



# Clinical Usefulness of Ultraperformance Liquid Chromatography-Tandem Mass Spectrometry Method for Low Serum Testosterone Measurement

Sung-Eun Cho , M.D.<sup>1</sup>, Jungsun Han , M.S.<sup>1</sup>, Ju-Hee Park , B.S.<sup>1</sup>, Euna Park , M.T.<sup>1</sup>, Geun Young Kim , M.T.<sup>1</sup>, Jun Hyung Lee , M.D.<sup>1</sup>, Ahram Yi , M.D.<sup>1</sup>, Sang Gon Lee , M.D.<sup>1</sup>, Eun Hee Lee , M.D.<sup>1</sup>, and Yeo-Min Yun , M.D.<sup>2</sup>

<sup>1</sup>Department of Laboratory Medicine, GCLabs, Yongin, Korea; <sup>2</sup>Department of Laboratory Medicine, Konkuk University Medical Center, Konkuk University School of Medicine, Seoul, Korea

**Background:** Mass spectrometry methods exhibit higher accuracy and lower variability than immunoassays at low testosterone concentrations. We developed and validated an ultraperformance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) assay for quantifying serum total testosterone.

**Methods:** We used an ExionLC UPLC (Sciex, Framingham, MA, USA) system and a Sciex Triple Quad 6500+ (Sciex) MS/MS system in electrospray ionization and positive ion modes with multiple reaction monitoring transitions to evaluate precision, accuracy, linearity, lower limit of quantitation (LLOQ), carryover, ion suppression, stability, and reference intervals. For method comparison, we measured serum testosterone concentrations using this method in 40 subjects whose testosterone concentrations ranged from 0.14 to 55.48 nmol/L as determined using the Architect i2000 immunoassay (Abbott Diagnostics, Abbott Park, IL, USA) and in an additional 160 sera with testosterone concentrations < 1.67 nmol/L.

**Results:** The intra- and inter-run precision CVs were < 2.81%, and the accuracy bias values were < 3.85%, which were all acceptable. The verified linear interval was 0.03–180.84 nmol/L; the LLOQ was 0.03 nmol/L. No significant carryover and ion suppression were observed. The testosterone in serum was stable at 4°C, at –20°C, and after three freeze-thaw cycles. The reference intervals were successfully verified. The correlation was good at testosterone concentrations of 0.14–55.48 nmol/L; however, the Architect assay showed positive percent bias at concentrations < 1.67 nmol/L.

**Conclusions:** The UPLC-MS/MS assay shows acceptable performance, with a lower LLOQ than the immunoassay. This method will enable the quantitation of low testosterone concentrations.

**Key Words:** Tandem mass spectrometry, Testosterone, Quantitation, Performance

**Received:** November 13, 2021

**Revision received:** May 27, 2022

**Accepted:** August 3, 2022

**Corresponding author:** Sung-Eun Cho, M.D.  
Department of Laboratory Medicine,  
GCLabs, 107 Ihyeon-ro 30 Beon-gil,  
Giheung-gu, Yongin 16924, Korea  
Tel: +82-31-260-0947  
Fax: +82-31-260-9049  
E-mail: secho0824@gmail.com



© Korean Society for Laboratory Medicine

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## INTRODUCTION

Total testosterone is an important biomarker of various testosterone-related endocrine disorders, and a reliable method for measuring total testosterone is essential [1]. Immunoassays are available for measuring total testosterone on automated immuno-

chemistry platforms. However, they show significant interferences with positive bias at low concentrations, which are commonly observed in testosterone-deficient men, women, and children. Moreover, there is a poor agreement among the different immunoassays. Mass spectrometry (MS) methods exhibit higher accuracy and lower variability than immunoassays, espe-

cially at low analyte concentrations. The very low concentrations of total testosterone in certain populations necessitate the use of MS-based methods. Accurate measurements in patients with low testosterone concentrations have led to position and consensus statements from the Endocrine Society, along with efforts and guidelines to evaluate and improve testosterone testing [2-5].

A previous study evaluated selective performance characteristics of the following five automated immunoassays for total testosterone: the Architect 2nd Generation Testosterone assay on the Architect i2000sr system (Abbott Diagnostics, Abbott Park, IL, USA), the Testosterone (TSTO) assay on the ADVIA Centaur system (Siemens, Malvern, PA, USA), the Testosterone assay on the Beckman Coulter UniCel Dxl 800 system (Beckman Coulter, Brea, CA, USA), the Testosterone II assay on the Roche Modular E170 system (Roche Diagnostics, Indianapolis, IN, USA), and the Total Testosterone assay on the Immulite 2000 XPi system (SiemensTheSiemens) [6]. According to it, some methods showed acceptable performances by improving upon existing testosterone assays, but, for the concentrations less than 1.9 nmol/L ( $1.9 \times 0.2884 \div 0.55$  ng/mL), which is the upper reference limit for women by its LC-MS/MS, all methods had positive biases, ranging from 14.2 to 63.8%.

In order to improve consistency of different methods, Centers for Disease Control and Prevention (CDC) established the Hormone Standardization program (HoSt) for testosterone assays. Standardization of total testosterone measurements in serum would be established through method comparison and bias estimation between the CDC reference laboratory and the testing laboratory [7].

The CDC HoSt program consists of two phases. In phase 1, 40 samples are sent to the participant laboratory, with total testosterone concentrations assigned. The participating laboratory can use these 40 samples to perform a bias assessment and adjust its calibration as needed prior to the start of phase 2. In phase 2, the laboratory will receive 4 sets of 10 samples with unknown concentrations over the course of 12 months. The acceptable overall mean bias criterion of  $\pm 6.4\%$ , based on biological variability data, will be used to issue certification. Twenty assays are certified by the CDC HoSt program (updated on December 2021). Among these, 16 are LC-MS/MS assays and four are chemiluminescence immunoassays (CLIAs) from Siemens [8]. This data indirectly indicates that standardization of serum testosterone measurement is to some extent achieved among LC-MS/MS methods, but not achieved among immunoassays, yet.

We developed and validated an ultra-performance LC-MS/MS

(UPLC-MS/MS) assay for quantifying low total testosterone concentrations. We compared the results with immunoassay measurement results. Finally, we participated in the an Accuracy-Based Survey (ABS) of the College of American Pathologists (CAP, Northfield, IL, USA) and the CDC HoSt program to demonstrate the accuracy of our LC-MS/MS method.

## MATERIALS AND METHODS

This method was developed and evaluated in the Special Chemistry Department of GCLabs (Yongin-Si, Gyeonggi-do, Korea) from November, 2020 to August 2021. Experiments involving human subjects were performed according to the Declaration of Helsinki (2019) and approved by the Institutional Review Board (IRB) of GCLabs (IRB No. GCL-2021-1050-01).

### Reagents and chemicals

HPLC-grade methanol and water (Tedia, Fairfield, OH, USA), HPLC-grade acetonitrile (Fisher, Pittsburgh, PA, USA), formic acid (Fluka, Muskegon, MI, USA), and tert-butyl methyl ether (MTBE) (Acros, Fair Lawn, NJ, USA) were used. Calibrators (6Plus1 Multilevel Serum Calibrator Set MassChrom Steroid Panel 2) and quality control (QC) materials (MassCheck Steroid Panel 2 Serum Control Levels I, II, and III) were purchased from Chromsystems (Gräfelfing, Bayern, Germany). Certified reference material (CRM) testosterone (1 mg/mL [3.47 nmol/L] in acetonitrile, m/w 288.42 g/mol) was purchased from Cerilliant (Round Rock, TX, USA). The internal standard (IS; testosterone-2,3,4- $^{13}\text{C}_3$ , 0.1 mg/mL [0.35 nmol/L] in methanol, m/w 291.40 g/mol) was purchased from Sigma-Aldrich (St. Louis, MO, USA). The standard reference material (Hormones in Frozen Human Serum, Male Serum and Female Serum, SRM 971a) was purchased from the National Institute of Standards and Technology (NIST) (Gaithersburg, MD, USA). Isolute SLE+ 400  $\mu\text{L}$  Supported Liquid Extraction (SLE) plates were purchased from Biotage (Uppsala, Uppsala, Sweden).

### Calibrator, QC material, and IS preparation

The calibrators were reconstituted by adding 3.0 mL of distilled water (DW) to the vials, and calibration curves were constructed using seven concentrations of each calibrator (0, 0.05, 0.25, 0.96, 2.94, 5.78, and 11.50 ng/mL [0, 0.16, 0.86, 3.33, 10.19, 20.04, and 39.88 nmol/L, respectively]). Three QC samples of low, medium, and high concentrations (0.20, 1.46, and 7.92 ng/mL [0.69, 5.06, and 27.46 nmol/L, respectively]) were prepared by adding 3.0 mL of DW to the vials. NIST SRM 971a ma-

terials stored at  $-80^{\circ}\text{C}$  were thawed and used as internal QC materials. The IS at 0.1 mg/mL (0.35 nmol/L) in methanol was diluted with acetonitrile to yield a 100-mL working solution with a final concentration of 50 ng/mL (173.37 nmol/L).

### Sample preparation

Twenty microliters of IS working solution was transferred into each well of a 96-well plate. Then, 200  $\mu\text{L}$  of calibrators, controls, NIST materials, and sera from study subjects were added into each well. The solution in each well was mixed by pipetting up and down 10 times using a 12-channel pipette. A SLE plate was placed on a deep-well plate to collect the extraction solvent. The mixed samples were loaded onto the SLE plate and drawn into the plate by applying vacuum using a vacuum pump. Then, 800  $\mu\text{L}$  of MTBE solvent was added for extraction under vacuum. The solvent was evaporated to dryness and reconstituted in 100  $\mu\text{L}$  of 50% methanol by vortexing. Twenty microliters of the reconstituted eluate was injected into the LC-MS/MS instrument.

### LC-MS/MS

The ExionLC UPLC (Sciex, Framingham, MA, CA, USA) system equipped with a Kinetex C18 column (3.0 $\times$ 100 mm, 2.6  $\mu\text{m}$ ; Phenomenex, Torrance, CA, USA) was used for LC. The mobile phases were 0.1% formic acid in DW (A) and 0.1% formic acid in acetonitrile (B). The following gradients were applied to the column: 0 minute, 60% A and 40% B; 0.5 minute, 60% A and 40% B; 4 minutes, 5% A and 95% B, 4.5 minutes, 5% A and 95% B; 4.6 minutes, 60% A and 40% B; and 7 minutes, 60% A and 40% B. The total run time was 7 minutes; the flow rate was 0.30 mL/min.

We used the Triple Quad 6500+ (Sciex) MS/MS system equipped with an electrospray ionization source operated in positive ion mode. Nitrogen gas was used for nebulation, desolvation, and collision. The analytes were monitored in the multiple reaction monitoring (MRM) mode. The transitions and other MS/MS settings are shown in Table 1. Quantitation was performed based on the ratio of the integrated peak area of testosterone to that of

the IS and was calculated using Analyst Instrument Control and Data Processing Software (version 1.7.1 with HotFix1, AB Sciex).

### Assay performance analysis

Analytical performance characteristics, including precision, accuracy, linearity, lower limit of quantitation (LLOQ), carryover, matrix effects, and stability, were evaluated, and method comparisons were conducted in accordance with the US Food and Drug Administration Center for Drug Evaluation and Research (CDER) Bioanalytical Method Validation Guidance for Industry [9], CLSI guidelines [10, 11], a previous study [12], and review articles on LC-MS/MS laboratory development and operation [13, 14].

Intra-run precision was assessed using five replicates in a single run for five days, and inter-run precision was assessed using 20 separate runs over 20 days, with a single run per day and five analyte concentrations. These five concentrations included three of QC materials and two of NIST SRM 971a Hormones in Frozen Human Serum: 0.32 ng/mL (1.12 nmol/L) with an uncertainty limit of  $\pm 0.005$  ng/mL (0.01 nmol/L) for the Female Serum and 5.81 ng/mL (20.14 nmol/L) with an uncertainty limit of  $\pm 0.090$  ng/mL (0.31 nmol/L) for the Male Serum. Accuracy was assessed using same NIST SRM 971a materials. Acceptance limits were within  $\pm 5.3\%$  CV for precision and within  $\pm 6.4\%$  bias of nominal concentrations for accuracy. Analytical performance goals for testosterone measurement on the basis of biological variability were set according to the CDC HoSt program participant protocol acceptance criteria [8] quoted from the study of Yun, *et al.*'s study [15]. The linearity of the response was assessed by mixing a 1/800 dilution of NIST SRM 971a Male Serum as the low-concentration material and 50 ng/mL (173.37 nmol/L) CRM as the high-concentration material to achieve final concentrations of 0.007, 0.02, 0.03, 0.06, 12.54, 25.03, 37.52, and 50.00 ng/mL (0.02, 0.05, 0.10, 0.20, 43.50, 86.79, 130.09, and 173.37 nmol/L, respectively). The LLOQ was evaluated using NIST SRM 971a Male Serum diluted with DW, with  $<20\%$  CV (eight concentrations $\times$ five times). Carryover was evaluated according to the following equation, with four se-

**Table 1.** UPLC-MS/MS settings, including MRM transitions

Analyte	Q1 (m/z)	Q3 (m/z)	DP (V)	EP (V)	CXP (V)	CE (V)	CUR (Psi)	CAD (Psi)	TEM ( $^{\circ}\text{C}$ )	Dwell time (msec)
Testosterone 1 (quantifier)	289.1	97.1	50	11	5	30	45	9	450	100
Testosterone 2 (qualifier)	289.1	109.1	75	10	10	35	45	9	450	100
Testosterone $-2,3,4-^{13}\text{C}_3$ (IS)	292.2	100.0	75	10	5	30	45	9	450	100

Abbreviations: UPLC-MS/MS, ultraperformance liquid chromatography-tandem mass spectrometry; MRM, multiple reaction monitoring; IS, internal standard; DP, declustering potential; EP, exit potential; CXP, collision cell exit potential; CE, collision energy; CUR, curtain gas; CAD, collision gas; TEM, temperature.

rial measurements of two levels of calibrators (calibrators 1 and 6):

$$\text{Carryover (\%)} = \frac{[L1 - (L3 + L4)/2]}{[(H2 + H3)/2 - (L3 + L4)/2]} \times 100$$

Where, L indicates the low concentration, and H the high concentration. The acceptance criterion for carryover was  $\pm 1.0\%$ .

Ion suppression was evaluated by the post-column infusion method. In brief, a standard testosterone solution with a concentration of 200 ng/mL (693.48 nmol/L) was continuously infused directly into the mass detector at a flow rate of 7  $\mu\text{L}/\text{min}$ , while DW and six extracted participant samples were injected into the UPLC column at a flow rate of 0.2 mL/min. If there was a significant change in the detection level, ion suppression was considered to have occurred at the point at which the change was observed.

To evaluate the stability of testosterone concentrations in serum, 10 serum samples stored at  $-20^\circ\text{C}$  were exposed to one to three repeated freeze-thaw cycles, and testosterone concentrations were measured after each cycle. In addition, 10 serum samples were stored at  $4^\circ\text{C}$  and  $-20^\circ\text{C}$ , and testosterone concentrations were measured at 3, 7, 14, and 28 days and at 3, 7, 14, 28, and 60 days, respectively. The acceptance criterion was that the accuracy (% nominal) at each level should be  $\pm 15\%$ .

### Verification of reference intervals (RIs)

After comparing the lower limits of all reference values from Mayo Clinic Laboratories, Quest Diagnostics, and ARUP Laboratories based on LC-MS/MS and the in-kit reference values of a chemiluminescent microparticle immunoassay (CMIA) and electrochemiluminescent immunoassay (ECLIA) (Supplemental Data Table S1), we transferred the lowest reference value limits among them. For verification of the transferred RIs, we measured the testosterone concentrations in residual samples from 30 healthy men and 30 healthy women aged 20–49 years and 30 healthy men and 30 healthy women older than 50 years. The results were statistically analyzed and considered appropriate if 27 out of 30 (in the 95% confidence interval) measurements were within the desired RI according to CLSI guidelines [16].

### Method comparison of LC-MS/MS and CMIA

We measured serum testosterone concentrations in 40 random subjects having testosterone concentrations of 0.04–16.00 ng/mL (0.14–55.48 nmol/L) by LC-MS/MS and CMIA using the Architect 2nd Generation Testosterone assay on the Architect i2000sr system. Additionally, we measured serum testosterone concentrations in 160 subjects (40 men, 40 women, 40 boys [ $<20$  years], and 40 girls [ $<20$  years]) with unknown clinical histories, hav-

ing testosterone concentrations  $<0.48$  ng/mL (1.67 nmol/L), which is the upper reference limit for women  $<50$  years by LC-MS/MS.

### Statistical analysis

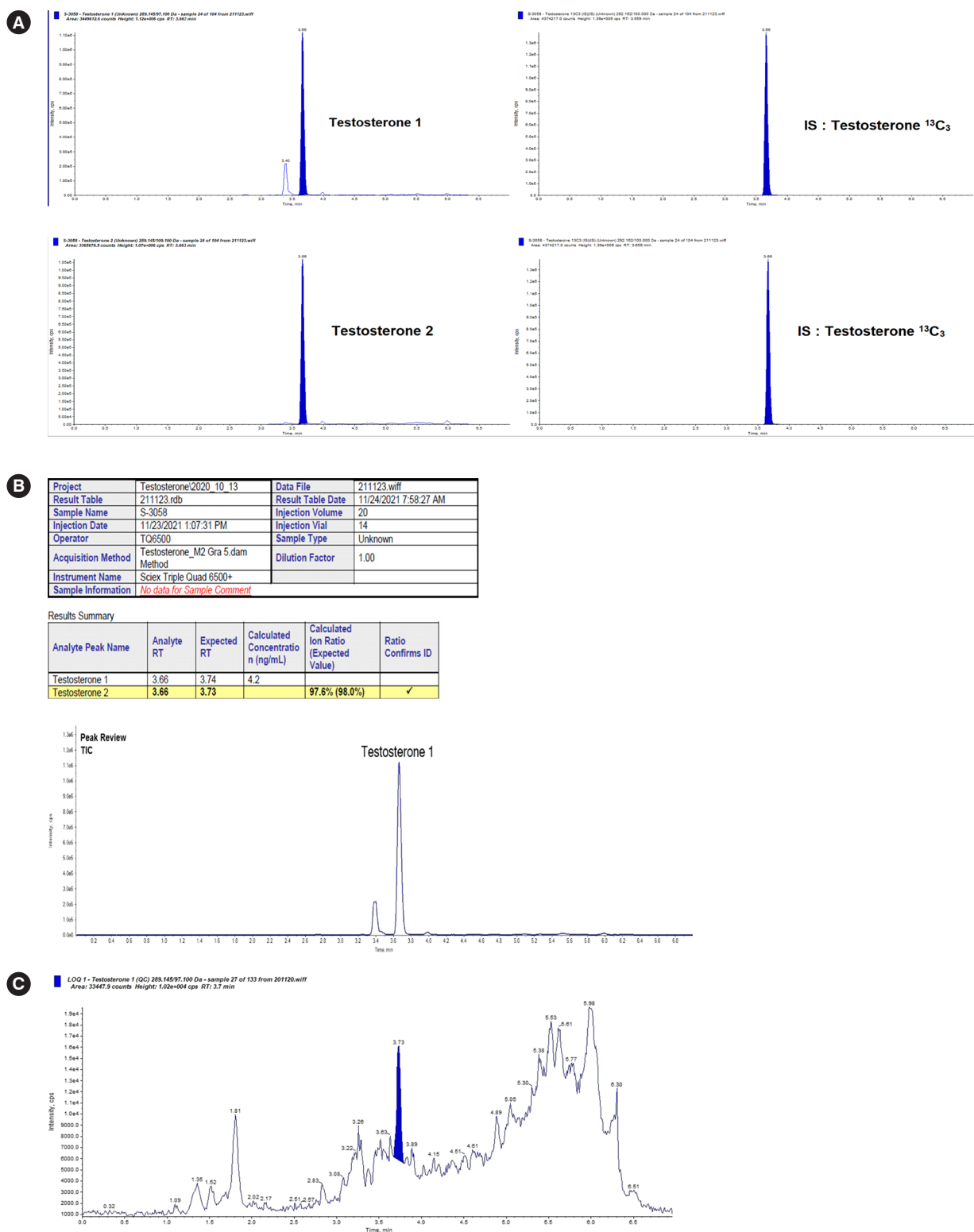
EP Evaluator (Data Innovations, Burlington, VT, USA) was used for validation of the selected RIs and for method comparisons using Passing–Bablok regression analysis and percent bias plots. *P*-values  $<0.05$  were considered significant.

## RESULTS

Representative UPLC-MS/MS chromatograms of serum total testosterone are shown in Fig. 1. The inter- and intra-run imprecision ranged from 0.48% to 2.81% CV, which is lower than the minimal requirement based on biological variation. The percentage bias for accuracy ranged from 1.55% to 3.85% and was within the acceptable range (Table 2). The linearity range was 0.008–52.16 ng/mL (0.03–180.84 nmol/L), with  $R^2=0.9999$ . The LLOQ was 0.008 ng/mL (0.03 nmol/L). The % bias and CV (%) calculated at eight concentrations using SRM 971a, Male Serum were as follows: measured mean concentrations (ng/mL) (% bias, CV (%)); 0.008 ng/mL (0.03 nmol/L) (6.85, 10.73), 0.02 ng/mL (0.06 nmol/L) (8.97, 11.32), 0.03 ng/mL (0.10 nmol/L) (4.83, 8.58), 0.06 ng/mL (0.21 nmol/L) (2.62, 0.92), 12.58 ng/mL (43.62 nmol/L) (0.30, 1.25), 26.18 ng/mL (90.77 nmol/L) (4.60, 1.28), 39.19 ng/mL (135.88 nmol/L) (4.45, 2.52), and 52.16 ng/mL (180.84 nmol/L) (4.31, 2.38). Although the diluted sample does not represent the authentic matrix, the absence of ionization suppression supports the determination of the LLOQ. There was no carryover effect. No significant ion suppression or enhancement was observed at the corresponding retention time (Supplemental Data Fig. S1). Testosterone in serum was stable over three freeze-thaw cycles, for 28 days at  $4^\circ\text{C}$ , and for 60 days at  $-20^\circ\text{C}$ .

After validation of the selected RIs, we determined them to be as follows: for men aged 20–49 years, 2.49–8.36 ng/mL (8.63–28.99 nmol/L), for men older than 50 years, 1.93–7.40 ng/mL (6.69–25.66 nmol/L), for women of 20–49 years, 0.08–0.48 ng/mL (0.29–1.67 nmol/L), and for women older than 50 years, 0.03–0.41 ng/mL (0.10–1.41 nmol/L).

The results of the method comparison of UPLC-MS/MS and the Architect CMIA are shown in Fig. 2 and Supplemental Data Table S2. The correlation was good ( $R=0.989$ , slope=0.995) in 40 random subjects having total testosterone concentrations of 0.04–16.00 ng/mL (0.14–55.48 nmol/L); however, the CMIA



**Fig. 1.** Representative MS/MS chromatograms. (A) Representative MS/MS chromatograms of testosterone in serum samples from study subjects (4.25 ng/mL [14.73 nmol/L]). (B) The calculated ion ratio of 97.6% was close to the expected value of 98.0% for the first and second MRM transitions, implying adequate identification and specificity of the assay. (C) Representative MS/MS chromatogram of testosterone in a 1/800 dilution of NIST SRM 971a, Male Serum (LLOQ: 0.008 ng/mL [0.03 nmol/L]). Abbreviations: MS/MS, tandem mass spectrometry; MRM, multiple reaction monitoring; LLOQ, lower limit of quantitation; RT, retention time.



**Table 2.** Assay performance evaluation results, including intra- and inter-run precision and accuracy

Analyte	Target value* (ng/mL <sup>†</sup> )	Precision				Accuracy
		Intra-run (N = 5)		Inter-run (N = 20)		
		Mean (ng/mL)	% CV	Mean (ng/mL)	% CV	% Bias
QC 1	0.20 <sup>†</sup>	0.20	0.57	0.20	2.81	
QC 2	1.46 <sup>†</sup>	1.45	1.06	1.47	2.15	
QC 3	7.92 <sup>†</sup>	7.87	0.62	7.78	1.50	
SRM 971a, Female Serum	0.32	0.33	0.48	0.33	1.82	1.55–2.18
SRM 971a, Male Serum	5.81	6.03	0.86	5.96	1.17	2.55 – 3.85

\*Target values of QC1, QC2, and QC3 are mean values reported in the Chromsystems insert sheet; acceptable ranges for the QC levels are as follows: 0.14–0.26 for QC1, 1.17–1.75 for QC2, and 6.34–9.51 for QC3, according to the insert sheet; <sup>†</sup>1 ng/mL  $\times$  3.47  $\div$  3.47 nmol/L. Abbreviation: QC, quality control.

showed positive percent bias compared with LC-MS/MS in men, women, boys, and girls having testosterone concentrations <0.48 ng/mL (1.67 nmol/L), i.e., 20.36% in all (N=160), 64.46% in men, 7.04% in women, 29.99% in boys, and 15.76% in girls.

Our results of the participation in the external proficiency tests of CAP ABS-B and the CDC HoSt program were all acceptable.

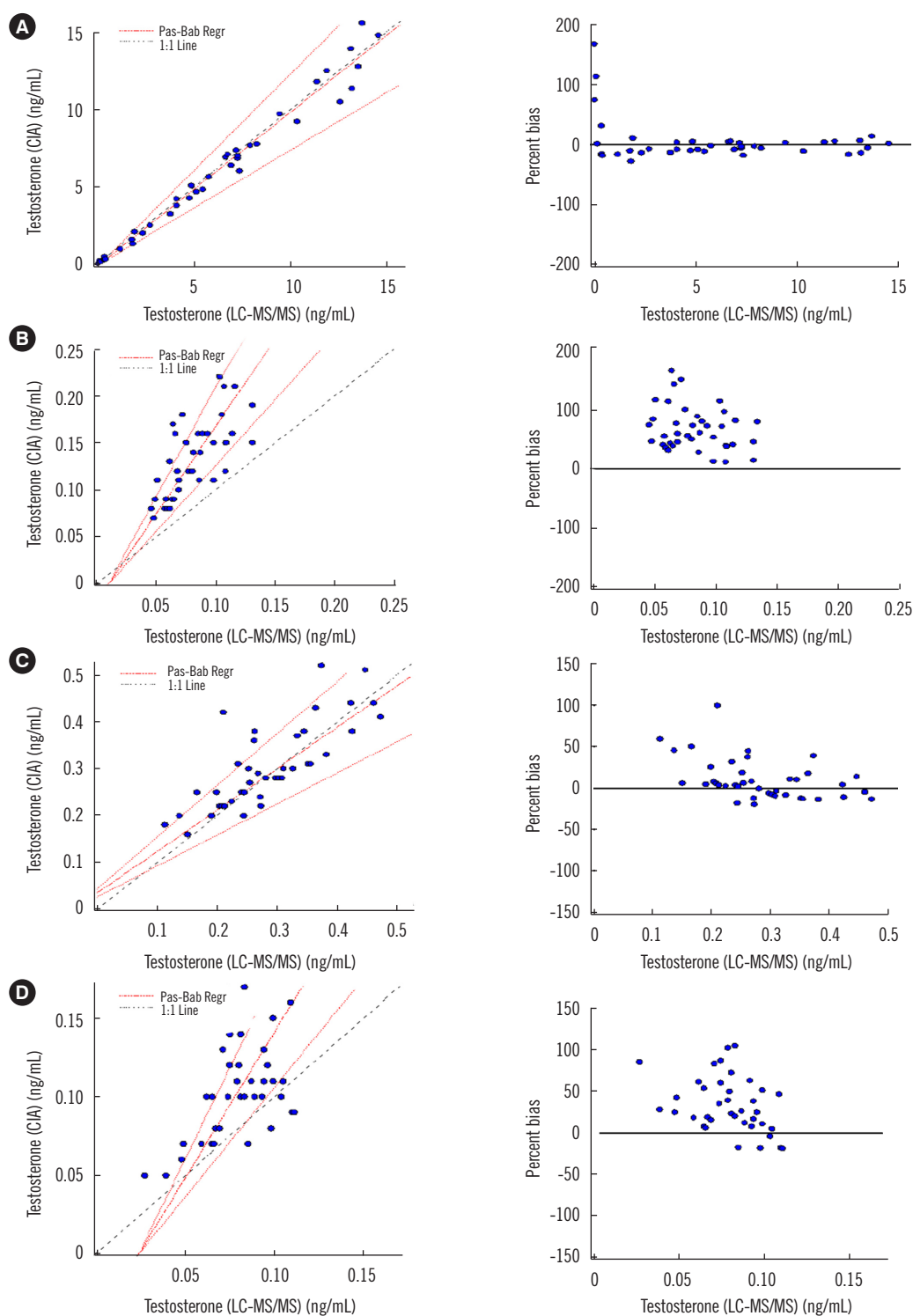
## DISCUSSION

We developed and validated a UPLC-MS/MS method for quantifying low total testosterone concentrations in serum. All assay performance characteristics were satisfactory. Most importantly, the LC-MS/MS method had lower LLOQ values and a wider linearity range than conventional immunoassays. The LLOQ of the LC-MS/MS method was 0.008 ng/mL (0.03 nmol/L), which is three times lower than that of 2nd Generation Testosterone assay on the Architect i2000sr system, which is 0.02 ng/mL (0.08 nmol/L) according to the manufacturer. The linearity range of our method was 0.008–52.15 ng/mL (0.03–180.84 nmol/L)—three times wider than that of the Architect CMIA, which is 0.04–18.62 ng/mL (0.13–64.56 nmol/L) according to the manufacturer [17]. In fact, one male participant had a testosterone concentration of 0.02 ng/mL (0.07 nmol/L) as measured by LC-MS/MS, which was not detected by the CMIA. Furthermore, the Architect CMIA showed positive percent bias compared to our method, which is consistent with a previous finding of 19.2% positive bias for the Architect CMIA [6].

In a study by Moal, *et al.* [18], none of the five immunoassays tested demonstrated sufficiently reliable results at testosterone concentrations <1.00 ng/mL (3.47 nmol/L), whereas LC-MS/MS precisely measured low testosterone concentrations. Kushnir, *et al.* [19] developed an LC-MS/MS method for measuring testosterone in women and children. The LLOQ of their method

was 0.04 nmol/L (0.01 ng/mL). The authors suggested that the sensitivity and specificity of their method were adequate for the analysis of testosterone in samples from women and children. The LLOQ of our method is even lower and thus adequate for the analysis of testosterone in samples from women and children with low testosterone concentrations. Testosterone concentrations measured by CLIA on the Immulite 2000 system (Siemens) tended to be higher than those obtained by LC-MS/MS [20]. However, two samples with low testosterone concentrations among 35 samples were not detected by CLIA but were detected by LC-MS/MS. This is in line with our findings.

In an inter-laboratory comparison study of serum total testosterone measurements using LC-MS/MS, the variability of total testosterone results among MS assays was substantially lower than that reported for immunoassays [21]. Our results of the participation in external proficiency tests were all acceptable. In the CAP Y-A 2021 Ligand-Special program, results are considered acceptable when they are within the peer group mean  $\pm$  3 SD, which is the allowable limit according to the CAP [22]. In the CAP ABS-B 2021, 22 of the 66 participating laboratories measured testosterone concentrations using MS (44 laboratories used EIA); however, as this was an educational challenge program, CAP did not report our grade [23]. Our results were all acceptable as evaluated by us, according to a bias goal of 6.4%. Accuracy-based proficiency testing can significantly contribute to improving testosterone testing by providing reliable data on accuracy in patient care to laboratories, assay manufacturers, and standardization programs [24]. MS methods in the 22 above-mentioned laboratories showed accurate median concentrations with narrow ranges, whereas the EIAs showed variable median concentrations with wider ranges according to the manufacturers [23]. Our results of the participation in the CDC HoSt program were also satisfactory. The mean (%) bias was 4.8% in



**Fig. 2.** Method comparison of UPLC-MS/MS and the Architect CMIA. Passing-Bablok regression scatter plots and percent bias plots. The correlation was good ( $R=0.989$ , slope 0.995) in 40 random subjects having testosterone concentrations of 0.04–16.00 ng/mL (0.14–55.48 nmol/L); however, positive percent bias was observed in the percent bias plot, especially at lower concentrations (A). The Architect CMIA exhibited positive percent bias in all groups with testosterone concentrations <0.48 ng/mL in the percent bias plots: 64.46% in 40 men (B), 7.04% in 40 women (C), 29.99% in 40 boys (D), 15.76% in 40 girls (E), and 20.36% in 160 total subjects (F). (G) Results from 200 subjects (40 random subjects and 160 subjects having testosterone concentrations <0.48 ng/mL [1.67 nmol/L]).

Abbreviations: CMIA, chemiluminescent microparticle immunoassay; UPLC-MS/MS, ultraperformance liquid chromatography-tandem mass spectrometry.

(Continued to the next page)

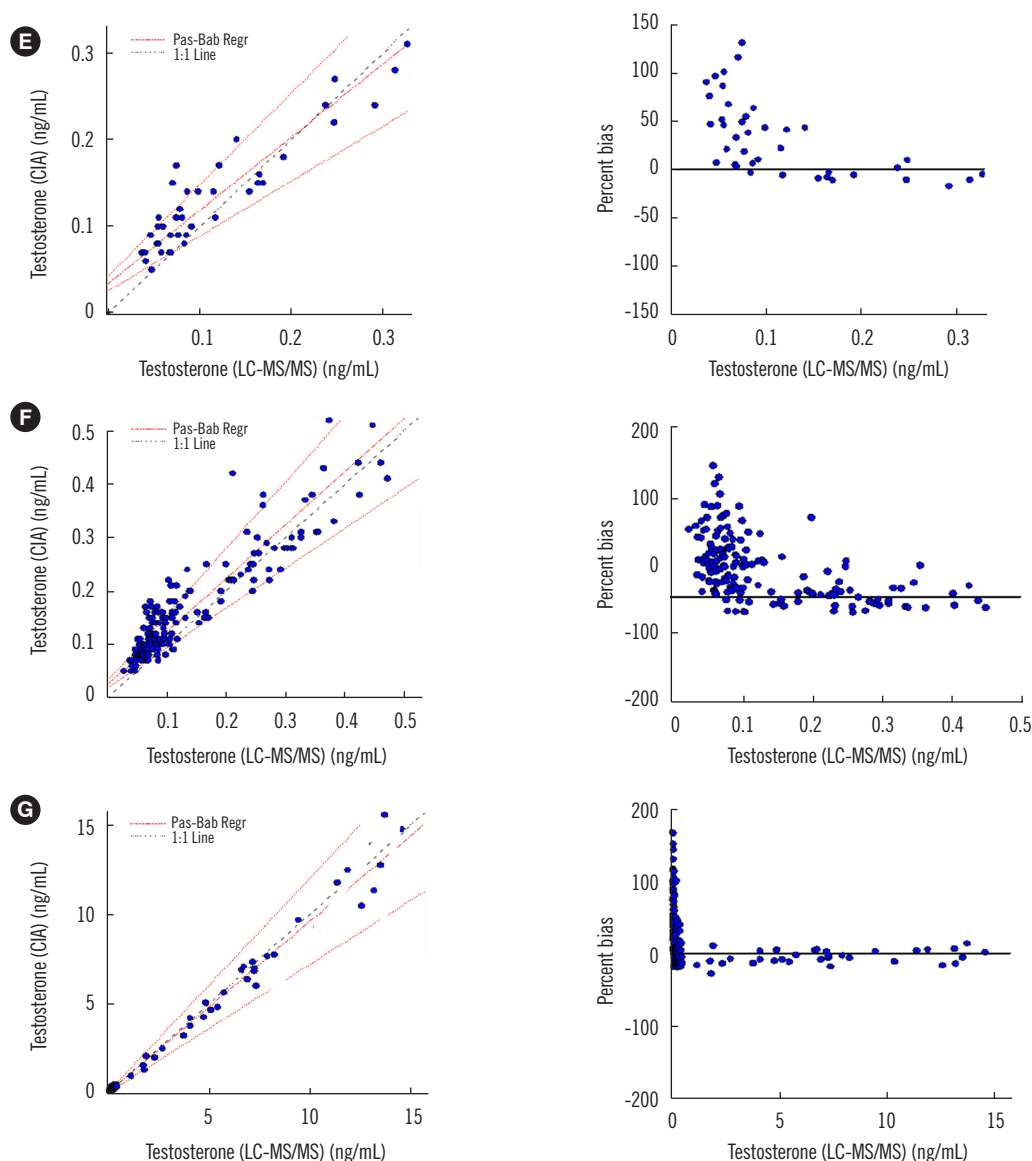


Fig. 2. Continued.

phase 1 and 3.2% in phase 2 (Q3–4 of 2021), which are all within the acceptable overall mean bias of 6.4%. The HoSt certificate will be issued for one year's performance after passing consecutive additional tests conducted in each quarter.

In future, LC-MS/MS should not only be applied as a routine primary measurement method for low testosterone concentrations but also as a reference measurement method in various national projects in which the accuracy of the results is of utmost importance. We are currently participating in the Biomarker Panel Data Production program supervised by National Biobank of Korea, in which serum testosterone concentrations from healthy men and women are measured using UPLC-MS/MS.

We recommend a critical medical decision point for selecting the best method to measure low testosterone concentrations between immunoassays and the LC-MS/MS method. Below the medical decision point, it would be better to measure testosterone concentrations using LC-MS/MS than using immunoassays. Based on our findings and those of other studies, the medical decision points for selecting the best testosterone measurement method as well as for showing the positive bias of various immunoassays are as follows: 0.48 ng/mL (1.67 nmol/L) (the present study), 0.55 ng/mL (1.90 nmol/L) [6], 1.00 ng/mL (3.47 nmol/L) [18], and 0.50 ng/mL (1.73 nmol/L) [19]. Thus, serum testosterone concentrations of 0.48–1.00 ng/mL (1.67–3.47 nmol/L)



can be considered medical decision points to select the best measurement method in clinical laboratories. We suggest a limit of 0.48 ng/mL (1.67 nmol/L), which is the upper limit of the RI for women of 20–49 years of age based on UPLC-MS/MS.

There are several Korean studies on LC-MS/MS-based measurement of testosterone concentrations [25–27]. Testosterone in neonates has been measured using dried blood spot multiplexed steroid profiling using LC-MS/MS [25]. Serum testosterone has been measured for monitoring chemical castration agents [26] and for serum steroid profiling in healthy children and adults [27] using LC-MS/MS. However, there are less than three institutions providing LC-MS/MS clinical services and that too in limited situations.

In conclusion, we developed a UPLC-MS/MS method for measuring low total testosterone concentrations and compared it with an immunoassay. We demonstrated its adequate performance characteristics. Method accuracy was demonstrated by the UPLC-MS/MS analysis results of SRM materials and by participating in several accuracy-based proficiency testing programs. Our method has lower LLOQ values and a wider linearity range than conventional immunoassays. It enables precisely measuring low total testosterone concentrations, especially in children and women, in Korea.

## ACKNOWLEDGEMENTS

None.

## AUTHOR CONTRIBUTIONS

Cho SE, Yun YM, Lee JH, Yi A, Han J, Park E, and Kim GY conceived and designed this study. Park JH and Han J investigated and validated this study. Cho S-E wrote this manuscript. Yun YM, Lee JH, Yi A, Par JH, and Han J provided valuable feedback. Cho SE, Lee SG, and Lee EH supervised this study. All authors have read and approved the final manuscript.

## CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

## RESEARCH FUNDING

None declared.

## ORCID

Jungsun Han	<a href="https://orcid.org/0000-0003-1160-0776">https://orcid.org/0000-0003-1160-0776</a>
Geun Young Kim	<a href="https://orcid.org/0000-0002-5867-5655">https://orcid.org/0000-0002-5867-5655</a>
Eun Hee Lee	<a href="https://orcid.org/0000-0002-2517-175X">https://orcid.org/0000-0002-2517-175X</a>
Jun Hyung Lee	<a href="https://orcid.org/0000-0002-8682-3694">https://orcid.org/0000-0002-8682-3694</a>
Sang Gon Lee	<a href="https://orcid.org/0000-0002-4672-5811">https://orcid.org/0000-0002-4672-5811</a>
Euna Park	<a href="https://orcid.org/0000-0002-3230-1364">https://orcid.org/0000-0002-3230-1364</a>
Ju-Hee Park	<a href="https://orcid.org/0000-0002-3751-4086">https://orcid.org/0000-0002-3751-4086</a>
Sung-Eun Cho	<a href="https://orcid.org/0000-0001-8082-3100">https://orcid.org/0000-0001-8082-3100</a>
Ahram Yi	<a href="https://orcid.org/0000-0002-3107-936X">https://orcid.org/0000-0002-3107-936X</a>
Yeo-Min Yun	<a href="https://orcid.org/0000-0002-5485-8331">https://orcid.org/0000-0002-5485-8331</a>

## REFERENCES

1. Zhou H, Wang Y, Gatcombe M, Farris J, Botelho JC, Caudill SP, et al. Simultaneous measurement of total estradiol and testosterone in human serum by isotope dilution liquid chromatography tandem mass spectrometry. *Anal Bioanal Chem* 2017;409:5943–54.
2. Rosner W, Auchus RJ, Azziz R, Sluss PM, Raff H. Position statement: utility, limitations, and pitfalls in measuring testosterone: an endocrine society position statement. *J Clin Endocrinol Metab* 2007;92:405–13.
3. Zhou Y and Wang S. A robust LC-MS/MS assay with online cleanup for measurement of serum testosterone. *J Sep Sci* 2019;42:2561–8.
4. Herold DA and Fitzgerald RL. Immunoassays for testosterone in women: better than a guess? *Clin Chem* 2003;49:1250–1.
5. Rosner W, Vesper H; Endocrine Society; American Association for Clinical Chemistry; American Association of Clinical Endocrinologists; Androgen Excess/PCOS Society, et al. eds. Toward excellence in testosterone testing: a consensus statement. *J Clin Endocrinol Metab* 2010;95:4542–8.
6. La'ulu SL, Kalp KJ, Straseski JA. How low can you go? Analytical performance of five automated testosterone immunoassays. *Clin Biochem* 2018;58:64–71.
7. Vesper HW, Botelho JC, Shacklady C, Smith A, Myers GL. CDC project on standardizing steroid hormone measurements. *Steroids* 2008;73:1286–92.
8. Centers for Disease Control and Prevention, Hormone Standardization Program (CDC HoSt). [https://www.cdc.gov/labstandards/hs\\_standardization.html](https://www.cdc.gov/labstandards/hs_standardization.html) (Updated on Dec 2021).
9. U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER). Bioanalytical method validation guidance for industry. <http://www.fda.gov/regulatory-information/serach-fda-guidance-documents/bioanalytical-method-validation-guidance-industry> (Updated on May 2018).
10. CLSI. Liquid chromatography-mass spectrometry methods. 1st ed. C62-A. Wayne, PA: Clinical and Laboratory Standards Institute, 2014.
11. CLSI. Evaluation of linearity of quantitative measurement procedures. 2nd ed. EP06-ED2. Wayne, PA: Clinical and Laboratory Standards Institute, 2020.
12. French D. Development and validation of a serum total testosterone liquid chromatography-tandem mass spectrometry (LC-MS/MS) assay calibrated to NIST SRM 971. *Clin Chim Acta* 2013;415:109–17.
13. Rappold BA. Review of the use of liquid chromatography-tandem mass spectrometry in clinical laboratories: Part I-Development. *Ann Lab Med* 2022;42:121–40.

14. Rappold BA. Review of the use of liquid chromatography-tandem mass spectrometry in clinical laboratories: Part II-Operations. *Ann Lab Med* 2022;42:531-57.
15. Yun YM, Botelho JC, Chandler DW, Katayev A, Roberts WL, Stanczyk FZ, et al. Performance criteria for testosterone measurements based on biological variation in adult males: recommendations from the Partnership for the Accurate Testing of Hormones. *Clin Chem* 2012;58:1703-10.
16. CLSI. Defining, establishing, and verifying reference intervals in the clinical laboratory. 3rd ed. EP28-A3c. Wayne, PA: Clinical and Laboratory Standards Institute, 2010.
17. Architect 2nd Generation Testosterone Instruction 2P13 G6-9345/R03 B2P1WZ. [www.abbottdiagnostics.com](http://www.abbottdiagnostics.com)
18. Moal V, Mathieu E, Reynier P, Malhi  ry Y, Gallois Y. Low serum testosterone assayed by liquid chromatography-tandem mass spectrometry. Comparison with five immunoassay techniques. *Clin Chim Acta* 2007; 386:12-9.
19. Kushnir MM, Rockwood AL, Roberts WL, Pattison EG, Bunker AM, Fitzgerald RL, et al. Performance characteristics of a novel tandem mass spectrometry assay for serum testosterone. *Clin Chem* 2006;52:120-8.
20. McCartney CR, Burt Solorzano CM, Patrie JT, Marshall JC, Haiseneder DJ. Estimating testosterone concentrations in adolescent girls: comparison of two direct immunoassays to liquid chromatography-tandem mass spectrometry. *Steroids* 2018;140:62-9.
21. Vesper HW, Bhasin S, Wang C, Tai SS, Dodge LA, Singh RJ, et al. Inter-laboratory comparison study of serum total testosterone measurements performed by mass spectrometry methods. *Steroids* 2009;74:498-503.
22. College of American Pathologists, Chemistry Resource Committee, Y-A 2021, Ligand special, participant summary 2021. [www.cap.org](http://www.cap.org)
23. College of American Pathologists, Accuracy Based Testing Committee, ABS-B 2021, Testosterone and estradiol accuracy, participant summary 2021. [www.cap.org](http://www.cap.org)
24. Cao ZT, Botelho JC, Rej R, Vesper H. Accuracy-based proficiency testing for testosterone measurements with immunoassays and liquid chromatography-mass spectrometry. *Clin Chim Acta* 2017;469:31-6.
25. Choi R, Park HD, Oh HJ, Lee K, Song J, Lee SY. Dried blood spot multiplexed steroid profiling using liquid chromatography tandem mass spectrometry in Korean neonates. *Ann Lab Med* 2019;39:263-70.
26. Ko DH, Lee K, Jeon SH, Song SH, Yun YM, Chun S, et al. Simultaneous measurement of serum chemical castration agents and testosterone levels using ultra-performance liquid chromatography-tandem mass spectrometry. *J Anal Toxicol* 2016;40:294-303.
27. Olisov D, Lee K, Jun SH, Song SH, Kim JH, Lee YA, et al. Measurement of serum steroid profiles by HPLC-tandem mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* 2019;1117:1-9.

**Supplemental Data Table S1.** Reference values for testosterone in several institutes measured based on LC-MS/MS

(A) ARUP

<https://ltd.aruplab.com/Tests/Pub/2004246> and <https://ltd.aruplab.com/Tests/Pub/0081058>

(B) Mayo Clinic Laboratories

<http://www.mayocliniclabs.com/test-catalog/Clinical+and+Interpretive/8533>

(C) Quest Diagnostics

<https://testdirectory.questdiagnostics.com/test/test-detail/15983/testosterone-total-ms?p=r&q=Testosterone,%20Total,%20MS&cc=MASTER>

(D) Abbott Architect

(ng/mL)	Men	Women
21–49 yr	2.40–8.71	0.14–0.53
≥ 50 yr	2.21–7.16	0.12–0.36

(E) Roche Elecsys

(ng/mL)	Men		Women	
	Median	5th–95th quantile	Median	5th–95th quantile
20–49 yr	5.360	2.490–8.360	0.271	0.084–0.481
≥ 50 yr	4.760	1.930–7.400	0.162	0.029–0.408

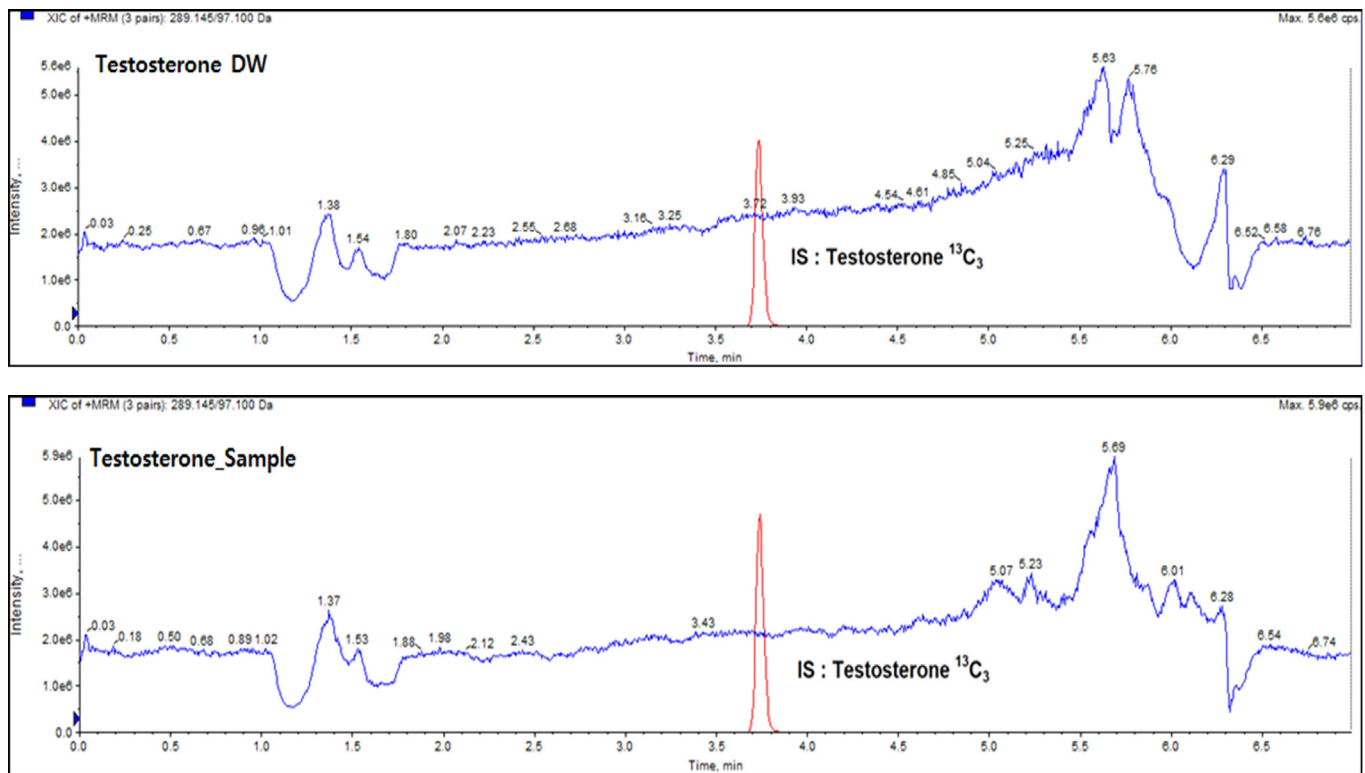
Abbreviation: LC-MS/MS, liquid chromatography-tandem mass spectrometry.

**Supplemental Data Table S2.** Method comparison of LC-MS/MS and the CMIA at low testosterone concentrations

Group	Age mean $\pm$ SD (yr)	Passing–Bablok regression analysis		<i>r</i>
		Slope	Intercept	
Men (N=40)	74.3 $\pm$ 12.7	1.868 (1.429–2.400)*	–0.0189 (–0.0616 to 0.0114)	0.7432
Women (N=40)	34.1 $\pm$ 9.6	0.885 (0.721–1.154)	0.0346 (–0.0238 to 0.0764)	0.7868
Boys (N=40)	10.5 $\pm$ 1.4	1.856 (1.212–3.000)	–0.0444 (–0.1270 to 0.0027)	0.5406
Girls (N=40)	11.6 $\pm$ 3.2	0.844 (0.732–1.000)	0.0342 (0.0205–0.0410)	0.9285

\*Numbers in parentheses are 95% confidence intervals of Passing–Bablok regression analysis.

Abbreviations: LC-MS/MS, liquid chromatography-tandem mass spectrometry; CMIA, chemiluminescent microparticle immunoassay.



**Supplemental Data Fig. S1.** Ion suppression analysis. No significant ion suppression or enhancement at the corresponding retention time was observed in tandem mass spectrometry analysis.