

Quantitative determination of human serum testosterone via isotope dilution ultra-performance liquid chromatography tandem mass spectrometry

GUANG SUN^{1,2}, JINMEI XUE³, LIUXU LI¹, XUE LI¹, YAQIONG CUI¹,
BIN QIAO³, DIANJUN WEI⁴ and HUIQIANG LI¹

¹Department of Clinical Immunology, School of Medical Laboratory, Tianjin Medical University, Hexi, Tianjin 300203; ²Department of Clinical Laboratory, Tianjin Gong'an Hospital, Heping, Tianjin 300042; ³School of Chemical Engineering and Technology, Tianjin University, Jinnan, Tianjin 300350; ⁴Department of Clinical Laboratory, Hebei Yanda Hospital, Yanda International Health City, Sanhe, Hebei 065201, P.R. China

Received December 24, 2019; Accepted May 21, 2020

DOI: 10.3892/mmr.2020.11235

Abstract. Quantification of testosterone serves an important role in the differential diagnosis of androgen-related endocrine diseases. Mass spectrometry exhibits higher accuracy and lower variability than immunoassays, especially at low testosterone concentrations. The present study developed and validated an isotope dilution ultra-performance liquid chromatography tandem mass spectrometry method for determination of human serum testosterone. The serum was equilibrated with an isotopic internal standard and treated with acidic buffer to release hormones from their binding proteins. Testosterone was extracted via two serial liquid-liquid extractions. In the first stage, the lipid fractions from an acidic buffer solution were isolated using ethyl acetate and n-hexane. The organic phase was evaporated and reconstituted in a basic buffer solution. In the second stage, the polar impurities of n-hexane extraction were removed. Total testosterone in serum was quantified via ultra-performance liquid chromatography tandem mass spectrometry in multiple reaction monitoring mode with positive electrospray ionization. The coefficient of variation of the method for intra- and inter-assay was 2.13% (1.40-2.77%) and 3.44% (3.06-3.66%), respectively. The recovery ranged from 94.32 to 108.60% for different samples. The limit of detection was 0.50 ng/dl and the linear range was from 1.00 to 1,000.00 ng/dl. In addition, the extraction efficiency in three different levels of quality control of the serum ranged from 85.02 to 93.29%. Moreover, structural analogues were investigated and were not indicated to affect the quantification

of testosterone. The present method may enable quantification of testosterone in a clinical setting with high precision and accuracy.

Introduction

Testosterone is an important steroid hormone. A recent study has indicated that the harmonized reference range for total testosterone in healthy, non-obese young males aged 19-39 years was 264-916 ng/dl (1). The physiological serum testosterone concentrations are <50 ng/dl in adult females and <10 ng/dl in infants and children of both sexes (2). The quantification of testosterone levels serves an important role in the diagnosis and treatment of various diseases, such as hypogonadism, polycystic ovary and hirsutism (1,3,4).

Immunoassays have been developed to quantify testosterone levels, including radioimmunoassay, chemiluminescence analysis, electrochemiluminescence analysis, and light-initiated chemiluminescent assay (5,6). A considerable inaccuracy of testosterone assays has been reported, especially regarding testosterone quantification in females, children and male patients with hypogonadism (7-9).

Liquid chromatography tandem mass spectrometry (LC-MS/MS) combines the separation ability of liquid chromatography with the detection ability of mass spectrometry, which enables LC-MS/MS to exhibit unique advantages in the detection of small molecules (10). Moreover, isotope dilution mass spectrometry (ID-MS) has been recommended as a reference measurement method by the Advisory Committee on Quantity of Materials (11). However, to the best of our knowledge, ID-MS has been rarely used in clinical laboratories. Isotope dilution comprises the addition of a known amount of isotope-labeled analyte to each calibrator, followed by quality control and sample assessment (12). Although isotope-labeled analytes exhibit similar recovery, chromatography and ionization characteristics compared with unlabeled analytes, they can be distinguished on the mass spectrum based on their different mass numbers (12).

Correspondence to: Professor Huiqiang Li, Department of Clinical Immunology, School of Medical Laboratory, Tianjin Medical University, 1 Guangdong Road, Hexi, Tianjin 300203, P.R. China
E-mail: lhq@tmu.edu.cn

Key words: testosterone, liquid-liquid extraction, isotope dilution, ultra-performance liquid chromatography tandem mass spectrometry

Using stable isotope-labeled analytes as internal standards (IS) may counterbalance the analyte loss and matrix effect (ME) during sample pretreatment and LC-MS/MS analysis. The variability in the precision and accuracy of different mass spectrometry methods (seven high performance liquid chromatography tandem mass spectrometry assays, and one gas chromatography tandem mass spectrometry assays) has been reported, and the results of the comparison of eight mass spectrometry methods showed that the within-run variability at 297 and 8.47 ng/dl ranged between 1.40-11.36% CV and 2.52-25.58% CV, respectively (13). Sample preparation includes several stages, such as derivatization procedure, liquid-liquid extraction (LLE) and solid phase extraction, which has been indicated to result in the poor sensitivity and low extraction efficiency of the different methods (14-16). Therefore, certain mass spectrometry methods for testosterone determination cannot meet the clinical requirements. In the present study, a method for quantitative analysis of total testosterone in human serum was developed via isotope dilution (ID) ultra-performance liquid chromatography (UPLC) tandem mass spectrometry (ID-UPLC-MS/MS). The results of the current study indicated that this method was suitable for detection of testosterone in males and females, with high accuracy and precision.

Materials and methods

Samples. Clinical serum samples were collected at the Gong'an Hospital (Tianjin, China), between March 2019 and June 2019. The serum samples (n=30) that were used to assess the applicability of the present method were obtained from both outpatient and inpatient individuals. Samples with sufficient serum (at least 2 ml), clear appearance, no jaundice, no hemolysis and no blood lipids were collected. The average age of 15 female patients was 38 years (range, 20-58), while the average age of 15 male patients was 49 years (range, 18-85). A total of three levels of quality control (QC) materials were prepared via combining human serum samples from healthy donors. A total of 8 subjects were recruited from Tianjin Gong'an Hospital, between April 2019 and June 2019. Upon study entry, all participants underwent a detailed medical examination and clinical history review. The donors were enrolled in the study to meet the following requirements: Age 18-55 years; normal renal and liver function; on no medications known to interfere with androgen synthesis or action (e.g., cimetidine, spironolactone, finasteride, testosterone, flutamide, lupron); not have participated in a drug study within the last 3 months; available for blood draws between 7:00 a.m. and 10 a.m. The total testosterone of healthy donor samples was tested on the ARCHITECT i2000 analyzer (Abbott Pharmaceutical Co. Ltd.), and the results of testosterone concentration met the expected requirements. The donors were 4 male and 4 female subjects. The average age of the male donors was 31 years (range, 25-36 years), and the average age of the female donors was 28 years (range, 24-37 years). The 8 donors exhibited a normal testosterone concentration. All serum samples were stored in aliquots at -80°C until used. The study protocol was approved by the Ethics Committee of Tianjin Medical University (approval no. TMUHMEC2017008) and written informed consent was obtained from each participant prior to study entry.

Apparatus and reagents. Testosterone ($\geq 99.6\%$) was purchased from National Pharmaceutical Engineering Research Center and served as a calibration standard. 16,16,17-d₃-testosterone (T-D₃) with an isotopic purity of 99.37% was purchased from Sigma-Aldrich (Merck KGaA) and was used as an IS. Methanol, acetonitrile, n-hexane and ethyl acetate were purchased from Merck KGaA, and formic acid was purchased from Waters Corporation. Ammonium acetate and sodium carbonate were purchased from Shanghai Fuchen Chemicals Co., Ltd. All solvents were high-performance liquid chromatography grade and chemicals were reagent grade. The steroids used for interference testing were obtained from Shanghai Yuanye Bio-Technology Co., Ltd. and National Institute of Metrology of China. Electrospray ionization UPLC-MS/MS analysis was performed using a Xevo® TQ-XS Triple Quadrupole Mass Spectrometry instrument (Waters Corporation) with an ACQUITY UPLC® I-Class PLUS system (Waters Corporation).

Calibrator preparation. Testosterone was dissolved in anhydrous methanol to prepare the primary stock solution (PSS) with a concentration of 1.00 mg/ml. A total of ten calibrator levels covering a range of 1.00-1,000.00 ng/dl were prepared with methanol from 1.00 mg/ml PSS. T-D₃ was dissolved in anhydrous methanol to prepare IS solution with a concentration of 1,000.00 ng/dl.

Sample preparation. The samples for testosterone determination were prepared via LLE as previously described, with certain modifications (17). The serum samples were processed together with QC samples, reagent blank and ten levels of calibrators. Samples (100 μ l) and IS solution (100 μ l; 1,000 ng/dl) were mixed for 15 min at room temperature. Acid buffer (100 μ l; 0.5 mol/l ammonium acetate; pH 5.5) was added, and then mixed for 2 h at room temperature. LLE was performed twice with 500 μ l ethyl acetate/n-hexane solution (3:2; v/v) to extract the analytes. The combined organic extracts were evaporated to dryness using a nitrogen blowing instrument and the sample extract was re-dissolved in basic buffer (200 μ l; 0.2 mol/l sodium carbonate; pH 9.8). This solution was extracted twice using n-hexane (500 μ l each). The combined organic layers were dried and reconstituted with methanol (100 μ l) for UPLC-MS/MS analysis.

UPLC-MS/MS conditions. For the chromatographic assay, an ACQUITY UPLC™ BEH C18 (2.1x100 mm; 1.7 μ m) analytical column was heated to 40°C with 0.1% formic acid in water (buffer A) and acetonitrile (buffer B). The injection volume was 5 μ l and the samples were maintained at 5°C in the autosampler. The gradient elution procedure is presented in Table SI.

For the mass spectrometry assay, the mass spectrometry instrument was operated with electrospray ionization in the positive ion mode with ion spray source temperature at 150°C. The temperature of desolvation gas was 400°C and the nebuliser pressure of the gas was 7.0 bar (101.5 psi). The flow rate of desolvation gas was 600 l/h, and the flow rate of counter blow gas with conical hole was 150 l/h. The ion selection parameters of testosterone and T-D₃ are presented in Table I.

Table I. Ion selection parameters for T and T-D3.

Materials	Quantitative ion pairs, m/z	Confirmation ion pairs, m/z	Cone voltage, V	Collision voltage, eV
T	289.2/97.0	289.2/108.9	38	24
		289.2/97.0	38	20
T-D3	292.2/97.0	292.2/108.9	44	24
		292.2/97.0	44	24

T, testosterone; T-D3, 16,16,17-d3-testosterone; m/z, mass-to-charge ratio; eV, electron volt.

Data analysis. The UPLC-MS/MS raw data were processed using MassLynx version 4.2 software (Waters Corporation). The level of testosterone was analyzed using MedCalc version 18.2.1 (MedCalc Software Ltd.). Origin version 9.0 (OriginLab) and GraphPad prism version 6.0 (GraphPad Software, Inc.) were used for the analysis of chromatograms and histograms. The peak area ratio of testosterone quantifier ion to IS quantifier ion was used for quantification. To establish the best fit for the calibration curve, the peak area ratio of testosterone and IS was obtained from four groups of calibration runs and plotted with the target concentration using a linear model with no weighting, weight of $1/X$, $1/X^2$ or $1/\text{variance of } Y$. The average sum of squared residuals and the average relative sum of squared residuals were estimated via comparing the calculated concentration with the target concentration. A linear calibration curve using a weight of $1/X$ was selected, as this model exhibited the smallest average sum of squared residuals among all linear models. The analyte concentration in serum was calculated using the peak area ratio for the sample and the regression parameters of the established $1/X$ weighted calibration curve.

Method validation. The method verification procedure followed the guidelines of Clinical and Laboratory Standards Institute (CLSI) C62-A (18) and China National Accreditation Committee GL037 for performance evaluation (19).

Precision. The method precision was assessed using six replicates/day of in-house QC samples at three concentrations (intra-assay) and in three different days (inter-assay). A coefficient of variation (CV) $<15\%$ for the intra-assay and $<20\%$ for the inter-assay was accepted.

Accuracy. Spike recovery was evaluated via supplementing a range of testosterone with different concentrations (10.00, 200.00 and 800.00 ng/dl) into two serum samples (mixed serum from female healthy donors). The testosterone concentrations of the two serum samples were 29.56 and 57.32 ng/dl, respectively. The recovery was calculated as [(final concentration-initial concentration)/added concentration] $\times 100\%$.

Sensitivity. The limit of quantification (LOQ) was determined using five replicates as the lowest concentration, which generated a signal-to-noise ratio (S/N) >10 , with an accuracy between 80 and 120% of the true value, and CV $<20\%$. The

limit of detection (LOD) represents the absolute limit of detection that produced an S/N >3 .

Linearity. The linear range of the present method was determined using a 10-point calibration curve (testosterone, 1.00-1,000.00 ng/dl) measured in four replicates. The applicability of sample dilution was assessed for samples with analyte concentrations above the upper limit of the linear range. The sample dilution experiments were performed with high QC samples that were diluted with saline using dilution factors ranging from 1 to 10 (a total of 9-level dilutions). Each dilution level was repeated in triplicate and the mean value was used for the analysis. The dilution factors were applied to calculate the final concentrations and the results were compared with the values obtained from undiluted samples.

Specificity. Structural steroid analogues were added to the samples to assess the co-elution. The absence of a peak with mass transitions (289.2 to 97.0 for testosterone; 292.2-97.0 for T-D3) at the retention times for testosterone and IS (2.40 min) confirmed the absence of interference with the quantification of testosterone in serum. The ratio of quantification ions to confirmation ions (QI/CI) of testosterone was monitored in the samples and compared to the QI/CI ratio of calibrator solutions to confirm that no interfering compounds were present. Co-elution was considered when the QI/CI ratio differed $>20\%$ (18,20).

ME. ME was evaluated in three matrices (charcoal-stripped serum, male serum and female serum) and one set of neat samples in methanol (matrix-free), according to previous studies (21,22). A 7-point calibration curve ranging from 10-1,000 ng/dl for testosterone was prepared in each matrix. The mass spectrometry response (area count ratios of analyte over IS) was compared in all three matrices with that of the neat samples prepared in methanol. The sample ME was calculated using the following formula: $ME \% = B/A \times 100$, where B corresponded with the area count ratios of analyte to IS obtained from samples in matrix and A corresponded with the area count ratios in matrix-free samples.

Extraction efficiency. The extraction efficiency was assessed using low, medium and high QC samples. In one set of QC samples, the IS solution was added at the beginning of sample preparation (A), while in another set of QC samples it was added at the end of the sample preparation (B). The efficiency

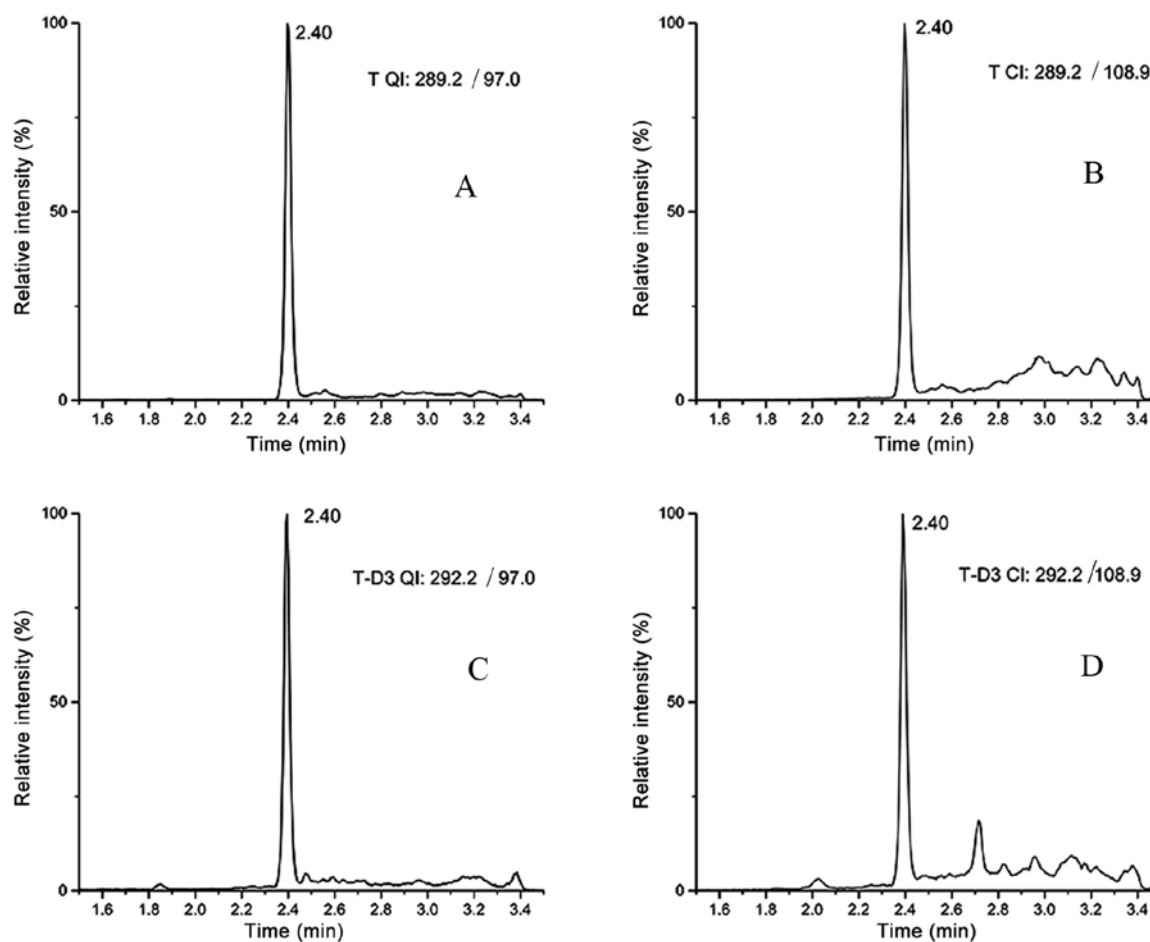


Figure 1. Quantitative and confirmation ion chromatography of T-D3 and T in serum samples. Selected ion chromatograms by UPLC-MS/MS of the (A) quantification ion and (B) confirmation ion transitions for testosterone. Selected ion chromatograms by UPLC-MS/MS of the (C) quantification ion and (D) confirmation ion transitions for 16,16,17-d3-testosterone. T, testosterone; T-D3, 16,16,17-d3-testosterone; QI, quantitative ion; CI, confirmation ion.

was calculated using the following equation: Measured concentration (B)/measured concentration (A) x100.

Stability of sample preparation. The stability of sample preparation was evaluated via comparing the measurement values of the medium QC samples that were obtained with the aforementioned method compared with those obtained following modification of three principal sample preparation parameters. The first parameter (P1) was the equilibration time, during which the sera were incubated with IS to achieve equilibration between free and protein-bound IS (incubation durations tested: Low, 10 min; medium, 15 min; high, 20 min). The second parameter (P2) was the time allowed for the removal of testosterone from proteins (incubation durations tested: Low, 90 min; medium, 120 min; high, 150 min). The third parameter (P3) was the buffer concentration used to remove testosterone from proteins (buffer concentrations tested; Low, 0.3 mol/l; medium 0.5 mol/l; high 0.7 mol/l).

Method applicability. The present method was used to analyze 30 samples from female and male patients to evaluate the applicability of the present method in the general population. The data are represented by the median and the concentration range of the test results.

Results

Chromatographic characteristics. The chromatographic retention time of testosterone and T-D3 was 2.40 min (Fig. 1), while the total analysis time of each sample was 5 min.

Precision. Precision was estimated via measuring QC samples at three concentrations in 3 days (six replicates/day). As presented in Table II, the intra-assay CV ranged from 1.40 to 2.77%, while the inter-assay CV was from 3.06 to 3.66%. The estimated intra-assay and inter-assay precision of this UPLC-MS/MS was within the acceptance criteria (23).

Accuracy. Testosterone (10.00, 200.00 and 800.00 ng/dl) was added to the serum samples. The actual concentrations, which were measured following addition of pure testosterone, were close to female and male patient levels (1,24). The recovery rates ranged from 94.32 to 108.60%, as presented in Table III, and all were within the acceptable range of 80-120% (19).

Sensitivity. LOD was evaluated via considering the lowest concentration at which the S/N was >3. LOQ was determined as the concentration providing S/N ≥ 10 within CV <20%. In the current method, LOD and LOQ were 0.50 and 1.00 ng/dl, respectively.

Table II. Intra-assay and inter-assay precision.

Group	Intra-assay, n=6			Inter-assay, n=18		
	Mean	SD	CV, %	Mean	SD	CV, %
Low	52.26	1.57	2.21	54.42	1.99	3.66
Medium	381.61	5.36	1.40	374.33	11.47	3.06
High	670.10	18.55	2.77	678.38	24.41	3.60

CV, coefficient of variation; SD, standard deviation.

Table III. Recovery of testosterone added in human serum samples.

Sample	Added, ng/dl	Measured, ng/dl	Recovery, %
1	0	29.56	- ^a
	10.00	40.10	105.40
	200.00	223.15	96.80
	800.00	810.06	97.56
2	0	57.32	- ^a
	10.00	68.18	108.60
	200.00	245.96	94.32
	800.00	822.68	95.67

^aNot applicable.

Linearity. For the evaluation of linearity, the ratio of the analyte peak area to the IS peak area was plotted against testosterone to generate the calibration curves. A calibration curve ($Y=0.0044X+0.0667$) was generated with ten concentrations of calibrators. The calibration curve was linear within the range of 1.00-1,000.00 ng/dl of testosterone ($R^2>0.999$). The weighted regression parameters from four replicate calibration curves were consistent [mean slope, 0.0044; 95% confidence interval (CI), 0.0043-0.0045; mean intercept, 0.0667; 95% CI, 0.0665-0.0669]. Dilutions of samples up to 10X with saline were indicated to result in accurate measurements (Table SII). The accuracy of diluted samples was 100.37% (95%CI: 99.30-101.44%). The high QC samples were diluted to obtain accurate results, which was to support the expansion of the measurement range. The measurement range could be extended from 1.00 to 10,000.00 ng/dl.

Specificity. Fig. 1 depicts the ion chromatograms that were observed for the quantitative and confirmation ion chromatography of T-D3 and testosterone in serum samples. The structural analogues of testosterone either did not contain the same mass transitions that were used for quantification or were chromatographically separated from testosterone with the UPLC conditions described in the current study. Structural analogues of testosterone include dihydrotestosterone, estradiol, estriol, progesterone, estrone, cortisol and corticosterone (2). Other potential interferences can be detected using the QI/CI ratio. For the present method, the QI/CI ratios in

Table IV. Assessment of ME.

Matrix	Slope	R ² -value	ME, %
Methanol (matrix-free)	0.0044	0.999	- ^a
Charcoal	0.0042	0.997	102.2
Female serum	0.0046	0.999	101.8
Male serum	0.0043	0.998	99.7

^aNot applicable. ME, matrix effect.

Table V. Extraction efficiency of testosterone in three levels of quality control.

Samples	A, ng/dl	B, ng/dl	Extraction efficiency, %
Low	60.58	56.52	93.29
Medium	382.15	324.91	85.02
High	667.70	598.64	89.66

30 single-donor serum samples were compared with the neat calibrators. The mean QI/CI ratio of the calibrators was 1.15 (95% CI, 1.10-1.21) and mean ratio of the 30 samples was 1.21 (95% CI, 1.15-1.26). The QI/CI ratios of all 30 serum samples were within $\pm 20\%$ of the mean QI/CI ratio of the calibrators, indicating that the current method was not affected by interfering compounds.

ME. The mean ME % determined in four different matrices ranged from 99.7 to 102.2% (Table IV). The slopes of all matrices and matrix-free curves were similar (~ 0.004). The results of the present study indicated that the current method was not affected by different matrices.

Extraction efficiency. The extraction efficiency of three different levels of QC ranged from 85.02 to 93.29% (Table V).

Stability of sample preparation. The medium QC samples were obtained using the aforementioned method. Following alteration in the sample preparation parameters, the measured values exhibited a good consistency (Fig. 2). For value assignment, three replicate preparations of the medium QC samples were prepared in each independent run. All data were presented as the mean \pm standard deviation. The results indicated that the current method was not affected by the alterations in the sample processing parameters.

Method applicability. The minimum value and the maximum value were 12.45 and 874.47 ng/dl, respectively. Samples with wide range of testosterone concentrations were detected via the aforementioned method (Table VI).

Discussion

In the current study, an ID-UPLC-MS/MS method was developed to detect testosterone in human serum. The method

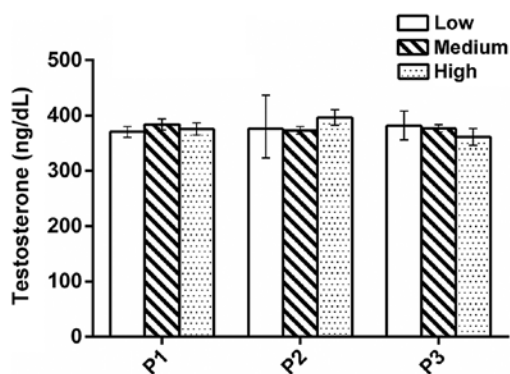


Figure 2. Stability test of sample preparation ($n=3$; error bar: Standard deviation). All data were presented as the mean \pm standard deviation and 3 repeats were included for each parameter at each duration/concentration tested.

used isotope-labeled IS, two-step LLE and UPLC to provide the desirable accuracy, specificity, precision and matrix-independent measurements. The intra-assay CV of high, medium and low QC samples were 2.77, 1.40 and 2.21%, respectively, while the inter-assay CV were 3.60, 3.06 and 3.66%, respectively. The method precision was indicated to be below the suggested maximum variation for total testosterone measurements of 5.3% (23). Moreover, the current method exhibited a good consistency between low and high concentrations of testosterone. In addition, the present method required a simple sample preparation and a small sample volume, therefore it may be suitable for routine clinical practice.

The isotope-labeled IS compensates for the potential sample loss during sample preparation. The IS exhibited the same chromatographic properties as the corresponding analytes. The sample preparation included two steps of LLE. The first LLE step was to separate lipid and protein. The second LLE step was to remove acid impurities, such as fatty acids and phospholipids. These impurities are often the source of ion inhibition and ME (25). Introducing the second LLE step resulted in detection limits and S/N ratios that were more consistent across individual samples and minimized the contamination of the UPLC-MS/MS system, thereby prolonging the column lifetime. The outstanding performance of the UPLC-MS/MS system provided a support for analysis, a sample retention time of 2.40 min and an injection volume of 5 μ l, which shortened the analysis time compared with a previous study (17) and increased the possibility of testing multiple replicates of the sample.

The extraction efficiency of the current method ranged from 85.02 to 93.29%. If an equilibrium is reached, the recovery of testosterone and the IS should be equal. Therefore, the extraction efficiency was not indicated to affect the accuracy of the method, as the quantification was based on the area ratio of testosterone and IS. However, maximizing recovery during sample preparation ensured an adequate signal strength during the UPLC-MS/MS analysis. In addition, if sufficient analytes and IS are not recovered, the extraction efficiency may affect the detection limit of the method. The detection limit of the current method was calculated to be 0.50 ng/dl.

The linear range of the present method was indicated to be 1 to 1,000 ng/dl, and the extended measurement range was 1 to 10,000 ng/dl. The current method was revealed to quantify

Table VI. Detection of testosterone in serum samples.

Samples	Age range, years	n	Median, ng/dl	Range, ng/dl
Female	20-58	15	26.04	12.45-56.80
Male	18-85	15	463.34	269.81-874.47

testosterone over a wide concentration range that covers the low concentrations observed in children, females and males. As pediatric specimens were not collected and were not assessed for the applicability of the method, additional studies are required in the future to increase the credibility.

In conclusion, a reliable ID-UPLC-MS/MS method with precision, specificity and sensitivity was developed for clinical determination of total testosterone in human serum. The measurement values that were obtained with the current method were indicated to meet the performance criteria for accuracy in clinical mass spectrometry, which are provided by CLSI (18) and the Hormone Standardization Program of the Center for Disease Control (23). The present ID-UPLC-MS/MS method may be used as a routine method in clinical laboratories to detect the total serum testosterone concentration in humans.

Acknowledgements

Not applicable.

Funding

The present study was funded by National Natural Science Foundation of China (grant no. 81772259).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

HL, BQ and DW were involved in the conception of the study and guided the experiment. GS designed and drafted the manuscript. GS and JX performed the experiments. LL, XL and YC conducted data analysis. HL and SG were involved in revising the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

The present study was approved by the Ethics Committee of Tianjin Medical University (approval no. TMUHMEC2017008) and was conducted in accordance with the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Bhasin S, Brito JP, Cunningham GR, Hayes FJ, Hodis HN, Matsumoto AM, Snyder PJ, Swerdloff RS, Wu FC and Yialamas MA: Testosterone therapy in men with hypogonadism: An endocrine society clinical practice guideline. *J Clin Endocrinol Metab* 103: 1715-1744, 2018.
2. Kushnir MM, Rockwood AL, Roberts WL, Pattison EG, Bunker AM, Fitzgerald RL and Meikle AW: Performance characteristics of a novel tandem mass spectrometry assay for serum testosterone. *Clin Chem* 52: 120-128, 2006.
3. Salonia A, Rastrelli G, Hackett G, Seminara SB, Huhtaniemi IT, Rey RA, Hellstrom WJG, Palmert MR, Corona G, Dohle GR, *et al*: Paediatric and adult-onset male hypogonadism. *Nat Rev Dis Primers* 5: 38, 2019.
4. Martin KA, Anderson RR, Chang RJ, Ehrmann DA, Lobo RA, Murad MH, Pugeat MM and Rosenfield RL: Evaluation and treatment of hirsutism in premenopausal women: An endocrine society clinical practice guideline. *J Clin Endocrinol Metab* 103: 1233-1257, 2018.
5. La'ulu SL, Kalp KJ and Straseski JA: How low can you go? Analytical performance of five automated testosterone immunoassays. *Clin Biochem* 58: 64-71, 2018.
6. Cui Y, She T, Zhao H, Li J, Li L, Gao W and Li H: Competitive light-initiated chemiluminescent assay: Using 5- α -dihydrotestosterone-BSA as competitive antigen for quantitation of total testosterone in human sera. *Anal Bioanal Chem* 411: 745-754, 2019.
7. Taieb J, Mathian B, Millot F, Patricot MC, Mathieu E, Queyrel N, Lacroix I, Somma-Delpero C and Boudou P: Testosterone measured by 10 immunoassays and by isotope-dilution gas chromatography-mass spectrometry in sera from 116 men, women, and children. *Clin Chem* 49: 1381-1395, 2003.
8. Rosner W, Auchus RJ, Azziz R, Sluss PM and Raff H: Position statement: Utility, limitations, and pitfalls in measuring testosterone: An Endocrine Society position statement. *J Clin Endocrinol Metab* 92: 405-413, 2007.
9. Wang C, Catlin DH, Demers LM, Starcevic B and 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* 89: 534-543, 2004.
10. Kushnir MM, Rockwood AL and Bergquist J: Liquid chromatography-tandem mass spectrometry applications in endocrinology. *Mass Spectrom Rev* 29: 480-502, 2010.
11. International Organization for Standardization: In vitro diagnostic medical devices-measurement of quantities in samples of biological origin-metrological traceability of values assigned to calibrators and control Materials. ISO 17511. Geneva, ISO, 2003.
12. Stanczyk FZ and Clarke NJ: Advantages and challenges of mass spectrometry assays for steroid hormones. *J Steroid Biochem Mol Biol* 121: 491-495, 2010.
13. Vesper HW, Bhasin S, Wang C, Tai SS, Dodge LA, Singh RJ, Nelson J, Ohorodnik S, Clarke NJ, Salameh WA, *et al*: Interlaboratory comparison study of serum total testosterone [corrected] measurements performed by mass spectrometry methods. *Steroids* 74: 498-503, 2009.
14. Star-Weinstock M, Williamson BL, Dey S, Pillai S and Purkayastha S: LC-ESI-MS/MS analysis of testosterone at sub-picogram levels using a novel derivatization reagent. *Anal Chem* 84: 9310-9317, 2012.
15. Keski-Rahkonen P, Huhtinen K, Poutanen M and Auriola S: Fast and sensitive liquid chromatography-mass spectrometry assay for seven androgenic and progestagenic steroids in human serum. *J Steroid Biochem Mol Biol* 127: 396-404, 2011.
16. Yuan TF, Le J, Cui Y, Peng R, Wang ST and Li Y: An LC-MS/MS analysis for seven sex hormones in serum. *J Pharm Biomed Anal* 162: 34-40, 2019.
17. Tai SS, Xu B, Welch MJ and Phinney KW: Development and evaluation of a candidate reference measurement procedure for the determination of testosterone in human serum using isotope dilution liquid chromatography/tandem mass spectrometry. *Anal Bioanal Chem* 388: 1087-1094, 2007.
18. Lynch KL: CLSI C62-A: A New standard for clinical mass spectrometry. *Clin Chem* 62: 24-29, 2016.
19. Guidance on the verification of quantitative measurement procedures used in the clinical chemistry. GL037. CNAS, 2019.
20. Rodriguez M and Orescan DB: Confirmation and quantitation of selected sulfonylurea, imidazolinone, and sulfonamide herbicides in surface water using electrospray LC/MS. *Anal Chem* 70: 2710-2717, 1998.
21. Botelho JC, Ribera A, Cooper HC and Vesper HW: Evaluation of an isotope dilution HPLC tandem mass spectrometry candidate reference measurement procedure for Total 17- β estradiol in human serum. *Anal Chem* 88: 11123-11129, 2016.
22. Matuszewski BK, Constanzer ML and Chavez-Eng CM: Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS. *Anal Chem* 75: 3019-3030, 2003.
23. Yun YM, Botelho JC, Chandler DW, Katayev A, Roberts WL, Stanczyk FZ, Vesper HW, Nakamoto JM, Garibaldi L, Clarke NJ and Fitzgerald RL: Performance criteria for testosterone measurements based on biological variation in adult males: Recommendations from the partnership for the accurate testing of hormones. *Clin Chem* 58: 1703-1710, 2012.
24. Legro RS, Schlaff WD, Diamond MP, Coutifaris C, Casson PR, Brzyski RG, Christman GM, Trussell JC, Krawetz SA, Snyder PJ, *et al*: Total testosterone assays in women with polycystic ovary syndrome: Precision and correlation with hirsutism. *J Clin Endocrinol Metab* 95: 5305-5313, 2010.
25. Neville D, Houghton R and Garrett S: Efficacy of plasma phospholipid removal during sample preparation and subsequent retention under typical UHPLC conditions. *Bioanalysis* 4: 795-807, 2012.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.