The importance of measuring the uncertainty of second-generation total testosterone analysis

Sema Nur Ayyildiz
Department of Biochemistry, University Faculty of Health Sciences, Ankara Training and Research Hospital, Ankara, Turkey

Abstract

Objectives: Testosterone is present in both genders; however, the level is low in women, whereas it is high in men. A measurement of total testosterone (TT) is widely requested by clinicians in daily practice, and is used in both diagnosis and treatment. The aim of this study was to calculate the measurement uncertainty of TT and to make a contribution of new research to the literature.

Methods: The identification of measurement, factors that can affect the measured value, laboratory reproducibility bias (uRw²), laboratory and method bias measurement uncertainty (Ubias), uncertainty of calibration (uCref), uncertainty of external quality control data (uEQA), combined measurement uncertainty (Uc), and extended measurement uncertainty (U) were evaluated and reported.

Results: The results of determining the measurement uncertainty of TT measurement in this laboratory were as follows: 1) uRw² = (RSDlevel 1² + RSDlevel 2² + RSDlevel 3²) / n: (20.90² + 17.48² + 7.00²) / 3 = 263.79; 2) uCref = 0.012; Cref = 2.35; k=2. RuCref = (100 x uCref) / (k x Cref): (100 x 0.012) / (2 x 2.35) = 0.255; RuCref² = 0.065; 3) Ubias² = RuCref² + uEQA²: 0.065 + 105.60 = 105.67; 4) Uc=√ (uRw² + Ubias²): √ (263.79 + 105.67) = 19.22; 5) U = 2√ (uRw² + Ubias²): 2 x 19.22 = 38.44%; 6) Result report (serum TT): ± 38.44%.

Conclusion: The measurement uncertainty result for TT calculated using the top-down method was ± 38.44% with a 95% confidence interval. The individual measurement uncertainty result for each test should be given to the clinician and the patient together with the test results. The TT measurement uncertainty of other methods should also be known at the international level.

Keywords: Second generation immunoassay, total testosterone, uncertainty of measurement
also lead to different outcomes in TT level [5]. When interpreting test results, it is necessary to consider individual variations. Measurement uncertainty is an important entity in all laboratories. It affects test results and reflects the quality of the result given to the patient and doctor [6,7]. However, there is still no full consensus on the calculation and methodology of uncertainty. Currently, the International Vocabulary of Metrology 2 (VIM2), the Guide to Expression of Uncertainty in Measurement 1 (GUM1), the International Organization for Standardization/International Electrotechnical Commission (ISO/IEC), and the VIM3 guidelines are used in the evaluation and identification of measurement uncertainty [6,8]. In addition, some countries and laboratories have their own methods of assessing uncertainty [9,10]. Testosterone results are important to both the patient and the clinician. TT is widely requested in daily practice by clinicians for use in diagnosis and treatment. The aim of this study was to calculate the measurement uncertainty for TT and to discuss these findings, which have not been published before.

### Materials and Methods

This study was performed using the top-down method, one of the methods of determining measurement uncertainty [6,7,9,11].

1. **Identification of Measurement**

   Abbott Architect i-2000SR (Serial No: ISR 04243; Abbott Laboratories, Abbott Park, IL, USA) immunoassay and Abbott Architect 2nd Generation Total Testosterone reagent kit (2P13, reagent lot no: 10413UP00, 10406UP00, 10391UP00, 10375UP00, 10369UP00, 10368UP00, and 10353UP00; Abbott Laboratories, Abbott Park, IL, USA) were used in the Ordu University Hospital clinical biochemistry laboratory between 2014 and 2015. TT was determined with chemiluminescent microparticle immunoassay (CMIA) technology and a delayed one-step method using flexible assay protocols (Chemiflex; Abbott Laboratories, Abbott Park, IL, USA).

2. **Factors that may affect measurement**

   Several factors may influence the results when the measurement is carried out. These factors include equipment failure, the electrical system, calibration errors, contamination-induced errors, temperature changes, pressure and humidity changes, vibration, employee error, and method effects. Other undetermined factors may also influence measurement values.

3. **Laboratory reproducibility bias (uRw²)**

   As control materials, 2G-Testo level 1, level 2, level 3 internal control serum (lot number G3-9729; Abbott Laboratories, Abbott Park, IL, USA) and Technopath level 1, level 2, level 3 internal control serum (lot number 33902150, 33607140, 33505140; Technopath Clinical Diagnostics, Ballina, Co. Tipperary, Ireland) were used. Control samples covered the 15-month period between 2014 and 2015. Internal quality control results were checked for normal and pathological low and high controls. The results of 15 months were calculated as standard deviation (SD) and relative standard deviation (RSD) according to the results of 3 evaluations (Table 1).

4. **Laboratory and method bias measurement uncertainty (Ubias)**

   If measurement results are given with reference and deviation values, the measurement uncertainty values must be written in the report [7]. However, even if the bias value is 0, it should be added to the measurement uncertainty [6]. Here, uncertainty from the calibration and external control performance data were used. The formula was follows: \( Ubias^2 = RuCref^2 \) (relative uncertainty of calibration) + \( uEQA^2 \) (uncertainty of external quality control data).

5. **Uncertainty of calibration (uCref)**

   The uCref data were requested from Abbott Laboratories, and the information provided was applied. In the calibration, the average of the uncertainties from the total testosterone calibrator (Arc Testo, product number 02P1301; Abbott Laboratories, Abbott Park, IL, USA) at calibrations A, B, C, D, E, and F at

---

### Table 1. Uncertainty between reproducibility and analytical processes of internal control data in the laboratory

<table>
<thead>
<tr>
<th>TT (n=157) level 1 control (ng/mL)</th>
<th>TT (n=168) level 2 control (ng/mL)</th>
<th>TT (n=124) level 3 control (ng/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.134</td>
<td>0.761</td>
</tr>
<tr>
<td>SD</td>
<td>0.028</td>
<td>0.133</td>
</tr>
<tr>
<td>RSD</td>
<td>20.90</td>
<td>17.48</td>
</tr>
<tr>
<td>( 100\times(\text{SD/mean}) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( uR^2_w )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RSD: Relative SD; SD: Standard deviation; TT: Total testosterone; uRw²: Laboratory reproducibility bias.
6. Uncertainty of external quality assessment (uEQA)
This indicates the percentage of laboratory test results deviating from the average of the test results of the control group. The averages of the percent deviations were obtained using external quality control data of the tests done with Abbott Architect. The percentage deviation from the mean is given below and is included in the external quality control data sheets. The formula used was: Percent deviation from mean = [(test result comparator group mean / comparative group mean) * 100]. For each single value emerging, the following equation was used for the uncertainty arising from the overall performance data of the laboratory: uEQA = \( \sqrt{\sum \text{bias}_{EQA}^2} / n \).
The uncertainty of the serum TT level external quality control data was determined according to the results of RIQAS (Randox External Quality Control Material; Randox Laboratories, Crumlin, County Antrim, Northern Ireland) program analysis of 20 samples every month from April 2014 to October 2015. The calibrations and the values calculated with external quality control data and the \( \text{bias}^2 \) were calculated according to the formula given in the laboratory and the methodological bias measurement uncertainty paragraph above (Table 3).

7. Combined measurement uncertainty (Uc)
When the data were examined, the combined Uc was calculated using \( u_{Rw} \) and \( \text{bias}^2 \) in the following formula: Uc = \( \sqrt{u_{Rw}^2 + \text{bias}^2} \).

8. Extended measurement uncertainty (U)
The Uc was multiplied by the factor k. The formula was: U = 2\( \sqrt{u_{Rw}^2 + \text{bias}^2} \).

9. Reporting of extended measurement uncertainty
TT (ng/mL) should be reported with a 95% CI.

Statistical Analysis
Basic calculations for measurements of uncertainty and laboratory reproducibility bias were performed using Microsoft Excel 2010 (Microsoft Corp., Redmond, WA, USA) software.

Results
The results of determining the measurement uncertainty in TT measurement in this laboratory were as follows:

1) \( u_{Rw2} = (\text{RSD}_{\text{level}1}^2 + \text{RSD}_{\text{level}2}^2 + \text{RSD}_{\text{level}3}^2) / n \)
   \( = (20.90^2 + 17.48^2 + 7.00^2) / 3 \)
   \( = 263.79 \) (Table 1)

2) \( \text{Cref} = 0.012; \text{Cref} = 2.35; k = 2 \) (Table 2)
   \( \text{RuCref} = (100 \times \text{uCref}) / (k \times \text{Cref}) \)

Table 2. Nominal uncertainty values and mean values for total testosterone according to calibrator level (nmol/L-ng/mL)

<table>
<thead>
<tr>
<th>Calibrator level</th>
<th>Nominal values nmol/L</th>
<th>Uncertainty nmol/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>0.1</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.03 ng/mL)</td>
<td>(0.00 ng/mL)</td>
</tr>
<tr>
<td>C</td>
<td>0.2</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.06 ng/mL)</td>
<td>(0.00 ng/mL)</td>
</tr>
<tr>
<td>D</td>
<td>1.6</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.46 ng/mL)</td>
<td>(0.00 ng/mL)</td>
</tr>
<tr>
<td>E</td>
<td>12.5</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>(3.61 ng/mL)</td>
<td>(0.02 ng/mL)</td>
</tr>
<tr>
<td>F</td>
<td>30.0</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(8.65 ng/mL)</td>
<td>(0.05 ng/mL)</td>
</tr>
<tr>
<td>Mean</td>
<td>7.4</td>
<td>0.047</td>
</tr>
<tr>
<td>( C_{\text{ref}} - \text{uC}_{\text{ref}} )</td>
<td>(2.35 ng/mL)</td>
<td>(0.012 ng/mL)</td>
</tr>
</tbody>
</table>

Table 3. Bias of total testosterone derived from external quality control data

<table>
<thead>
<tr>
<th>EQA samples</th>
<th>TT bias (deviation %)</th>
<th>TT biasEQA&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.2</td>
<td>67.24</td>
</tr>
<tr>
<td>2</td>
<td>-2.4</td>
<td>5.76</td>
</tr>
<tr>
<td>3</td>
<td>5.3</td>
<td>28.09</td>
</tr>
<tr>
<td>4</td>
<td>13.2</td>
<td>174.24</td>
</tr>
<tr>
<td>5</td>
<td>17.6</td>
<td>309.76</td>
</tr>
<tr>
<td>6</td>
<td>1.2</td>
<td>1.44</td>
</tr>
<tr>
<td>7</td>
<td>11.6</td>
<td>134.56</td>
</tr>
<tr>
<td>8</td>
<td>15.2</td>
<td>231.04</td>
</tr>
<tr>
<td>9</td>
<td>8.4</td>
<td>70.56</td>
</tr>
<tr>
<td>10</td>
<td>27.7</td>
<td>767.29</td>
</tr>
<tr>
<td>11</td>
<td>4.6</td>
<td>21.16</td>
</tr>
<tr>
<td>12</td>
<td>-11.6</td>
<td>134.56</td>
</tr>
<tr>
<td>13</td>
<td>1.7</td>
<td>2.89</td>
</tr>
<tr>
<td>14</td>
<td>-1.9</td>
<td>3.61</td>
</tr>
<tr>
<td>15</td>
<td>-1.7</td>
<td>2.89</td>
</tr>
<tr>
<td>16</td>
<td>0.6</td>
<td>0.36</td>
</tr>
<tr>
<td>17</td>
<td>11.6</td>
<td>134.56</td>
</tr>
<tr>
<td>18</td>
<td>3.9</td>
<td>15.21</td>
</tr>
<tr>
<td>19</td>
<td>-2.5</td>
<td>6.25</td>
</tr>
<tr>
<td>20</td>
<td>0.7</td>
<td>0.49</td>
</tr>
<tr>
<td>( \sum \text{bias}_{EQA}^2 )</td>
<td>2111.96</td>
<td></td>
</tr>
<tr>
<td>UEQA&lt;sup&gt;2&lt;/sup&gt; = (( \sum \text{bias}_{EQA}^2 )) / n</td>
<td>105.598</td>
<td></td>
</tr>
</tbody>
</table>

EQA: External quality assessment; TT: Total testosterone.

0.0, 0.1, 0.2, 1.6, 12.5, and 30.0 nmol, the relative value obtained with the following formula was used: RuCref = (100 x uCref) / (k x Cref). The k factor (k) is 2 for a 95% confidence interval (CI) for TT. The nmol/L value was converted to ng/mL (Table 2).
Özmen [22] defined an extended mean measurement uncertainty of ±38.44%. The New York State Department of Health has set a tolerable error limit of ±20 ng/mL or 25% for testosterone [23].

The testosterone assay has been a subject of debate for many years [24]. A high rate of variation, especially at low concentrations, and cross-reactions with other steroids affect testosterone results. For this reason, it plays a key role in determining net TT in serum and in the analysis of incorrect results where the SHBG concentration is incorrect [24]. Measurement of serum TT, especially with a short incubation time, is difficult due to the shift of testosterone from SHBG. Solvents that cause the release of testosterone from SHBG may not be compatible with automated immunoassays. Alternatives may be inadequate in this respect [25,26].

The high measurement uncertainty of serum TT in this study supports the literature on testosterone analysis [24]. The results of the measurement uncertainty should be presented in the report to guide clinicians and patients. If testosterone results are supported by clinical findings, treatment should be planned accordingly.

In the scientific arena, the contribution of test results to the regulation of treatment plans is thought to be the most important; honesty should be the forerunner in the development and enforcement of all scientific paradigms and methods [17]. There may be many measurements in a process or analysis of a test and there is an uncertainty in each measurement. Any source of error that might affect the preanalytical, analytical, and postanalytical stages should be identified and an uncertainty value should be given [27]. The goal is to improve the measurement with the greatest uncertainty in the system in order to improve the quality of the measurement result [28].

In our study, we found the measurement of uncertainty value for serum TT to be high. Methods of TT measurement should be reviewed and improved. However, measurement uncertainty values are not calculated and reported for parameters currently being studied in clinical laboratories. Measurement uncertainty is often used for research purposes, but is not used routinely. However, measurement uncertainty is an important issue for all parties: laboratories, clinicians, and patients, since the results will affect diagnosis and treatment procedures. For this reason, every laboratory should routinely determine measurement uncertainty. When the clinician associates these results with the patient’s clinical evaluation, it could affect decisions. In short, measurement uncertainty could determine the clinician’s decision.

The methods used to assess measurement uncertainty, the ISO and the GUM guidelines, are still being discussed, in an effort to provide the best guidelines [29,30]. These calculations are complicated and elaborate, so their applicability in practice is diminishing. Workgroups for measurement uncertainty using simpler and easier formulas have been established in different countries [27]. These study groups recommend the percent coefficient of variation (CV) x 2 formula for measurement uncertainty. The CV used in this calculation is based on the existence of many parameters already contributing to un-

\[
\text{Result report (serum TT)} = \text{Measured value } \pm 38.44\%
\]

\[
U = 2\sqrt{(u_{Rw}^2 + ubias^2)} = 2\sqrt{(263.79 + 105.67)} = 19.22
\]

\[
RuCref = (100 \times 0.012) / (2 \times 2.35) = 0.255
\]

\[
RuC = \sqrt{(263.79 + 105.67)} = 19.22
\]

Discussion

Testosterone measurements are used in the diagnosis and treatment of primary and secondary hypogonadism; male sex hormone-related diseases, such as delayed or premature puberty, loss of libido, and impotence; tumor-induced hirsutism and virilization in women; polycystic ovary syndrome; and adrenogenital syndromes [12,13].

Testosterone is carried in the blood with sex hormone-binding globulin (SHBG) and albumin. Testosterone has a high capacity-low affinity for albumin binding and a high capacity-low affinity for SHBG binding. Men have lower SHBG relocation rates than women. In vivo tissue release studies have shown that virtually all albumin-bound testosterone is bioavailable for tissue uptake [14].

Measurements in laboratories are not always exact and precise. It is necessary to express this in absolute numbers. In taking measurements and determining results, factors such as calibrations, including internal and external laboratory practices; internal and external quality; total error; and measurement uncertainty as well as patient-related features play an important role [15-21]. Measurement uncertainty is a statistical parameter that is generated by arithmetic operations that describe the fluctuations that occur in a test result given a certain CI. It is the whole of a series of operations, and may be calculated from either the top-down or bottom-up method [6,19].

We used the top-down method to calculate the measurement uncertainty of TT. In this method, the use of both in-laboratory and out-of-laboratory, as well as inter-laboratory performance data, are said to facilitate uncertainty calculations [11,16,17,19].

In our study, we found a serum TT measurement uncertainty value of ±38.44%. Özmen [22] defined an extended mean measurement uncertainty in TT of about 12% in his study. In addition, some institutions and organizations recommend the use of total tolerable error and biological variation. The New
certainty [31, 32]. Or, whenever clinical internal and external data are obtained, measurement uncertainty should be provided automatically through software programs.

In the literature, measurement uncertainty studies are limited in number. There are even fewer regarding hormonal tests. Measurement uncertainty calculations have been mandatory since 2002 in accordance with the In Vitro Diagnostic Medical Devices directives and the related European Commission decision. Clinical chemistry and biochemistry laboratories are not considered to be excluded from these applications [27]. Providing the clinician with the measurement uncertainty of the results of the analysis will be of great benefit in the treatment of the disease and in the construction of the patient’s follow-up care.

Uncertainty of measurement normally indicates that measurement results are made with validated methods and provides the level of reliability [6]. It never shows that the accuracy of the results is doubtful [6]. If the serum testosterone level is sub-normal according to the American Endocrinology Association guidelines, the test should be repeated [33]. Immunoassay, massspectrometry, modular E170 electroimmunassay, or gas chromatography mass spectrometry methods may be used for TT measurement. Each of these methods has advantages and disadvantages [12]. The Endocrine Society suggests that a single method may detect a low TT value, and morning TT measurements should be repeated twice [12]. The highest level of serum TT is seen between 7 am and 11 am [12,33]. In this study, the measurement uncertainty of TT measurements was performed using a second-generation CMIA method. The uncertainties of TT measurement using other methods should also be conducted and published in the literature. These issues remain unclear. Although the guidelines of the Endocrine Society and our country’s health policies state that the liquid chromatography/mass-mass spectrometry method is the best method [14], we have no data about the measurement uncertainties of this procedure.

**Conclusion**

Measurement uncertainty is a measure of the distribution of measurement results. Clinician notification is important, both for increasing the reliability of laboratories and for planning the treatment options of patients. The measurement uncertainty results of TT were 38.44% with a 95% CI using the top-down method. This rate is high, and needs to be reflected in TT results. When these test results are given, international standards should be applied; however, countries, and even laboratories, should carry out measurement uncertainty studies involving all tests appropriate to their own conditions, which will increase the reliability of the results. The individual measurement uncertainty results for each test should be given together with the test results to the clinician and the patient. The TT measurement uncertainties of other methods should also be determined at the international level.

**Conflict of interest:** There are no relevant conflicts of interest to disclose.

**Peer-review:** Externally peer-reviewed.


**References**


