serum testosterone levels measured by ID-LC-MS/MS in postmenopausal women versus those in women who underwent bilateral oophorectomy

Hong N. Bui, Eduard A. Struys, Frans Martens, Willem de Ronde, Linda M. Thienpont, Peter Kenemans, Marieke O. Verhoeven, Cornelis Jakobs, Hilde M. Dijstelbloem, Marinus A. Blankenstein

*Annals of Clinical Biochemistry* 2010;47:248-252
ABSTRACT

Differentiation between subtle changes in low serum testosterone concentrations, common in women and children, is not possible with current commercially available assays. The objectives of the study were to develop a method based on stable isotope dilution-liquid chromatography-tandem mass spectrometry (ID-LC-MS/MS) with adequate sensitivity and specificity and to investigate the applicability of this assay in serum samples from pre- and postmenopausal women.

METHOD. For 16 women, testosterone levels were measured in blood samples drawn two years before and after physiological menopause, and for eight women in samples drawn before and after bilateral oophorectomy. Testosterone was extracted from serum, derivatized and analysed on an LC-MS/MS.

RESULTS. The developed ID-LC-MS/MS method allowed for specific and reproducible measurement of testosterone. Comparison with stable isotope dilution-gas chromatography coupled to mass spectrometry detection by Deming regression analysis gave a slope of 1.025 and an intercept of 0.055 nmol/L ($r = 0.9998$). A significant decrease was found in testosterone concentrations before and after bilateral oophorectomy ($p = 0.02$), whereas no significant difference was found before and after natural menopause ($p = 0.4$).

CONCLUSION. The ID-LC-MS/MS assay measures serum testosterone with acceptable accuracy and is useful in female samples, supporting the conclusion that the postmenopausal ovary contributes to circulating testosterone. To our knowledge, our analytical method compares favorably to similar published methods in terms of sensitivity. The sensitivity and specificity of this method comply with the reference method for measurement of testosterone in serum samples of women, children and men suffering from hypogonadism and can also be used for men with testosterone in the reference range.

INTRODUCTION

Measuring serum testosterone levels in women and children has always been an analytical challenge, as small changes in their already low testosterone concentrations have strong physiological relevance. For instance, Judd et al. [1] demonstrated a statistically significant decline in serum testosterone levels in women after bilateral oophorectomy. Subsequent research established the ovarian androgen production in postmenopausal women [2–6]. Furthermore, differentiation of several testosterone-related conditions, such as hirsutism in women, gender assignment of newborns with ambiguous genitalia and pubertal stage in children, depends on small changes in testosterone levels. Additionally, endocrine treatment of prostate cancer is aimed at lowering testosterone to ‘castration levels’. To measure subtle changes in low testosterone concentrations, an adequate and reliable assay is necessary. Androgen
concentrations are generally measured using immunoassays, but unfortunately these types of assays are known to lack sensitivity and specificity, which results in poor performance at the low concentration range. This issue has been addressed extensively in the literature [7–17]. Large sample volumes, i.e. 1–2 mL, are needed to compensate for the lack of sensitivity, while extensive sample preparation is required to compensate for low specificity. The reference method for testosterone is based on stable isotope dilution-gas chromatography coupled to mass spectrometry (ID-GC-MS) detection [9], but this method is laborious and also requires large sample volumes. An alternative upcoming technique in clinical chemistry, stable isotope dilution-liquid chromatography-tandem mass spectrometry (ID-LC-MS/MS), eliminates the need for large sample volumes due to its high sensitivity and specificity. We therefore developed an ID-LC-MS/MS-based assay for quantifying a broad range of serum testosterone concentrations typically found in women, children, hypogonadal and healthy men. In this paper, we report on the analytical validation and first clinical application of our ID-LC-MS/MS method.

EXPERIMENTAL SECTION

Laboratory methods. All sample preparations were performed in borosilicate glass tubes (13 x 100 mm²). A stable isotopically labeled internal standard (testosterone-16,16,17-D₃; CDN Isotopes, Pointe-Claire, Canada) was dissolved in ethanol (Merck, Darmstadt, Germany, absolute p.a.) and diluted to a concentration of 1.67 ng/mL. Prior to work-up, 50 µL of the internal standard solution was evaporated to dryness and 100 µL of the specimen (sample; control; calibrator) was added. Testosterone was extracted using 2 mL of hexane/ diethyl ether 4:1 (v/v) (hexane: Rathburn Chemicals Ltd, Walkerburn, Scotland, glass distilled grade; diethyl ether: Merck, spectroscopy grade). After vortex mixing, centrifugation and freezing of the aqueous layer in a dry ice/ethanol bath, the upper organic layer was decanted into a clean tube. The organic phase was subsequently dried under a stream of nitrogen and re-dissolved for derivatization in 200 µL of aqueous 3% (w/v) methoxylamine hydrochloride (MOX, Sigma-Aldrich, St Louis, MO, USA) solution, vortex-mixed and incubated at 80°C for one hour. The derivatized testosterone was extracted using the same protocol and re-constituted in 100 µL methanol/ water 1:1 (v/v) (methanol: BDH, Dorset, UK, high performance liquid chromatography [HPLC] grade).

Sample analysis was performed on an API 4000 HPLC tandem-MS (Applied Biosystems Inc, Foster City, CA, USA) with electrospray ionization in the positive mode using an injection volume of 40 µL. The instrument settings were optimized for maximum signal intensity: ion spray voltage 4500 V, declustering potential 65 V,
entrance potential 15 V, collision energy 45 V, and collision cell exit potential 10 V. Nitrogen was used as the collision gas. Separation was achieved on a C<sub>18</sub> analytical column (SymmetryShield RP<sub>18</sub>, 2.1 x 100 mm<sup>2</sup>, 3.5 mm particle size: Waters, Milford, MA, USA) with an isocratic elution at a flow rate of 0.2 mL/min methanol/water 95:5 (v/v) containing 0.5% formic acid (Merck, p.a.). Total analysis time was four minutes. The monitored transitions were m/z 318 &gt; 126 and m/z 321 &gt; 126 for derivatized testosterone (T-MOX) and the internal standard (D<sub>3</sub>T-MOX), respectively. The seven-point calibration curve ranged from 0.10 to 25 nmol/L (0.030–7.3 ng/mL). The lower limit of quantification (LLOQ) was defined as the testosterone concentration expected at a signal-to-noise ratio (SNR) of 10. This was calculated on the basis of a female serum sample containing 0.11 nmol/L testosterone measured with an SNR of 41. The inter-assay variation was determined using pooled human serum samples: the serum samples were aliquoted, stored at -20°C and analyzed as described in 10 independent runs. Ion suppression was assessed by post-column infusion of derivatized internal standard at a constant flow, while injecting prepared serum samples without internal standard. Linearity was demonstrated by serial dilution of serum samples with phosphate-buffered saline (0.9% NaCl, pH 7.4, B Braun, Melsungen, Germany) with 0.2% human serum albumin (Sigma-Aldrich, 96–99% albumin) and recovery was determined by spiking different serum samples with different quantities of testosterone. The specificity of the method was tested by adding approximately 1.5 ng of androstenedione (Sigma-Aldrich) and 1.5 ng of dihydrotestosterone (DHT: Sigma-Aldrich) to serum samples prior to work-up to reach a final concentration of approximately 50 nmol/L of each steroid. These steroids are known to cause interference in immunoassays. The ID-LC-MS/MS assay was validated for accuracy as suggested by Thienpont et al. [10], i.e. by blind analysis of 38 serum samples (stored at -20°C and transported on dry ice), assigned with target values by an ID-GC-MS reference measurement procedure listed in the database of the Joint Committee for Traceability in Laboratory Medicine (http://www.bipm.org/jctlm/). The samples covered testosterone concentrations from 0.20 to 24 nmol/L. Total testosterone in serum was also determined with the ‘Coat-a-Count, Total Testosterone’ radioimmunoassay (Siemens, Munich, Germany).

Experimental subjects. Samples for the surgical menopause were obtained from a database of women with a hereditary high risk of breast and/or ovarian cancer based on their family history or established breast cancer 1 (BRCA1) or BRCA2 germline mutation. This group comprised eight women who underwent a prophylactic bilateral salpingo-oophorectomy. The average age at the time of surgery was 44.8 ± 6.6 y (mean ± standard deviation [SD]). Blood samples were drawn no more than one year before/after surgery. The women were required to have a regular menstrual cycle at the time of surgery. Samples for the natural menopause were obtained within the context of the project ‘Aging Women’; in a longitudinal study from a group of 16
women who experienced a natural menopause, defined as the last menstruation followed by at least 12 months of amenorrhea. The average age at the time of menopause was $50.7 \pm 2.5$ y. Blood samples from the menopausal women were drawn no more than two years before and after the menopause date. All serum samples were stored at $-80^\circ$C. Because the samples were collected during observational studies in patient care programs, no institutional review board approval was needed to be obtained. However, all women involved had given informed consent [18]. Women of both groups did not use hormonal contraception, hormone therapy or other medication interfering with the parameters of this study as described by Verhoeven et al. [18]. Results were evaluated with Wilcoxon’s rank (paired samples) using Medcalc 9.3 software (Medcalc Software, Mariakerke, Belgium).

**RESULTS**

Analytical validation of the ID-LC-MS/MS method showed inter-assay variations ($n = 10$) of 4% and 5% at 7.2 and 0.69 nmol/L, respectively. No ion suppression occurred at the retention time of testosterone. Dilution linearity gave a correlation coefficient ($r$) of 1.0000. Analysis of serum samples supplemented with testosterone to concentration ranging from 0.31 to 25 nmol/L ($n = 10$) gave an overall analytical recovery of $100 \pm 5\%$ (mean $\pm$ SD). The ID-LC-MS/MS method tested negative for interference from androstenedione and DHT. The assay featured a calculated LLOQ of 27 pmol/L. As
shown in Figure 1, Deming regression analysis of the results obtained for the accuracy assessment gave a slope of $1.025 \pm 0.007$ (mean ± standard error) and an intercept of $0.055 \pm 0.018$ nmol/L ($r = 0.9998$). In the natural menopause group, seven women showed an increase, while nine women showed a decrease in testosterone level after menopause. Statistical analysis showed no significant difference in testosterone levels before and after natural menopause ([mean ± SD] 0.94 ± 0.51 versus 0.86 ± 0.39 nmol/L; $p = 0.4$). Out of eight subjects in the surgical group, seven women showed a decrease in testosterone level after surgery, while one woman showed a slight increase in testosterone concentration. Overall, a significant decrease was found for levels before ([mean ± SD] 1.48 ± 0.85 nmol/L) and after (0.82 ± 0.42 nmol/L) surgery in women who had their ovaries removed ($p = 0.02$) (Figure 2). In the radioimmunoassay, only 10 samples (out of 48) yielded values within the detection range of the assay (data not shown).

**DISCUSSION**

Thorough validation of the described ID-LC-MS/MS assay for measuring total testosterone in serum demonstrates its reproducibility, linearity, specificity, analytical recovery, accuracy, and precision. These characteristics are in accordance with other methods published in literature [9,10]. The method is validated for a broad detection range (from 0.1 to 25 nmol/L), and is therefore suitable for the measurement of male samples. It was found, an additional hallmark of the method was that it had sufficient sensitivity to allow the use of small sample volumes.

To compare the method’s sensitivity, the absolute lower limit of quantification (LLOQ) of various published ID-LC-MS/MS methods for serum testosterone has been calculated. Sample volume, concentration and dilution steps, as well as injection volume, determine the absolute detected quantity. For the calculation of the absolute LLOQ at a SNR of 10, the LLOQ claimed by the authors, the definition of the LLOQ, initial sample volume, concentration and dilution steps, and injection volume have been taken into account. In accordance with this calculation, our ID-LC-MS/MS method, with a functional LLOQ of 0.10 nmol/L (SNR 40), injection volume of 40 µL and no dilution/concentration steps during sample preparation, is able to measure as low as 1.2 fmol testosterone at an SNR of 10. This comparison is summarized in Table 1 and shows that the sensitivity of our method compares favorably to published methods. We are currently investigating the possibilities for further optimization with a focus on automation of sample preparation.
When the assay is applied for analysis of samples from menopausal women, a decline in serum testosterone levels was found in women after bilateral oophorectomy, in contrast to the unchanged concentrations found in women experiencing natural menopause. As anticipated, we were unable to confirm this finding using a
commercial immunoassay due to its lack of sensitivity and specificity, highlighting the importance of state-of-the-art technology in the field of androgen research. Our results support the conclusion that the postmenopausal ovary is hormonally active and that its contribution to circulating testosterone levels is to such an extent that it may be responsible for half of the circulating testosterone, as indicated in Figure 2.

Data were not corrected for age-related decline of testosterone levels as no significant decreases of testosterone concentrations were found in women with natural menopause. The samples from women who underwent surgery were drawn no more than two years apart. The age-related decline in this group is negligible. The samples were drawn at random during the menstrual cycle, which explains the spread. When comparing the menopausal testosterone values with the postsurgical values, there is no significant difference in testosterone levels. The adrenal gland is responsible for the production of circulating testosterone in ovariectomized women. This organ is probably more active in the post-surgery women than in the postmenopausal women, due to the significantly younger age of the women who underwent surgery, 44.8 ± 6.6 y (mean ± SD) compared with 50.7 ± 2.5 y in the menopause group.

**TABLE 1**: Comparison of the calculated absolute LLOQ at an SNR of 10.

<table>
<thead>
<tr>
<th></th>
<th>LLOQ (nmol/L)</th>
<th>SNR defined</th>
<th>Sample volume (µL)</th>
<th>Calculated absolute LLOQ at SNR of 10 (fmol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present study*</td>
<td>0.025</td>
<td>10</td>
<td>100</td>
<td>1.1</td>
</tr>
<tr>
<td>Kushnir et al.* [13]</td>
<td>0.03</td>
<td>6.6 (0.1 nmol/L)</td>
<td>100</td>
<td>3.0</td>
</tr>
<tr>
<td>Turpeinen et al. [14]</td>
<td>0.15</td>
<td>10</td>
<td>500</td>
<td>3.8</td>
</tr>
<tr>
<td>Cawood et al. [15]</td>
<td>0.25</td>
<td>10</td>
<td>50</td>
<td>3.3</td>
</tr>
<tr>
<td>Moal et al. [16]</td>
<td>0.17</td>
<td>10</td>
<td>500</td>
<td>23</td>
</tr>
<tr>
<td>Vicente et al. [17]</td>
<td>0.21</td>
<td>12</td>
<td>1000</td>
<td>47</td>
</tr>
</tbody>
</table>

LLOQ: lower limit of quantification; SNR: signal-to-noise ratio
*Methods using derivatization

Beside the sensitivity and associated applicability in clinical settings, an additional important advantage of ID-LC-MS/MS is its potential for measuring concentrations of multiple steroid hormones in one run, thus creating a steroid profile [11,12]. The ID-LC-MS/MS technique allows such measurements without loss of sensitivity and specificity. This development could eliminate the need for multiple immunoassays,
not only saving sample, time and money, but also enhancing quality. The ability to measure low testosterone levels in combination with other steroid concentrations creates new perspectives, particularly for analysis of small sample volumes that may be collected from children or in research settings. Since testosterone concentrations are generally much lower in tissues than in serum, our newly developed ID- LC-MS/MS method may contribute to measure testosterone levels with adequate sensitivity and specificity.

REFERENCES


isotopically labeled testosterone derivatives as internal standards in LC-MS/MS

Hong N. Bui, Marinus A. Blankenstein, Annemieke C. Heijboer


We read with interest the recent paper by Owen and Keevil [1] regarding the importance of internal standard choice for testosterone measurement by liquid-chromatography tandem-mass spectrometry (LC-MS/MS). The authors investigated the performance of three isotopically labelled internal standards for testosterone, i.e. D₂-testosterone, ¹³C₃-testostosterone, and D₅-testosterone, using the same protocol and a substantial number of samples covering the whole clinical range. While we agree that a proper internal standard is imperative for accurate LC-MS/MS measurements, the observed differences illustrated by their data might have clinical impact and left a few questions.

Stable isotope-dilution with ¹³C-isotopes to correct for losses during sample preparation is considered to be the standard because ¹³C-labelled compounds co-elute exactly with their endogenous counterpart. Moreover, the stability of deuterium labelled-isotopes is highly dependent on the location of the labels which might exchange with hydrogen-atoms. Nevertheless, it is generally accepted that any stable isotope could be sufficient. Comparison between D₂ and ¹³C₃ by Owen and Keevil [1] resulted in differences that are greater than the reported imprecision of the assay. It seemed fair to assume that results obtained with D₂ were considered the target as
these results have previously been shown to concur excellently with those obtained with a reference method [2]. However, two different D$_3$ manufacturers are reported and it is unclear whether the same D$_2$ was used in the comparison with the reference method [2] as in the recent paper [1].

The discrepancy between D$_2$ and D$_5$ is thought-provoking because, in our experience, an unexpectedly large difference was reported. According to their data, differences up to 11 nmol/L were observed when changing the internal standard from D$_2$ to D$_5$. Isotopically labelled components tend to elute slightly shifted from their endogenous counterpart, especially when high resolution chromatography such as UPLC is used. Ion suppression, which is known to be sample-specific, could cause these differences in outcomes. For proper interpretation, it is essential to know to what degree D$_5$ and endogenous testosterone co-elute in the method of Owen and Keevil, and if the deviated samples show a different ion suppression profile.

![Figure 1: Bland-Altman bias plot of D$_3$-testosterone compared with D$_5$-testosterone results. The solid line at -0.09 nmol/L represents the mean difference. The dotted lines at 0.77 nmol/L and -0.94 nmol/L represent the ± 1.96 SD interval, respectively.](image)

We tested the differences between D$_3$ and D$_5$ on testosterone measurements using a method that was reported previously and also showed excellent agreement with a reference method [3]. The internal standards were tested for purity prior to the development of the method and both internal standards co-eluted with endogenous testosterone. Furthermore, ion suppression was tested on multiple samples during method validation. Samples from men and women (n = 40), covering the whole clinical range (0.16 – 50.7 nmol/L), were prepared using D$_3$ and subsequently D$_5$ as internal standard. To exclude possible instrumental shifts, the samples were measured alternately. Comparison by Passing-Bablok analysis resulted in the equation: testosterone(D$_5$) = 1.00 * testosterone(D$_3$) − 0.00 nmol/L. As shown in figure 1, minimal differences were observed.

Even though our observations differ from those reported by Owen and Keevil, we agree that the choice of internal standard is crucial. We were surprised to see the
differences reported. In general, the use of any internal standard should be validated. If necessary, the method could be adjusted to ensure proper functioning of the internal standard.

REFERENCES


