 to 1500 ng/dl, the reportable range is from 10 to 7500 ng/dl with two calibrators at 67.3 and 698 ng/dl and two controls with means of 137 and 513 ng/dl. This method shows no significant interference up to 500 mg/dl hemoglobin, 300 mg/dl triglycerides and 20 mg/dl bilirubin.

2.8. Data analysis

All analyses and figure preparation were carried out using EP Evaluator® Release 9 (Data Innovations Inc., South Burlington, VT), the statistical environment R2.15.1 (R Core Development Team, http://www.r-project.org) or Microsoft Office Excel 2003 (Microsoft Corp., Redmond, WA).

3. Results

3.1. Analytical results

The newly developed LC–MS/MS assay for serum total testosterone performed well with a mean extraction recovery of 90% and mean ion suppression of 5%. The assay was linear for testosterone between 2 and 1263 ng/dl (slope: 0.99, intercept: 0.52, $R^2 = 0.99$; Fig. 1) with a mean total error of 0.52 ng/dl or 2.6% across the linear range. The calculated recovery for the measured concentrations compared to the assigned concentrations of the calibrators ranged from 97 to 104%. The LOD was 1 ng/dl and the LOQ was determined to be 2 ng/dl based on a total precision of 14.5% CV by spiking double charcoal stripped serum with testosterone. Additionally, the signal to noise ratio (S/N) of the testosterone peak in 19 patients with calculated testosterone concentrations below 10 ng/dl (median = 4 ng/dl, interquartile range = 4, range = 1–10 ng/dl) was determined using the Analyst® software. The median S/N ratio was 30:1 (interquartile range = 73, range = 10:1–25:1) and all results were above 10:1. The reportable range of the assay was determined to be 2–3000 ng/dl. This is based upon the concentration of the 1:5 dilution of the serum with water before extraction being calculated as +/− 10% of the actual concentration spiked into the serum. Therefore, if the testosterone concentration of a patient sample is >1263 ng/dl, the original serum sample will be diluted 1:5 with water, extracted by LLE and run on the LC–MS/MS assay again. The calculated concentration will be multiplied by 5 in order to give an exact value. If the calculated concentration based upon the dilution is >3000 ng/dl, the testosterone concentration for that patient will be reported as >3000 ng/dl. Within-run precision studies yielded CVs of 3.2%, 5.3% and 4.0% for the quality controls with calculated mean concentrations of 10, 172 and 538 ng/dl respectively. Total precision studies yielded CVs of 6.9%, 4.8% and 4.7% for the quality controls with calculated mean concentrations of 9, 168 and 532 ng/dl respectively.

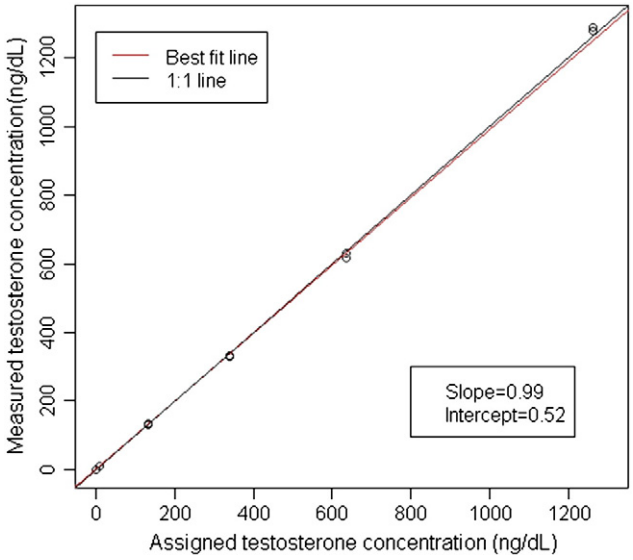

Fig. 1. Linear regression analysis of the actual concentration of the testosterone calibrators versus the calculated concentration as measured by the response of the mass spectrometer (testosterone peak area divided by internal standard peak area).

Testosterone was quantified in five patients who had serum drawn in both a clot activator tube (red top) and a serum separator tube (gold top) during one blood collection. As can be seen in Fig. 2a, in the serum collected in a red top tube, there is one obvious peak corresponding to testosterone when we monitor the testosterone quantifier transition. Fig. 2b shows the same testosterone quantifier transition being monitored in serum from a patient collected in a gold top tube. As can be seen, there are a number of other much larger, potentially interfering peaks, but the testosterone peak is baseline separated from them. Dehydroepiandrosterone and epitestosterone were chromatographically separated from testosterone and so did not interfere with the quantification (Fig. 3). Estradiol and androstenedione did not interfere with testosterone quantification at 800 pg/ml and 500 ng/dl respectively. The assay was not affected by hemolysis (up to 320 mg/dl of hemoglobin), lipemia (up to 1000 mg/dl triglycerides), icterus (up to 20 mg/dl bilirubin) or cholesterolemia (up to 500 mg/dl cholesterol). Additionally, no carryover was observed after injection of a sample containing 5000 ng/dl testosterone. Testosterone was found to be stable in patient samples under all conditions that were tested with recoveries between 97 and 111%.

3.2. Method comparison studies

3.2.1. New LC–MS/MS assay vs chemiluminescent immunoassay

Thirty patient samples covering a concentration range of 10 to 722 ng/dl, with a mean concentration of 294 ng/dl and a median concentration of 297 ng/dl based upon the immunoassay results were run on the newly developed LC–MS/MS assay. The correlation across the whole range yielded a Deming regression slope of 1.09 (95% confidence interval (95% CI): 0.98 to 1.19), an intercept of −8.16 (95% CI: −45.86 to 29.55), and r = 0.97 (Fig. 4a). The median percent bias was 3.48% (95% CI: −9.50 to 8.72) and the mean percent bias was −11.35% (95% CI: −24.69 to 1.99). When 9 of these samples with

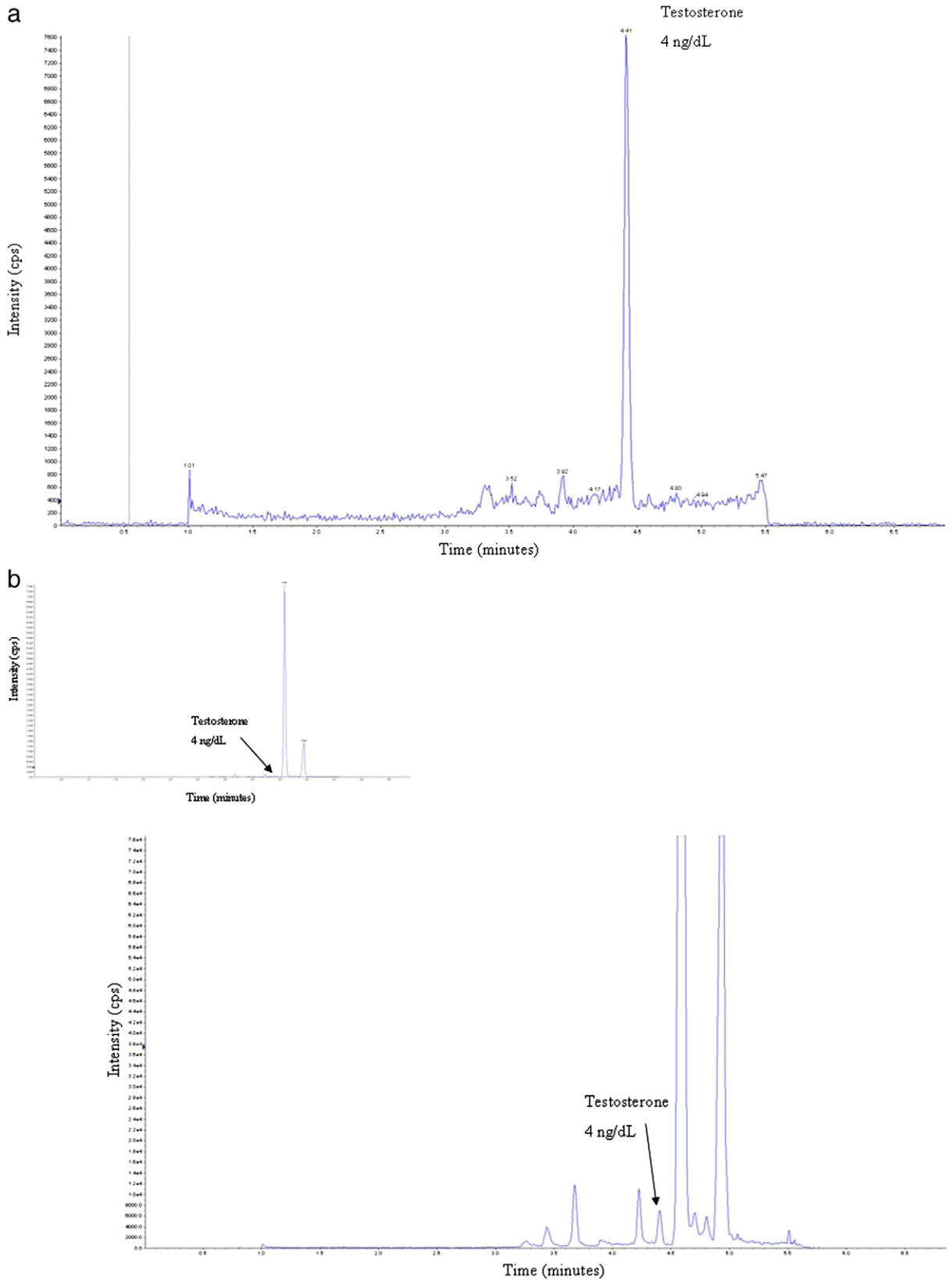

Fig. 2. Comparison between collecting serum in a clot activator tube versus a serum separator tube. a) The extracted ion chromatogram of a patient sample collected in a clot activator tube with a concentration of 4 ng/dl. b) The extracted ion chromatogram of a sample collected from the same patient at the same time in a serum separator tube.

concentrations <60 ng/dl were compared, the correlation yielded a Deming regression slope of 0.63 (95% CI: 0.23 to 1.02), an intercept of −5.03 (95% CI: −16.99 to 6.94), and r = 0.78. At these lower testosterone concentrations, the new LC–MS/MS assay showed a marked negative percent bias compared to the immunoassay method except for one sample (Fig. 4b). The median percent bias was −62.75% (95% CI: −70.95 to −61.67) and the mean percent bias was −59.86% (95% CI: −72.77 to −46.95).

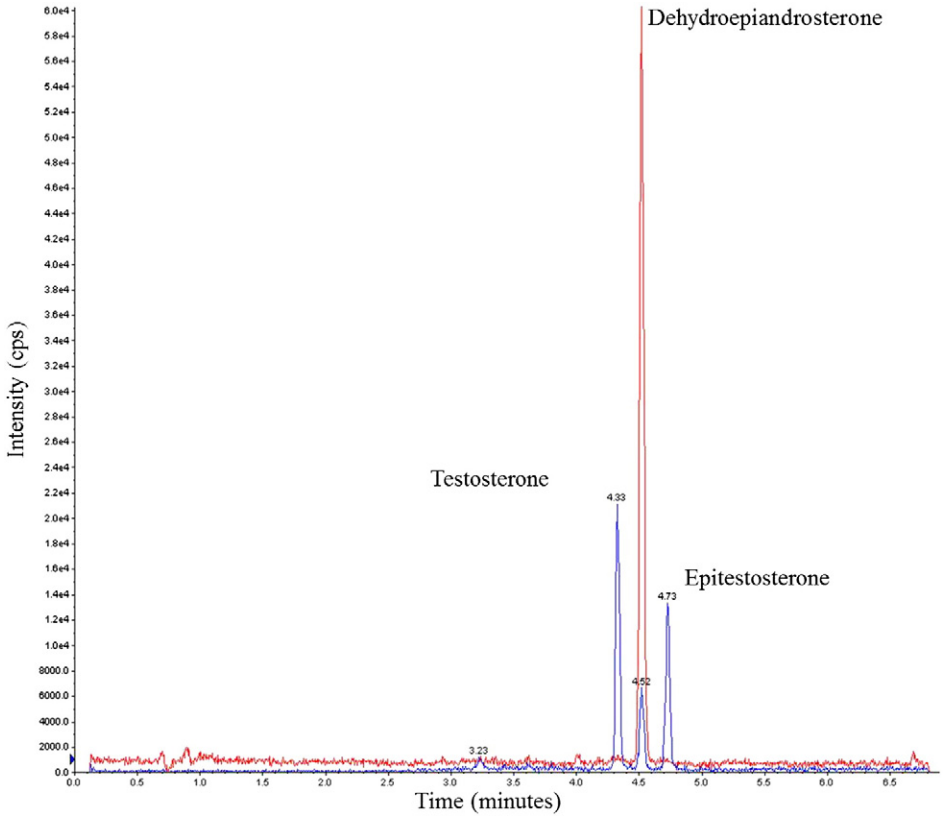

Fig. 3. Example chromatogram: chromatogram displaying the baseline separation of testosterone, dehydroepiandrosterone and epitestosterone. The blue trace shows the peaks obtained from monitoring the testosterone quantifier transition (*m/z* 289.3–97.0) and the red trace shows the peak obtained from monitoring a dehydroepiandrosterone transition (*m/z* 289.2–271.1).

3.2.2. New LC–MS/MS assay vs LC–MS/MS assay at a reference laboratory

Sixty-six patient samples covering a concentration range of 2–948 ng/dl, with a mean concentration of 169 ng/dl and a median concentration of 57 ng/dl based upon the reference laboratory results were run on the newly developed LC–MS/MS assay. The correlation across the whole concentration range yielded a Deming regression slope of 1.04 (95% CI: 1.01 to 1.06), an intercept of 0.91 (95% CI: −5.67 to 7.49), and r = 0.99. Across the range of testosterone concentrations, the new LC–MS/MS method showed a mostly positive percent bias (Fig. 5a). The median percent bias was 4.32% (95% CI: 0.00 to 7.36) and the mean percent bias was 2.81% (95% CI: −0.76 to 6.39). When 45 of these samples with concentrations <100 ng/dl were compared, the correlation yielded a Deming regression slope of 1.03 (95% CI: 0.97 to 1.10), an intercept of 0.08 (95% CI: −3.15 to 3.19), and r = 0.98. At these lower testosterone concentrations, the new LC–MS/MS assay again showed a mostly positive percent bias (Fig. 5b). The median percent bias was 4.08% (95% CI: −3.97 to 8.33) and the mean percent bias was 1.84% (95% CI: −3.13 to 6.80).

4. Discussion

Standardization of laboratory tests has recently been a topic of great discussion. A notable example is that of the steroid hormones where standardization of testosterone measurement has initially been the focus of a number of organizations, including The Endocrine Society, The Centers for Disease Control and Prevention, the American Association for Clinical Chemistry and the National Institute for Standards and Technology.

A number of articles have been published highlighting the variation between total testosterone immunoassays and LC–MS/MS assays [9–12] and indeed between different LC–MS/MS assays [22]. A further example of the discrepant quantification lies in the CAP proficiency testing results. In the September 2011 survey, approximately 1500 laboratories participated, of which approximately 99% were immunoassay methods and the remaining 1% were mass spectrometry methods. In the first challenge sample, the mean reported for each method group ranged from 98 to 476 ng/dl; in the second sample, the mean reported for each method group ranged from 145 to 518 ng/dl; and in the third sample, the mean reported for each method group ranged from 363 to 1333 ng/dl. The issue of inter-method standardization is very concerning, although the commutability of CAP samples may add to the lack of comparability in this survey. Due to this inherent problem, our laboratory wanted to use the NIST SRM 971 to assign values to the purchased calibrators in order for our assay to be more accurate.

As can be seen from the method comparison data, our assay compared well to a clinical validated LC–MSMS assay certified by the CDC HoST program. It also compared well to the Siemens Centaur XP assay in samples with testosterone concentrations >60 ng/dl. Since the analytical sensitivity of the Siemens Centaur assay was 10 ng/dl, it was expected that the LC–MS/MS assay would not compare well to this assay at concentrations <10 ng/dl, but we found that it did not compare well at concentrations <60 ng/dl — a value at the high end of the normal reference range for females. However, it must be mentioned that only 8 out of 30 samples had testosterone concentrations <60 ng/dl.

Our LC–MS/MS method has some advantages over previously published, clinically utilized LC–MS/MS methods [1,13–17]. The primary advantage is that no derivatization was required to obtain the limit of quantitation of 2 ng/dl which is required to accurately measure total testosterone in female and pediatric specimens. Additionally, using a simple, low cost liquid–liquid extraction the required

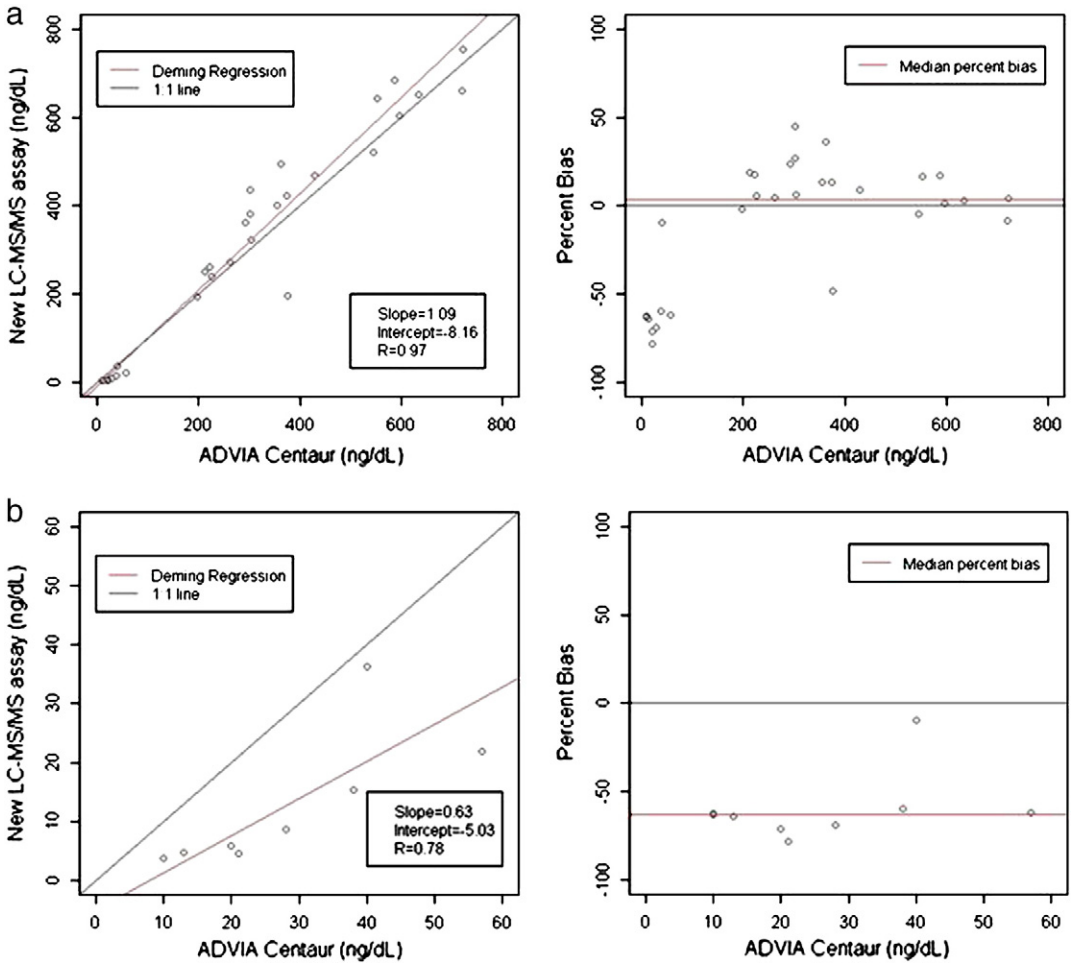

Fig. 4. Comparison between the new LC–MS/MS assay and an immunoassay. a) Scatter plot and percent bias plot of the method comparison between the ADVIA Centaur® XP immunoassay and the new LC–MS/MS assay using 30 patient samples (10–722 ng/dl); b) scatter plot and percent bias plot of a concentration sub-range (<60 ng/dl; n = 9) of the method comparison between the ADVIA Centaur® XP immunoassay and the new LC–MS/MS assay.

sensitivity was achieved without having to purchase additional expensive online extraction equipment. Use of the electrospray ionization source in this method is also an advantage given that it is the source that most mass spectrometry operators are familiar with and it is one that is most commonly utilized for other applications, such as toxicology testing. Another advantage is that using the 7-min chromatography method, gel serum separator tubes can be used to collect the patient specimen as the interference that is present from these tubes in the testosterone MRM transitions is separated from the testosterone peak [19]. Since this is the tube type that patient samples are normally collected in, the laboratory does not need to ask for the sample to be collected in another tube type making blood collection simpler for the clinicians and nurses and also there is no need to take a separate sample from the patient reducing the number of blood draws. Although it should be noted that changes in the lot number of serum separator tubes have been known to cause interference in mass spectrometry transitions that were not previously affected [personal communication]. Further, given the reportable range of 2–3000 ng/dl (upon dilution), this assay can measure total testosterone in all clinical samples, negating the need for different assays for male, female and pediatric samples. Finally, the calibration of our LC–MS/MS method was verified by the NIST SRM 971 assuring the accuracy of our method.

The disadvantages of our method are the sample volume of 200 µl, which for pediatric patients is a fairly significant quantity of serum. Additionally, liquid–liquid extraction requires the use of a number of organic solvents and therefore also requires the disposal of these solvents and it is also a fairly labor intensive process taking approximately 1 h and 20 min for the extraction of 21 patient samples, 6 calibrators and 3 controls. Further, as mentioned above, a 7-min chromatography method is utilized which is prohibitive if high throughput is required.

5. Conclusions

A sensitive LC–MS/MS assay was developed for the detection of total testosterone in all patients. The accuracy of this method was verified using the NIST SRM 971. Bringing this assay in-house reduces the turnaround time for our patients and clinicians, and additionally saves our institution valuable funds.

Acknowledgments

Kirk Sujishi and Ted Kurtz are acknowledged for insightful discussions and advice, and Witold Woroniecki and Hua-Fen Liu for technical assistance. This work was funded by the University of California San Francisco Medical Center and the Department of Laboratory Medicine.

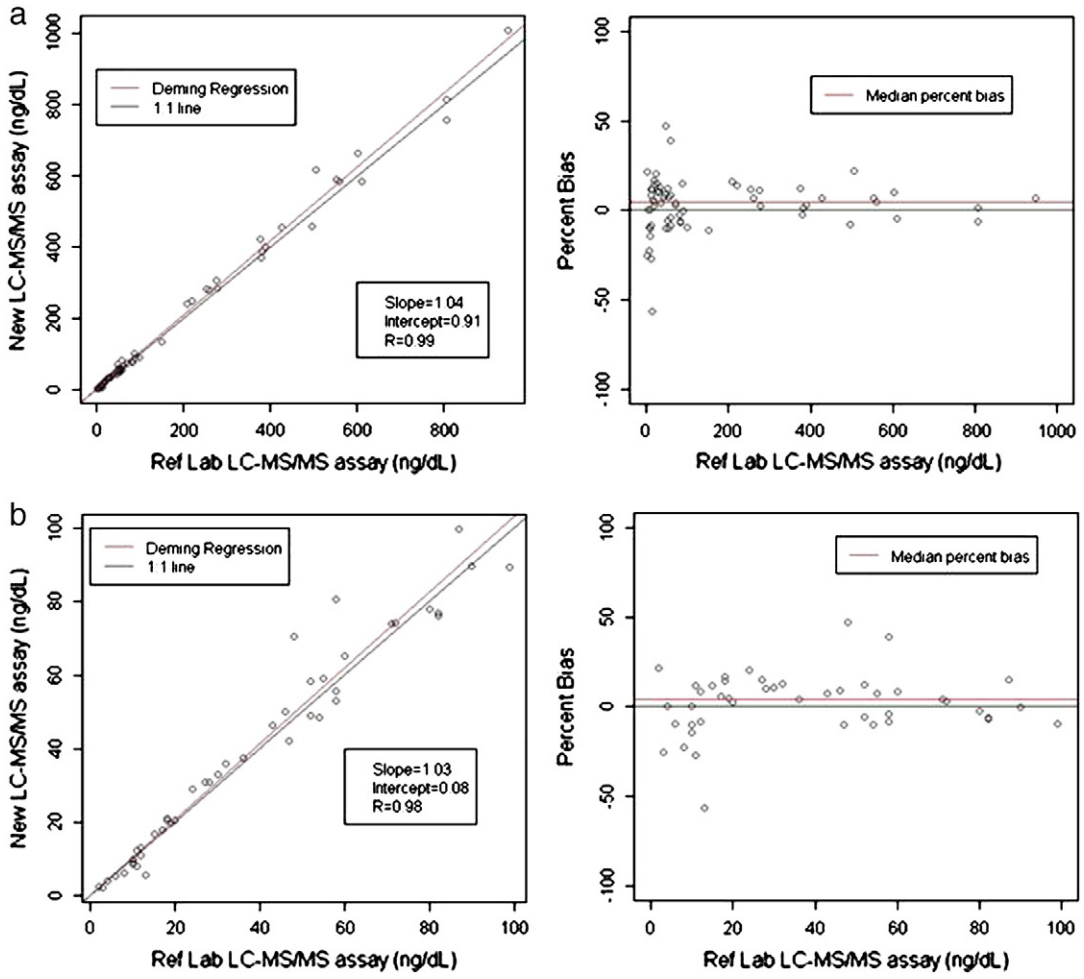

Fig. 5. Comparison between the new LC–MS/MS assay and an LC–MS/MS assay from a reference laboratory. a) Scatter plot and percent bias plot of the method comparison between the LC–MS/MS assay at a reference laboratory and the new LC–MS/MS assay using 66 patient samples (2–948 ng/dl); b) scatter plot and percent bias plot of a concentration sub-range (<100 ng/dl; n = 45) of the method comparison between the LC–MS/MS assay at a reference laboratory and the new LC–MS/MS assay.

References

[1] Kushnir MM, Blamires T, Rockwood AL, et al. Liquid chromatography–tandem mass spectrometry assay for androstenedione, dehydroepiandrosterone and testosterone with pediatric and adult reference ranges. Clin Chem 2010;56:1138-47.
[2] Barth JH, Field HP, Yasmin E, Balen AH. Defining hyperandrogenism in polycystic ovary syndrome: measurement of testosterone and androstenedione by liquid chromatography–tandem mass spectrometry and analysis by receiver operator characteristic plots. Eur J Endocrinol/Eur Fed Endocr Soc 2010;162:611-5.
[3] Elmlinger MW, Kuhnel W, Wormstall H, Doller PC. Reference intervals for testosterone, androstenedione and SHBG levels in healthy females and males from birth until old age. Clin Lab 2005;51:625-32.
[4] Surampudi PN, Wang C, Swerdloff R. Hypogonadism in the aging male diagnosis, potential benefits, and risks of testosterone replacement therapy. Int J Endocrinol 2012;2102:625434.
[5] Huhtaniemi IT, Tajar A, Lee DM, et al. Comparison of serum testosterone and estradiol measurements in 3174 European men using platform immunoassay and mass spectrometry; relevance for the diagnostics in aging men. Eur J Endocrinol 2012;166(6):983-91.
[6] Norman RJ, Dewailly D, Legro RS, Hickey TE. Polycystic ovary syndrome. Lancet 2007;370:685-97.
[7] Iughetti L, Predieri B, Ferrari M, et al. Diagnosis of central precocious puberty: endocrine assessment. J Pediatr Endocrinol Metab 2000;13(Suppl. 1):709-15.
[8] Palmert MR, Dunkel L. Delayed puberty. N Engl J Med 2012;366:443-53.
[9] Taieb J, Mathian B, Millot F, et al. Testosterone measured by 10 immunoassays and by isotope-dilution gas chromatography–mass spectrometry in sera from 116 men, women and children. Clin Chem 2003;49(8):1381-95.
[10] Wang C, Catlin DH, Demers LM, Starcevic B, Swerdloff RS. Measurement of total serum testosterone in adult men: comparison of current laboratory methods versus liquid chromatography–tandem mass spectrometry. J Clin Endocrinol Metab 2004;89(2):534-43.
[11] Rahhal SN, Fuqua JS, Lee PA. The impact of assay sensitivity in the assessment of diseases and disorders in children. Steroids 2008;73:1322-7.
[12] Rosner W, Auchus RJ, Azziz R, Sluss PM, Raff H. Utility, limitations and pitfalls in measuring testosterone: an Endocrine Society position statement. J Clin Endocrinol Metab 2007;92(2):405-13.
[13] Salameh WA, Redor-Goldman MM, Clarke NJ, Reitz RE, Caulfield MP. Validation of a total testosterone assay using high-turbulence liquid chromatography tandem mass spectrometry: total and free testosterone reference ranges. Steroids 2010;75:169-75.
[14] Savolainen K, Kiimamaa R, Halonen T. High-throughput analysis of testosterone in serum samples by on-line solid phase extraction liquid chromatography–tandem mass spectrometry. Clin Chem Lab Med 2011;49(11):1845-8.
[15] Shiraishi S, Lee PWN, Leung A, Goh VHH, Swerdloff RS, Wang C. Simultaneous measurement of serum testosterone and dihydrotestosterone by liquid chromatography–tandem mass spectrometry. Clin Chem 2008;54(11):1855-63.
[16] Cawood ML, Field HP, Ford CG, et al. Testosterone measurement by isotope-dilution liquid chromatography–tandem mass spectrometry: validation of a method for routine clinical practice. Clin Chem 2005;51(8):1472-9.
[17] Guo T, Taylor RL, Singh RJ, Soldin SJ. Simultaneous determination of 12 steroids by isotope dilution liquid chromatography–photospray ionization tandem mass spectrometry. Clin Chim Acta 2006;372:76-82.
[18] Soldin SJ, Soldin OP. Steroid hormone analysis by tandem mass spectrometry. Clin Chem 2009;55(6):1061-6.
[19] Wang C, Shiraishi S, Leung A, et al. Validation of a testosterone and dihydrotestosterone liquid chromatography tandem mass spectrometry assay: interference and comparison with established methods. Steroids 2008;73(13):1345-52.
[20] Annesley TM. Ion suppression in mass spectrometry. Clin Chem 2003;49:1041-4.
[21] Little JL, Wempe MF, Buchanan CM. Liquid chromatography–mass spectrometry/mass spectrometry method development for drug metabolism studies: examining lipid matrix ionization effects in plasma. J Chromatogr B 2006;833:219-30.

[22] Vesper HW, Bhasin S, Wang C, et al. Interlaboratory comparison study of serum total testosterone measurements performed by mass spectrometry methods. Steroids 2009;74:498-503.
[23] https://www-s.nist.gov/srmors/view_detail.cfm?srm=971; 2012. [accessed on July 25 2012].
[24] Vesper HW, Botelho JC, Shacklady C, Smith A, Myers GL. CDC project on standardizing steroid hormone measurements. Steroids 2008;73:1286-92.
[25] Matuszewski BK, Constanzer ML, Chavez-Eng CM. Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS. Anal Chem 2003;75:3019-30.
[26] Stroud LR, Solomon C, Shenassa E, et al. Long-term stability of maternal prenatal steroid hormones from the National Collaborative Perinatal Project: still valid after all these years. Psychoneuroendocrinology 2007;32:140-50.