Head-To-Head Assessment of Diagnostic Performance of Testosterone Immunoassays in Patients With Polycystic Ovary Syndrome

Andreas N. Schüring,¹ Stefan Nolte,² Manfred Fobker,² Frank Kannenberg,²
and Jerzy-Roch Nofer^{2*}

¹UKM Kinderwunschzentrum, Department of Gynecology and Obstetrics, University Hospital Münster, Münster, Germany

²Center for Laboratory Medicine, University Hospital Münster, Münster, Germany

Background: Determination of plasma testosterone is critical for the proper diagnosis of polycystic ovary syndrome (PCOS), but the interpretation of biochemical tests is hampered by inadequate specificity and precision of available immunoassays. We here compared the diagnostic performance of three testosterone immunoassays (Advia Centaur, Immulite 2000 XPi, Cobas e411) in PCOS patients using receiver operator characteristics curve analysis. **Methods and Results:** Plasma levels of testosterone, androstenedione, dehydroepiandrosterone sulfate, 17-hydroxyprogesterone, estradiol, progesterone, steroid hormone binding globulin, luteinizing hormone, and follicular stimulating hormone were determined in 188 patients with PCOS and 202 controls. Free testosterone (fT) levels and free androgen

index (FAI) were calculated. Testosterone levels measured on Advia Centaur, Immulite 2000 XPi, and Cobas e411 showed clear linear relationship to each other. Testosterone measured with Advia Centaur showed discriminatory performance superior to Immulite 2000 XPi and Cobas e411. Calculation of fT or FAI improved the performance of Advia Centaur and Immulite 2000 XPi, which nevertheless performed better than Cobas e411. The performance of other parameters was inferior to that of testosterone, fT, and FAI. **Conclusion:** Present study documents striking differences between testosterone immunoassays with respect to their capacity to identify PCOS patients and favors the use of calculated parameters reflecting active testosterone in plasma. *J. Clin. Lab. Anal.* **30**:479–484, 2016. © 2015 Wiley Periodicals, Inc.

Key words: testosterone; polycystic ovary syndrome; receiver operating characteristic curves; immunoassay

INTRODUCTION

Polycystic ovary syndrome (PCOS) is a common ovarian pathology affecting up to 15–20% of women in reproductive age (1, 2). Despite the high prevalence, there is much controversy regarding proper diagnostics of PCOS because of its heterogeneous phenotypic presentation. Biochemical hyperandrogenemia represents a major component of PCOS and determination of androgen levels in plasma have been recommended by most guidelines (3, 4). However, clear-cut definition of androgen excess based on circulating hormone levels revealed to be difficult. The accurate measurement of testosterone is technically demanding and the availability of gold-standard methods involving organic extraction, chromatography,

and radioimmunoassay is limited to few reference laboratories. On the other hand, platform immunoassays—though rapid, automatable, and easy to perform—are characterized by inadequate specificity due to antibody cross-reaction with various steroids species and often

Grant sponsor: SIEMENS Health Care; Grant sponsor: Center for Laboratory Medicine.

*Correspondence to: Jerzy-Roch Nofer, Centrum für Laboratoriumsmedizin, Universitätsklinikum Münster, Albert Schweizer Campus 1, Gebäude A1, 48149 Münster, Germany. E-mail: nofer@uni-muenster.de

Received 17 June 2014; Accepted 27 July 2015

DOI 10.1002/jcla.21882

Published online in Wiley Online Library (wileyonlinelibrary.com).

suffer from poor performance in lower concentration ranges. These shortcomings may severely impede the accurate identification of PCOS patients and lead to false-positive or false-negative diagnosis depending on the immunoassay used for androgen determination.

To facilitate the interpretation of biochemical tests applied for the diagnosis of hyperandrogenemia, in the present study we compared three widespread platform immunoassays (Advia Centaur, Immulite 2000 XPi, Cobas e411) used for testosterone determination in large groups of PCOS patients and healthy subjects using receiver operator characteristics (ROC) curve analysis. Our results document considerable differences between immunoassays with respect to their capacity to identify PCOS patients and emphasize the necessity to use calculated parameters such as free testosterone (fT) or free androgen index (FAI) for the accurate diagnosis of this syndrome.

MATERIAL AND METHODS

Three hundred ninety female patients attending the Reproductive Medicine Unit (Kinderwunschzentrum) at the University Hospital of Münster between years 2007 and 2010 were examined in this study as described previously (5). Written informed consent was obtained from all patients following ethical board approval. Patients were evaluated regarding their clinical, endocrine, and sonographic parameters indicating the presence of PCOS. A PCOS group was defined according to the ESHRE/ASRM consensus and compared with a control group not fulfilling the PCOS criteria (The Rotterdam ESHRE/ASRM sponsored PCOS consensus workshop group 2004, (4). The following parameters were recorded: age, body mass index (BMI), menstrual history, cycle length, and the presence of clinical androgenization (hirsutism). After cessation of any hormonal medication for 2 months, an endocrine status encompassing serum estradiol, luteinizing hormone (LH), follicle stimulating hormone (FSH), testosterone, dihydroepiandrosterone sulfate (DHEAS), and steroid hormone binding globuline (SHBG) was assessed in the early follicular phase (day 3–5) of a spontaneous or progestin-induced cycle. Basal 17-hydroxyprogesterone (17-OHP), prolactin, and thyroid-stimulating hormone (TSH) were assessed to exclude other endocrine disorders causing anovulation. Nonclassical congenital adrenal hyperplasia was excluded by corticotropin stimulation test. The diameter of a dominant follicle and the presence of polycystic ovaries were evaluated with transvaginal scan on day 12–14. Patients with unclear sonographic records were excluded. Serum progesterone was determined on day 20–24 of a spontaneous or progestin-induced cycle.

After collecting blood, samples were immediately centrifuged and sera were analyzed for standard parameters

or remained frozen at -80°C until further analysis. Initial determination of laboratory parameters (testosterone, SHBG, DHEAS, LH, FSH, estradiol, progesterone) used for identifying biochemical hyperandrogenemia was performed as a part of routine clinical examination during sample collection phase using electrochemiluminescence immunoassays on the Modular e170 analyzer (Roche Diagnostics, Mannheim, Germany). The first-generation immunoassay (Testo Roche) was used for the determination of testosterone concentrations. To ensure the proper and unambiguous attribution of study subjects to the PCOS positive or negative group, we subsequently validated testosterone measurements in all reposit samples using mass spectroscopy as a “gold standard” method. To this purpose, sera were extracted with ethyl acetate and additionally cleaned up using solid-phase extraction on RP18 material. Steroids were derivatized by *O*-(2,3,4,5,6-pentafluorobenzyl)hydroxylamine hydrochloride and *N,O*-bis(trimethylsilyl)-acetamide. Analysis was performed using a Shimadzu QP2010 gas chromatography-mass spectroscopy (GC-MS) system with negative chemical ionization equipped with a Restek Rtx5MS column ($15\text{ m} \times 0.25\text{ mm} \times 0.1\text{ }\mu\text{m}$) at 160°C to 300°C ramped by $30^{\circ}\text{C}/\text{min}$. Three microliters were injected at 300°C . The measurements were calibrated using testosterone standard solutions and d_3 -testosterone as an internal standard in SIM mode using the specific values $m/z = 535$ and $m/z = 538$. The limits of detection and limits of quantitation were determined using MVA 2.0 software and amounted to 0.021 and 0.155 ng/ml, respectively (using rules of DIN 32645). Recovery of testosterone was assessed using spiked pool serum and ranged from 97.0% to 100.4% ($<10.0\text{ ng/ml}$). Interassay precisions were 4.3% at 1.7 ng/ml and 5.4% at 4.7 ng/ml. Samples with discrepant testosterone determination (positive with first-generation testosterone immunoassay and negative with mass spectrometry or vice versa) were excluded from subsequent analysis.

For the head-to-head performance comparison, testosterone levels were measured with three commercially available chemiluminescence immunoassays: Testosterone (TSTO, Siemens Healthcare, Eschborn, Germany), Testosterone (L2KTW2, Siemens), and Testo II (second-generation immunoassay, Roche) using following platforms: Advia Centaur, Immulite 2000 XPi, and Cobas e411, respectively. Androstendione levels were measured on Immulite 2000 XPi. FAI was calculated using the formula (testosterone/SHBG) $\times$ 1000. fT was calculated according to Vermeulen et al. (6). SHBG concentrations determined on Advia Centaur were used for fT and FAI calculations.

An exploratory statistics was performed using the MedCalc Statistical Software version 12.7.7 (MedCalc

TABLE 1. Clinical Characteristics and Endocrine Profile of PCOS Patients and Control Subjects

	Control (<i>n</i> = 202)	PCOS (<i>n</i> = 188)	Significance (<i>P</i>)
Height (m)	1.67 ± 0.06	1.69 ± 0.06	n.s.
Weight (kg)	64.8 ± 10.7	71.3 ± 15.8	<0.001
BMI	22.8 ± 3.5	24.8 ± 5.3	<0.001
Waist circumference (cm)	80.0 ± 9.5	84.2 ± 12.6	<0.001
Androstendione (nmol/l)	6.87 (6.51–7.36)	9.72 (8.60–10.26)	<0.001
DHEAS (nmol/l)	6078 (5648–6477)	8083 (7187–8733)	<0.001
Oestradiol (pmol/l)	139 (128–150)	150 (136–165)	n.s.
Progesterone (nmol/l)	52.9 (49.9–55.4)	32.3 (25.7–34.2)	<0.001
17-OHP (nmol/l)	1.91 (1.81–2.13)	2.32 (2.10–2.54)	<0.01
SHBG (nmol/l)	65.4 (61.2–71.0)	59.0 (52.1–65.8)	<0.01
LH (U/l)	6.31 (5.92–6.55)	6.15 (5.60–6.88)	n.s.
FSH (U/l)	7.25 (7.00–7.74)	6.10 (5.70–6.48)	<0.001
LH/FSH ratio	0.81 (0.76–0.85)	1.05 (0.91–1.13)	<0.001
Prolactin (ng/ml)	16.4 (4.8–16.8)	149 (13.8–16.7)	n.s.
TSH (mU/l)	2.18 (1.95–2.41)	2.06 (1.84–2.31)	n.s.
Advia Centaur			
Testosterone (nmol/l)	1.52 (1.38–1.63)	2.42 (2.18–2.49)	<0.001
Free testosterone (pmol/l)	13.1 (12.3–14.9)	21.4 (19.4–25.1)	<0.001
FAI (%)	1.69 (1.54–1.92)	2.87 (2.45–3.26)	<0.001
Immolute 2000 XPI			
Testosterone (nmol/l)	1.00 (0.91–1.08)	1.48 (1.40–1.71)	<0.001
Free testosterone (pmol/l)	10.7 (9.9–12.4)	18.4 (16.2–20.8)	<0.001
FAI (%)	1.47 (1.31–1.67)	2.62 (2.19–2.69)	<0.001
Cobas e411			
Testosterone (nmol/l)	1.21 (1.07–1.24)	1.35 (1.10–1.38)	<0.05
Free testosterone (pmol/l)	9.7 (9.7–11.5)	12.7 (10.9–13.6)	<0.05
FAI (%)	1.36 (1.29–1.45)	1.61 (1.38–1.83)	<0.01
GC-MS			
Testosterone (nmol/l)	1.11 (1.04–1.17)	1.43 (1.35–1.45)	<0.001
Free testosterone (pmol/l)	9.3 (8.8–10.2)	14.3 (13.1–16.1)	<0.001
FAI (%)	1.22 (1.10–1.34)	1.86 (1.67–2.09)	<0.001

Values represent mean ± SD.

n.s., not significant.

Software bvba, Ostend, Belgium). The distribution of continuous variables was assessed for normality using Shapiro–Wilk test. Student's *t*-test (two-sided) or Mann–Whitney *U*-test (two-sided) was used to compare groups with normally or nonnormally distributed data, respectively. Passing–Bablok regression analysis with Cusum test for linearity was used for method comparison (7). ROC curve analysis was performed according to Hanley and McNeail (7). The local significance level (*P*) was set to 0.05. For descriptive purposes, means (± SD) or medians (95% CI) for normally or nonnormally distributed variables, respectively, are provided.

RESULTS

Table 1 shows the main clinical characteristics and endocrine features of the examined groups. There was no difference with respect to height but PCOS patients presented with significantly increased weight, BMI, and waist circumference. Serum levels of total testosterone and fT were significantly higher in patients with PCOS regardless of the method used for determination. Moreover, PCOS patients were characterized by increased FAI. In addition, the PCOS group displayed significantly higher plasma levels of androgens (DHEAS, 17-OHP, and androstendione) and

TABLE 2. Comparison of Methods Used for Testosterone Determination by Passing–Bablok Regression Analysis

	Intercept	95% CI	Slope	95% CI	Significance ^a
GC-MS vs. Advia Centaur	−0.22	−0.31 to −0.15	2.12	1.92–2.36	0.071
GC-MS vs. Immolute 2500 XPI	−0.07	−0.14 to −0.05	1.30	1.14–1.48	n.s.
GC-MS vs. Cobas e411	−0.06	−0.11 to −0.01	1.11	1.00–1.29	n.s.

^aCusum's test for linearity.

Fig. 1. ROC plots of total testosterone (A), fT (B), and FAI (C) in 188 female patients with PCOS and 202 controls. Plasma levels of total testosterone were determined with Advia Centaur, Immulite 2000 XPi, and Cobas e411 as described in Material and Methods.

LH/FSH ratio and significantly lower plasma levels of SHBG, progesterone, and FSH. No differences between both groups were observed with respect to estradiol, prolactin, and TSH.

Table 2 demonstrates the relationship between testosterone results obtained with GC-MS and Advia Centaur, Immulite 2000 XPi, or Cobas e411 as assessed by Passing–Bablok regression analysis. While all three platform immunoassays used for testosterone determination showed clear linear relationship to results obtained with GC-MS, no significant deviations from linearity were seen in case of Immulite 2000 XPi and Cobas e411, whereas borderline significant deviation was noted for Advia Centaur.

ROC curve analysis was used to assess the diagnostic performance of testosterone, fT, and FAI to discriminate between patients with PCOS and control subjects. As shown in Figure 1 and Table 3, total testosterone measured with Advia Centaur showed discriminatory performance superior to Immulite 2000 XPi, which in

TABLE 3. ROC Analysis of Total and Free Testosterone Levels and FAI in Patients with PCOS and Control Subjects

	AUC	SE
Testosterone		
Advia Centaur	0.729	0.027
Immulite 2000 XPi	0.699	0.031
Cobas e411	0.564	0.027
fT		
Advia Centaur	0.738	0.027
Immulite 2000 XPi	0.735	0.026
Cobas e411	0.594	0.030
FAI		
Advia Centaur	0.736	0.027
Immulite 2000 XPi	0.733	0.027
Cobas e411	0.601	0.030

turn performed better than total testosterone determined with Cobas e411. Calculation of fT or FAI improved and equalized the performance of Advia Centaur and Immulite 2000 XPi, which nevertheless performed better than Cobas e411. The discriminatory performance of other steroid hormones including androstendione, as well as protein hormones including LH/FSH ratio, was inferior to that of total testosterone, fT, and FAI determined either with Advia Centaur or Immulite 2000 XPi with the area under curve (AUC) <0.6 for all parameters tested except DHEAS (AUC = 0.632) and SHBG (AUC = 0.621).

DISCUSSION

The objective of the present study was the head-to-head evaluation of diagnostic performance of three immunochemical tests widely used for the determination of testosterone concentrations in plasma. To avoid the recruitment bias, the investigation was performed in a consecutive group of women attending the Reproductive Medicine Unit and presenting with symptoms suggesting PCOS as a diagnosis. The control group was randomly selected from consecutive patients of the same unit showing no indications of PCOS. The diagnostic performance of examined methods was assessed using ROC analysis, which has several advantages for evaluation of laboratory tests including that it is not dependent on the disorder prevalence in the population and the decision threshold employed. To our knowledge, this is the first head-to-head comparison of immunoassays used for routine testosterone measurement in a large PCOS cohort.

The direct comparison of three platform immunoassays with GC-MS using Passing–Bablok regression analysis revealed the linear relationship between testosterone concentrations measured with Immulite 2000 XPi, Cobas e411, and GC-MS, whereas borderline significant deviation from linearity was observed in case of Advia

Centaur. While reasons accounting for this finding remain unclear, it is worth noticing that higher testosterone concentrations were determined across the entire analytical measurement range, which may be related to the more-pronounced cross-reactivity of the Advia Centaur assay with steroids present in the samples that are structurally similar to testosterone. Actually, several steroids with potential to interact with testosterone-specific antibodies, which might be elevated in PCOS, were identified previously, but their cross-reactivity in Advia Centaur assay has not been specifically addressed (8). Another possible explanation could be the interaction between the labeled testosterone derivative in the assay, the assay antibody, and SHBG. Such interaction has been previously observed to influence testosterone concentrations in direct immunoassays (9).

The ROC analysis revealed Advia Centaur assay for total testosterone to best discriminate between patients with PCOS and healthy controls followed by Immulite 2000 XPi and Cobas e411. The AUC calculated for Advia Centaur assay was comparable to values previously reported in the literature for other immunoassays and mass spectroscopic measurement protocols (10–14). Interestingly, the superior performance of this assay in comparison to Immulite 2000 XPi was no longer evident, when indirect measures such as fT or FAI that better reflect active testosterone levels in plasma were calculated. This finding is consistent with the notion that Advia Centaur assay is more susceptible to interference from other steroids and that testosterone value measured with this assay may integrate level alterations of other steroids also affected in PCOS. This might improve its capacity to discriminate between PCOS patients and healthy individuals, as patients with hyperandrogenism due to nontestosterone steroids might exhibit an abnormal result. Conversely, the improved specificity characterizing the second-generation Cobas e411 assay might paradoxically contribute to its less satisfactory performance, which was observed both in case of total testosterone and calculated parameters fT and FAI (8).

The present study documents that regardless of the assay used for testosterone determination, calculated parameters such as fT or FAI have better capacity than total testosterone to discriminate between PCOS patients and healthy individuals. This observation corroborates the results of previous studies, which have repeatedly showed that fT, FAI, or bioavailable testosterone are better suited for diagnosing PCOS (10–14). The superior diagnostic performance of these parameters is likely related to the fact that they give direct or indirect information about active testosterone, which is accessible to bind to the androgen receptor and to induce intracellular signaling cascade. By contrast, other androgens such as androstendione and DHEAS, the use of which has been

previously advocated for diagnosis of PCOS in some studies (10, 14), showed relatively poor performance as diagnostic markers for PCOS in the investigated cohort. Likewise, both LH levels and LH/FSH ratio were found to inadequately discriminate between PCOS patients and healthy individuals.

In conclusion, using a large study cohort the present study reveals considerable differences between testosterone immunoassays with respect to their capacity to identify PCOS patients and suggests that increasing assay specificity may not necessarily translate into improved diagnostic performance. The performance characteristics of a testosterone immunoassay should be taken into consideration when diagnosing biochemical hyperandrogenemia in subjects with suspected PCOS.

ACKNOWLEDGMENTS

The expert technical assistance of Elke Börger is acknowledged. This work was supported by Siemens Healthcare, Eschborn, Germany, and intramural resources of the Center for Laboratory Medicine (support to J.-R.N.).

REFERENCES

1. March WA, Moore VM, Willson KJ, Phillips DI, Norman RJ, Davies MJ. The prevalence of polycystic ovary syndrome in a community sample assessed under contrasting diagnostic criteria. *Hum Reprod* 2010;25:544–551.
2. Lauritsen MP, Bentzen JG, Pinborg A, et al. The prevalence of polycystic ovary syndrome in a normal population according to the Rotterdam criteria versus revised criteria including anti-Müllerian hormone. *Hum Reprod* 2014;29:791–801.
3. Azziz R, Carmina E, Dewailly D, et al. Criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: An androgen excess society guideline. *J Clin Endocrinol Metab* 2006;91:4237–4245.
4. Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Hum Reprod* 2004;19:41–47.
5. Schüring AN, Welp A, Gromoll J, et al. Role of the CAG repeat polymorphism of the androgen receptor gene in polycystic ovary syndrome (PCOS). *Exp Clin Endocrinol Diabetes* 2012;120:73–79.
6. Vermeulen A, Verdonck L, Kaufman JM. The apparent free testosterone concentration, an index of androgenicity. *J Clin Endocrinol Metab* 1999;84:3666–3672.
7. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983;148:839–843.
8. Owen WE, Rawlins ML, Roberts WL. Selected performance characteristics of the Roche Elecsys testosterone II assay on the Modular analytics E 170 analyzer. *Clin Chim Acta* 2010;411:1073–1079.
9. Taieb J, Mathian B, Millot F, et al. Testosterone measured by 10 immunoassays and by isotope-dilution gas chromatography-mass

- spectrometry in sera from 116 men, women, and children. *Clin Chem* 2003;49:1381–1395.
10. Escobar-Morreale HF, Asunción M, Calvo RM, Sancho J, San Millán JL. Receiver operating characteristic analysis of the performance of basal serum hormone profiles for the diagnosis of polycystic ovary syndrome in epidemiological studies. *Eur J Endocrinol* 2001;145:619–624.
11. Hahn S, Kuehnel W, Tan S, et al. Diagnostic value of calculated testosterone indices in the assessment of polycystic ovary syndrome. *Clin Chem Lab Med* 2007;45:202–207.
12. Iwasa T, Matsuzaki T, Minakuchi M, et al. Diagnostic performance of serum total testosterone for Japanese patients with polycystic ovary syndrome. *Endocr J* 2007;54:233–238.
13. Barth JH, Field HP, Yasmin E, Balen AH. Defining hyperandrogenism in polycystic ovary syndrome: Measurement of testosterone and androstenedione by liquid chromatography-tandem mass spectrometry and analysis by receiver operator characteristic plots. *Eur J Endocrinol* 2010;162:611–615.
14. Villarroel C, Trejo L, Muñoz A, Kohen P, Fuentes A, Devoto L. Assessment of diagnostic competence of plasmatic androgens on polycystic ovary syndrome based on receiver operator characteristic curves. *Gynecol Endocrinol* 2010;26:600–606.