Diagnostic Performance of Serum Total Testosterone for Japanese Patients with Polycystic Ovary Syndrome

TAKESHI IWASA, TOSHIYA MATSUZAKI, MASAIRO MINAKUCHI, NAOKO TANAKA, FUMI SHIMIZU, YOHKO HIRATA, AKIRA KUWAHARA, TOSHIYUKI YASUI, MASAHIRO MAEGAWA AND MINORU IRAHARA

Department of Obstetrics and Gynecology, Institute of Health Biosciences, The University of Tokushima Graduate School, Tokushima 770-8503, Japan

Abstract. It is reported that the incidence of clinical and biochemical hyperandrogenism may be lower in Japanese patients with PCOS. Hyperandrogenism is included as a referential but not as an essential factor in the diagnostic criteria of the Japanese Society of Obstetrics and Gynecology (JSOG 1993). However, some patients with the typical clinical features of PCOS are not diagnosed with PCOS using JSOG 1993 criteria because they do not have a high LH level, which is defined as essential for diagnosis. In this study, we compared total testosterone (T) levels between Japanese patients with PCOS diagnosed using the JSOG 1993 criteria and normal menstrual women (controls). Fifty controls and 46 patients with PCOS were enrolled in this study. Furthermore, we evaluated the sensitivity and specificity of each cut-off value of T. The mean T level of patients with PCOS was significantly higher than that of the control (86 ± 48 vs 68 ± 46, P<0.01), and the prevalence rates of hyperandrogenism (T >114 ng/dL; defined as the mean +2SD of the control) were 10.2% in patients with PCOS and 4% in controls. The area under the ROC curve of T was 0.72, and there was no decision threshold to diagnose PCOS by T alone with both high sensitivity and high specificity. If the threshold is set as 110 ng/dL in order to gain high specificity, 94% of women whose serum level passed the threshold will be patients with PCOS. Although T should not be used as an independent essential factor of Japanese PCOS, it might be useful as a complementary factor in order to diagnose patients who have typical clinical features of PCOS but does not fulfill the JSOG 1993 criteria for PCOS.

Key words: PCOS, Hyperandrogenism, Total testosterone

POLYCYSTIC ovary syndrome (PCOS) is a common ovulatory disorder in women of reproductive age and is reported to have a prevalence of 6%–10% [1, 2]. PCOS is a syndrome of ovarian dysfunction and its cardinal features are hyperandrogenism and polycystic ovary morphology. Other common features are insulin resistance and elevated serum LH levels.

The 2003 Rotterdam ESHRE/ASRM-sponsored PCOS consensus workshop group concluded that no single diagnostic criterion was sufficient for a clinical diagnosis of PCOS and that two out of three criteria had to be met to fit the definition: chronic anovulation, clinical and/or biochemical evidence of hyperandrogenism and polycystic ovaries [3]. These criteria have been widely accepted recently; however, clinical and biochemical hyperandrogenism is less prevalent in Japanese patient than in Western one, and a high luteinizing hormone (LH) level is common in Japanese patients with PCOS [4–6]. The PCOS criteria of the Japanese Society of Obstetrics and Gynecology (JSOG 1993) consist of the association of all three of the following factors: chronic anovulation, high LH level, and presence of polycystic ovaries (PCO) [4]. Hyperandrogenism is included as a referential factor in the criteria of JSOG 1993; however, the definition of hyperandrogenism, particularly biochemical hyperandrogenism, has not been evaluated adequately in Japanese women. The decision threshold of total testosterone (T), one of the various androgens in the blood, has not been precisely defined, although it is often measured in
patients with PCOS. The problems are 1) commercially available kits for T demonstrate significant between-kit variability, especially at low concentrations which is important for the diagnosis of hyperandrogenic women [7, 8], 2) the decision threshold should be defined based on the purpose, and that is not always identical to the upper limit of the reference range; when we determine the decision threshold, we should consider the trade off between sensitivity and specificity. In this study, we compared T levels between patients with PCOS diagnosed using the criteria of JSOG 1993 and normal menstrual women (controls), using the commercially available kit, ARCHITECT testosterone (ARCHITECT T; Abbott Japan Inc., Tokyo, Japan). Furthermore, we evaluated the sensitivity and specificity of each cut-off value of T from the receiver-operating characteristic (ROC) curve, and decided the decision threshold appropriate for the diagnosis of Japanese PCOS.

Materials and Methods

Subjects

Fifty controls and 46 patients with PCOS were enrolled in this study. Controls were euthyroid with normal serum levels of prolactin and TSH, and had a history of regular menstrual cycle (25–35 days in length) with evidence of ovulation (midluteal phase with progesterone >10 ng/mL). They had ultrasonically normal ovaries. PCOS was diagnosed using the criteria of JSOG 1993 described above. Written informed consent was obtained from all patients.

Blood sampling

Serum samples were obtained from these patients. Serum samples in controls were collected in the follicular phase, ovulatory phase and luteal phase. In patients with PCOS, sampling was not timed to the menstrual cycle because all patients were oligomenorrheic or amenorrheic.

Hormone Assays

All samples were assayed for total testosterone levels using ARCHITECT T, which uses a chemiluminescent immunoassay method based on a patented acridium carboxamide chemiluminescent assay. The analytical sensitivity of the assay was 8 ng/dL, and the intra-assay and inter-assay coefficients of variation were <4.5% and <8.0%, respectively. As commercially available kits for T level demonstrate significant between-kit variability [7, 8], samples were also assayed using DPC total testosterone kit (Diagnostic Products Corp. Tokyo, Japan) and liquid chromatography-mass spectrometry (LC-MS) (Teikokuzouki Co., Tokyo, Japan). The DPC total testosterone kit used iodinated tracer and T-specific antibodies immobilized to the wall of tube (sensitivity of 4 ng/dL). LC-MS is the most precise method for the measurement of steroid hormone (sensitivity of 0.14 ng/dL). Measured values of ARCHITECT T and the other kits were compared in order to assess the reliability of ARCHITECT T.

Statistical analysis

Data are expressed as the mean ± 2SD. For the analyses of clinical data, statistical significances were determined by $t$ test. All $P$ values <0.05 were considered statistically significant. Receiver-operating characteristic (ROC) curves were constructed to examine the diagnostic test performance, i.e., its ability to discriminate between controls and patients with PCOS. All calculations were performed using the Statview program (Cricket Software, Inc., Philadelphia, PA) and JMP v 5 (SAS, Cary, NC).

Results

No difference was found in body weight and body mass index (BMI) between controls and patients with PCOS (Table 1).

<table>
<thead>
<tr>
<th>Group</th>
<th>Women with normal menstrual cycles</th>
<th>Patients with PCOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>50</td>
<td>49</td>
</tr>
<tr>
<td>Weight (kg) (mean ± 2SD)</td>
<td>53.5 ± 16.0</td>
<td>56.9 ± 27.4</td>
</tr>
<tr>
<td>BMI (kg/m²) (mean ± 2SD)</td>
<td>21.4 ± 5.8</td>
<td>23.1 ± 11.0</td>
</tr>
<tr>
<td>Obese subjects (%) (BMI≥25)</td>
<td>10.0</td>
<td>18.4</td>
</tr>
<tr>
<td>Hirsutism (%)</td>
<td>2.0</td>
<td>8.2</td>
</tr>
</tbody>
</table>
Measurement values of ARCHITECT T were highly correlated with those of LC-MS (r = 0.79; slope, 0.91, intercept, 47.2). The regression analysis showed a slope of 0.45 and an intercept of 62.3 ng/dL between ARCHITECT T and DPC total testosterone kit (Fig. 1).

As no differences were found in T levels among the three phases of the menstrual cycle in the control (data not shown), measurement values in the follicular phase were used for comparison with PCOS. The mean level of T in patients with PCOS was significantly higher than that of the control (86 ± 48 vs 68 ± 46 ng/dL, $P<0.01$) (Table 2); however, the distribution of the measured values of these two groups overlapped (Fig. 2). The prevalence rates of hyperandrogenism (T

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>mean ± 2SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women with normal menstrual cycles</td>
<td>50</td>
<td>68 ± 46</td>
</tr>
<tr>
<td>Patients with PCOS</td>
<td>49</td>
<td>86 ± 48a</td>
</tr>
</tbody>
</table>

Note. Total testosterone was measured by ARCHITECT T.

a) $P<0.01$ vs Women with normal menstrual cycles

Fig. 1. Comparison of LC-MS and ARCHITECT T. Serum samples obtained from patients with PCOS were used.

Fig. 2. Frequency distribution of measured value of total testosterone in women with normal menstrual cycles and patients with PCOS. Testosterone was measured by ARCHITECT T.
>114 ng/dL; mean +2SD in control) were 10.2% in patients with PCOS and 4% in controls. The diagnostic potency of T measured by ARCHITECT T was tested by the ROC procedure (Fig. 3). The area under the ROC curve (AURC) was 0.72. Several cut-off values of T were analyzed in terms of specificity and sensitivity from ROC curve data (Table 3).

### Discussion

In recent years, the diagnostic criteria of the 2003 Rotterdam ESHRE/ASRM-sponsored PCOS consensus workshop group have been widely accepted. Clinical and/or biochemical evidence of hyperandrogenism is included as one of the essential characteristics in these criteria, whereas hyperandrogenism is included as a referential but not as an essential factor in PCOS criteria of JSOG 1993 used widely in Japan. The criteria of JSOG are based on a nationwide questionnaire, and consist of chronic anovulation, presence of PCO, and high LH level as essential factors [4]. The establishment of these criteria was an important step toward standardizing the diagnosis of Japanese patients with PCOS, and these criteria are indeed expected to cover most Japanese patients with PCOS. However, some anovulatory patients who have typical clinical features of PCOS (i.e., chronic anovulation, presence of PCO, and hyperandrogenism) would not be diagnosed as PCOS using the criteria of JSOG 1993, due to lack of a high LH level, which is defined as essential for diagnosis. Accordingly, we need to re-evaluate these criteria in accordance with the actual clinical need, and particularly with emphasis on hyperandrogenism because hyperandrogenic anovulatory women should be diagnosed with PCOS despite their lack of a high LH level, as in international diagnostic criteria (NIH/NICHD 1990 or ESHRE/ASRM 2003). However, as mentioned above, biochemical hyperandrogenism has not been evaluated adequately. In this study, we evaluated total T for the diagnosis of Japanese patients with PCOS in order to assist the re-evaluation of the criteria of JSOG 1993.

As commercially available T assay kits do not require extraction steps, compounds present in serum, such as lipids, proteins, and binding globulin may interfere with the immunoassay [7]. First, we assessed the reliability of the measurement kit ARCHITECT T. The correlation coefficient between ARCHITECT T and LC-MS was 0.79, which is better than other kits evaluated in previous reports [8]. Another problem was between-kit variability. So, we compared measured values between ARCHITECT T and DPC total testosterone kit. Measured values obtained from ARCHITECT T were higher than those obtained from the DPC total testosterone kit at low concentration. Compared with previous reports, mean T level obtained from the controls was higher in ARCHITECT T [9]. Based on these findings, the information of the standard values of assay systems is important for assessment of T levels.

Although the mean T level was significantly higher in patients with PCOS than in control, the distribution of measured values greatly overlapped with those of controls. This indicates that there is no complete cutoff value that distinguishes between controls and pa-
patients with PCOS by T level alone. Then, we evaluated the sensitivity and specificity of each cut-off point using ROC analysis. ROC analysis has been used to determine the performance of several serum measurements in detecting PCOS [10, 11]. Koskinen et al. showed that T as well as LH, the LH/FSH ratio, and androstenedione had an AURC above 0.90 in Finnish patients with PCOS [10]. They also showed that the use of total T alone yielded a sensitivity and specificity of 70% and 97%, respectively, if 72 ng/dL was considered as the decision threshold. This indicated that T in itself would be useful for discriminating patients with PCOS and controls. In our results, AURC of T was 0.72, and there were no decision threshold that had both high sensitivity and high specificity. If we used the upper limit of the reference range (114 ng/dL) as the decision threshold, the sensitivity and specificity were 14% and 94%, respectively. Thus, T should not be used as an independent diagnostic factor of Japanese PCOS. It seems that the differences of race and diagnostic criteria of PCOS (diagnosis of PCOS was based on the presence of oligomenorrhea and PCO in a previous study) were the major reasons for these discrepancies. For example, clinical and biochemical hyperandrogenism are highly prevalent in Western patients [4–6]. This is true even in non-obese PCOS patients in Turkey [12]. Therefore T is useful for discriminating patients with PCOS and controls in Western people. On the other hand, the prevalence of hyperandrogenism is much less in Japanese patients. Therefore, we should not use T levels in diagnosis of Japanese patients with PCOS in the same way as Western countries. If we use T as an exclusion criterion for PCOS, namely, patients with low T levels would not be diagnosed with PCOS, the decision threshold should have high sensitivity, whereas if we use T as a complementary factor of other essential factors, the decision threshold should have high specificity. We should use T with the latter function. Among the three essential factors of the criteria of JSOG 1993, high LH level is most changeable because LH levels are influenced by endogeneous estrogen feedback, BMI, and between-kit variability [13–16]. In fact, 47% of PCOS patients showed normal LH levels at the point of blood sampling for T measurements in this study, although all patients had showed high LH levels at diagnosis. When we use the criteria of JSOG 1993 for diagnosis, it is a difficult problem that high LH level is not always detectable in patients. Because there was no significant difference of mean T levels between high LH group and normal LH group (data not shown), we did not distinguish these two groups. Clear separation of the patients with PCOS (high LH levels) and patients of PCOS suspicious (normal LH levels) is impossible because of variability of LH levels. Furthermore, it has been suggested that morphological PCO finding and biochemical hyperandrogenism but not high LH level are associated with insulin resistance [17]. T might be more useful than LH in obese patients with insulin resistance. Anyway, the patients with PCO finding, chronic anovulation, and hyperandrogenism should be diagnosed as PCOS irrespective of their LH levels. Consequently, it seems better for T to be used as a complementary factor of higher LH concentration in Japan (i.e., essential factor of ‘higher LH concentration’ is changed to ‘higher LH concentration and/or hyperandrogenism’). If T is used as a complementary factor of LH in future new criteria, the decision threshold of 110–120 ng/dL in ARCHITECT T, which yields sensitivity and specificity of 6–14% and 94–98%, respectively, might be appropriate for Japanese PCOS. The cut off value might also be appropriate in the upper limit of normal range of assay systems. Some patients who do not fulfill the present criteria of JSOG 1993 will be diagnosed with PCOS if the criteria are revised, with a false positive of 2–6%.

In conclusion, the sensitivity and specificity of T for the diagnosis of Japanese patients with PCOS varied considerably for each cut-off value. T might be used as a complementary factor of hyper LH concentration in order to diagnose patients with typical clinical features of PCOS but who do not fulfill the JSOG 1993 criteria.

Acknowledgement

We thank Abbott Japan Inc. for the measurements of T.
References