Clin. Lab. 2019;65:XXX-XXX

ORIGINAL ARTICLE

Comparison of Three Immunoassays Systems for Determining Serum Estradiol

Hongyu Zhang, Yiyong Huang, Ka Liang

Department of Clinical Laboratory, The Third Affiliated Hospital of Guangxi University of Chinese Medicine, Liuzhou Traditional Chinese Medical Hospital, The Third Clinical Faculty of Guangxi University of Chinese Medicine, Guangxi, China

SUMMARY

Background: **Estrogens are responsible for the development of the secondary female gender characteristics. Serum estradiol concentration measurements depend widely on automated immunoassay systems such as the Architect i2000SR, Cobas e601, and UniCel DxI 800. Here, we compare the performance of three automated immunoassay systems and investigate the correlation of serum estradiol levels.**
Methods: **The precision of the three automated immunoassay systems were evaluated by the Bio-Rad Laboratories Quality Control products according to the EP15-A2 document. Measurements of 81 serum samples from routine clinical specimens were completed on the Architect i2000SR, Cobas e601, and UniCel DxI 800 on the same day in our laboratory. Method comparison was performed using Bland-Altman and Passing-Bablok analyses. The correlation between the three detection systems was analyzed using the concordance correlation coefficient r.**
Results: **With the Bio-Rad Laboratories Quality Control products, we clearly demonstrated that the total inaccuracy of the Architect i2000SR detection systems is in the range of 1.61 - 5.64, the Roche Cobas e601 inaccuracy is in the range of 1.88 - 6.56, and the inaccuracy of the UniCel DxI 800 is in the range of 4.69 - 8.33. A Bland-Altman plot showed that serum estradiol concentrations determined by the Architect i2000SR were about 0.69 times those of the Roche Cobas e601, the Architect i2000SR concentrations were about 0.65 times those of the UniCel DxI 800, and the Roche Cobas e601 concentrations were about 0.96 times those of the UniCel DxI 800. The correlation coefficient r was 0.9936, 0.9857, and 0.9774 for the Architect i2000SR versus the Roche Cobas e601, the Architect i2000SR versus the UniCel DxI 800, and the Roche Cobas e601 versus the UniCel DxI 800, respectively.**
Conclusions: **The Architect i2000SR system had better precision. The three detection systems have good correlation with each other, but there is a large gap between the results of their detection of estradiol.**
(Clin. Lab. 2019;65:xx-xx. DOI: 10.7754/Clin.Lab.2018.180903)

Correspondence:
Hongyu Zhang
Department of Clinical Laboratory
The Third Affiliated Hospital of
Guangxi University of Chinese Medicine
Liuzhou Traditional Chinese Medical Hospital
The Third Clinical Faculty of
Guangxi University of Chinese Medicine
6 Honghu Road
Liuzhou 545000, Guangxi
China
Email: 360328467@qq.com

Manuscript accepted September 26, 2018

KEY WORDS

immunoassay, method comparison, serum estradiol

INTRODUCTION

As well-known, 17β-estradiol (E2) is a gender hormone responsible for female sexual maturation and maintenance of secondary sexual characteristics and reproductive functions. It is reported in the literature that estradiol is the key regulator of bone metabolism in both men and women and is also associated with inflammatory biomarkers, cancer, cardiovascular disease risk, and diabetes mellitus [1-5]. E2, in combination with

progesterone, pituitary prolactin, follicle stimulating hormone, luteinizing hormone, and testosterone, is widely used in the evaluation of endocrine functions and also in infertility, precocious puberty, menstrual disorders, spontaneous abortion, and more. In addition, E2, combined with progesterone and chorionic gonadotropin, is used for predicting the outcomes of *in vitro* fertilization in controlled ovarian hyperstimulation [6]. Therefore, accurate measurement of serum estradiol levels is crucial for the correct diagnosis, treatment, and prevention of disease and for embryo transfer outcome predictions.

Different methods for measuring E2 are available, such as capillary electrophoresis, liquid chromatography-electrospray tandem mass spectrometry, molecularly imprinted polymer grafted paper-based method, and gold nanoparticle-based fluorescence immunoassay [7-10].

However, chemiluminescence is widely used to detect estradiol in clinical practice. All of the methods are sandwich immunoassays, wherein one antibody called a capture antibody is used to bind estradiol in serum, another type of monoclonal antibody called a labeled antibody is also used in combination with estradiol in serum, and thus an immune complex of antibody-estradiol-labeled monoclonal antibody is formed. After further incubation and washing, a luminescent substrate solution is added to the reaction mixture. The resulting chemiluminescent reaction is measured as relative light units. There is a relationship between the amount of estradiol in the sample and the relative light units detected by the optics. Serum estradiol concentrations were calculated by the calibration curve.

Recently, there have been reports in the literature that the results of the same laboratory project detected by various detection systems exhibit large differences in measurements such as for serum human chorionic gonadotropin-beta, serum vitamin B12, and total 25-hydroxyvitamin D [11-13]. This raised the question whether detection results of estradiol are also different when measured in different detection systems.

In this study, we aimed to evaluate the analytical performance of three systems and compare the estradiol detection results obtained from these three fully automated chemiluminescent assays.

MATERIALS AND METHODS

Specimens

During the period from November 5, 2017, to December 9, 2017, we collected 81 samples from serum estradiol assay requests in the gynecological clinic of our hospital. The samples with different concentrations of serum estradiol were selected from clinical laboratory samples and lipid blood samples. Jaundice specimens and chylomic specimens were not included. All samples were centrifuged at 2,000 g for 10 minutes, then each serum sample was divided into three aliquots, and each aliquot was kept frozen at -80°C until they were analyzed. All studies using samples from human subjects obtained signed informed consent forms and were approved by the ethics committee of our hospital. The precision of the three instruments was validated with two levels of Bio-Rad Quality Control (Lyphochek® Immunoassay Plus Control).

Detection systems

Serum estradiol analysis was carried out using three different automated immunoassays: UniCel DxI 800 (Beckman Coulter, Brea, CA, USA), Architect i2000SR (Abbott Laboratories, Abbott Park, IL, USA), and Cobas e601 (Roche Diagnostics, Mannheim, Germany). The detection systems are closed systems, and all the reagents used in the instruments are original reagents. All instruments were operated in accordance with the standardized operating procedures provided by the instrument manufacturer. Chemiluminescence detection was generated by enzymatic reaction in the UniCel DxI 800, while the Architect i2000SR used chemiluminescence detection and the Roche Cobas e601 used an electrochemical luminescence detection system.

Performance characteristics of the three detection systems for estradiol measurement are displayed in Table 1.

Comparison design

The precision and accuracy of the above three instruments were verified by EPl5-A2. The precision was evaluated with two levels of Bio-Rad Quality Control products repeated four times a day for five days. Imprecision of batches were calculated along with the total imprecision, and these were compared with the within-batch precision and total imprecision provided by the instrument manufacturers. The estradiol detection of each specimen was carried out by the three instruments at the same time after the instruments were calibrated, and each instrument used a separate aliquot for recording estradiol test results.

Statistical analyses

Bland-Altman analysis was performed to determine the mean difference in serum estradiol results with 95% consistency between the three detection systems. Through the Bland-Altman analysis, we can understand whether there are differences between the three systems, and how many gaps exist. Different detection methods should not be considered interchangeable if these gaps are not accepted by the clinician. The Passing-Bablok regression analysis and concordance correlation analysis were used to evaluate the relationship among the three instrument methods. The r coefficients of ≤ 0.35 were considered to represent low or weak correlations; 0.36 - 0.67 represented modest or moderate correlations; 0.68 - 1.0 represented strong or high correlations; and with r coefficients of ≥ 0.90, very high correlations [12]. MedCalc Software (version 13.3.1, MedCalc Software, Mariakerke, Belgium) and IBM SPSS Statistics

Table 1. Summary of performance features according to information provided in the estradiol reagent instructions.

Performance features	UniCel DxI 800	Architect i2000sr	Roche Cobas E601
Detection principle	paramagnetic particle chemiluminescent immunoassay	chemiluminescent microparticle immunoassay	electrochemiluminescence immunoassay
Analytical sensitivity	20 pg/mL	10 pg/mL	5 pg/mL
Linearity	20 - 4,800 pg/mL	10 - 1,000 pg/mL	5 - 3,000 pg/mL
Precision, %	≤ 12 - 21	1.4 - 6.4	1.1 - 6.7
Reference range	follicular phase: 27 - 122 pg/mL ovulation phase: 95 - 433 pg/mL luteal phase: 49 - 291 pg/mL menopause: 20 - 40 pg/mL males: 20 - 47 pg/mL	follicular phase: 21 – 251 pg/mL ovulation phase: 38 - 649 pg/mL luteal phase: 21 - 312 pg/mL menopause: 10 - 28 pg/mL males: 11 - 44 pg/mL	follicular phase: 12.5 - 166 pg/mL ovulation phase: 85.8 - 498 pg/mL luteal phase: 43.8 - 211 pg/mL menopause: 5 - 54 pg/mL males: 7.63 - 42.6 pg/mL

Table 2. Imprecision summary for estradiol assays.

Method	Quality control level	Mean concentration (pg/mL)	% CV	
			Within-batch precision	Total imprecision
UniCel DxI 800	Level 1	132.86	7.25	8.33
	Level 2	695.56	4.69	6.25
Architect i2000sr	Level 1	95.64	3.52	5.64
	Level 2	344.36	1.61	2.24
Roche Cobas E601	Level 1	112.36	4.71	6.56
	Level 2	508.25	1.88	2.72

Table 3. Summary of serum estradiol concentrations in 81 samples (pg/mL).

	UniCel DxI 800	Architect i2000sr	Roche Cobas E601
Mean	838.9	562.4	798.4
Median	322	214	321
Concentration range	21 - 4,597	12 - 2,923	20 - 4,177

for Windows, Version 22.0. (IBM Corp.; Armonk, NY, USA) were used for statistical calculations. All tests were two-sided, and $p < 0.05$ was considered statistically significant.

RESULTS

Validation of precision

According to EPl5-A2 documents, precision was evaluated with two levels of Bio-Rad Quality Control products (Lyphochek® Immunoassay Plus Control 371, LOT:40331; and Lyphochek® Immunoassay Plus Control 373, LOT:40333), repeated four times a day for five days, calculating the batch imprecision and the total imprecision. The within-batch precision and total imprecision of the three instruments were less than those provided by the instrument manufacturers. The Architect i2000SR had the least imprecision, and the UniCel DxI 800 has the greatest imprecision. The results of the comparison of precision are shown in Table 2.

The general distribution of the comparison of specimens

Each aliquot of the same serum sample was detected by three instruments. A summary of serum estradiol concentrations detected by the three instruments in 81 samples is shown in Table 3.

Table 4. Passing-Bablok regression analysis and concordance correlation analysis of the three detection systems.

Method	Passing-Bablok regression analysis				Concordance correlation analysis			
	Intercept A	95% CI *	Slope B	95% CI *	CCC **	95% CI *	r ***	Cb ****
Architect i2000sr vs. Roche Cobas E601	-3.8785	-6.5918 to -0.5014	0.7066	0.6743 to 0.7347	0.9013	0.8753 to 0.9221	0.9936	0.9071
Architect i2000sr vs. UniCel DxI 800	-5.6834	-10.2954 to -0.0720	0.6887	0.6638 to 0.7146	0.8740	0.8395 to 0.9015	0.9857	0.8867
Roche Cobas E601 vs. UniCel DxI 800	-3.1141	-8.2359 to 2.5616	0.9538	0.9256 to 0.9826	0.9761	0.9633 to 0.9845	0.9774	0.9987

95% CI *: 95% confidence interval. CCC **: concordance correlation coefficient. r *: Pearson's correlation coefficient . Cb ****: bias correction factor.**

Bland-Altman analysis

The plot ratio was selected so that the ratios of the measurements would be plotted instead of the differences. In the Bland-Altman plot, serum estradiol concentrations determined by the Architect i2000SR were about 0.69 times that of the Roche Cobas e601 (Figure 1A). At the same time, serum estradiol concentrations determined by the Architect i2000SR were about 0.65 times that of the UniCel DxI 800 (Figure 1C). In addition, serum estradiol concentrations determined by the Roche Cobas e601 were about 0.96 times that of the UniCel DxI 800 (Figure 1E). The differences between these detection systems are considered clinically significant.

Passing-Bablok analysis and concordance correlation analysis

Passing-Bablok regression analysis showed that the regression equation for the Architect i2000SR versus the Roche Cobas e601 was y = -3.8785 + 0.7066 x (Figure 1B); for the Architect i2000SR versus the UniCel DxI 800, it was y = -5.6834 + 0.6887 x (Figure 1D); and for the Roche Cobas e601 versus the UniCel DxI 800, it was y = -3.1141 + 0.9538 x (Figure 1F). Concordance correlation analysis showed that they had good correlation with each other. A specific summary of the Passing-Bablok regression analysis and concordance correlation analysis of the three detection systems is given in Table 4.

DISCUSSION

Estradiol (17β-estradiol (1,3,5(10)-estratrien-3,17β-diol)) is a natural estrogen with a molecular mass of 272.3 daltons that has a variety of forms. It is estimated that only 1 - 3% of estradiol is free (unbound), with about 98% of estradiol bound to transport proteins (SHBG = gender hormone binding globulin). Estradiol is an important hormone index for evaluating ovarian function and also plays an important role in the outcome of embryo transfer. There are a number of instrument manufacturers that use a variety of detection methods to quantitate estradiol levels in serum.

Because the Beckman UniCel DxI 800, Abbott Architect i2000SR, and Roche Cobas e601 systems are the most widely used in clinical practice for detection of estradiol, it was necessary to assess the analytical performance of the three detection systems and compare the consistency of estradiol results. All three detection systems had good analytical performance, but the Architect i2000SR had the best precision, and the Beckman UniCel DxI 800 had the worst precision of the three detection systems. The three detection systems were well correlated with each other for detecting estradiol: the correlation coefficient r of estradiol between the Architect i2000SR and the Roche Cobas e601 system was 0.9936 (0.9900, 0.9959), while the Architect i2000SR versus the UniCel DxI 800 system was 0.9857 (0.9779, 0.9908) and the Roche Cobas e601 versus the UniCel DxI 800 was 0.9774 (0.9650, 0.9854). Although the correlation between the three detection systems was good, there was a significant difference when testing the same sample. The Bland-Altman plot analysis showed that the concentrations of serum estradiol determined by the Architect i2000SR was about 0.69 times that of the Roche Cobas e601, the Architect i2000SR concentrations was about 0.65 times that of the UniCel DxI 800, and the Roche Cobas e601 concentrations was about 0.96 times that of the UniCel DxI 800.

This clinically significant difference deserves our deep consideration. At present, although steroid hormone measurement techniques have been significantly im-

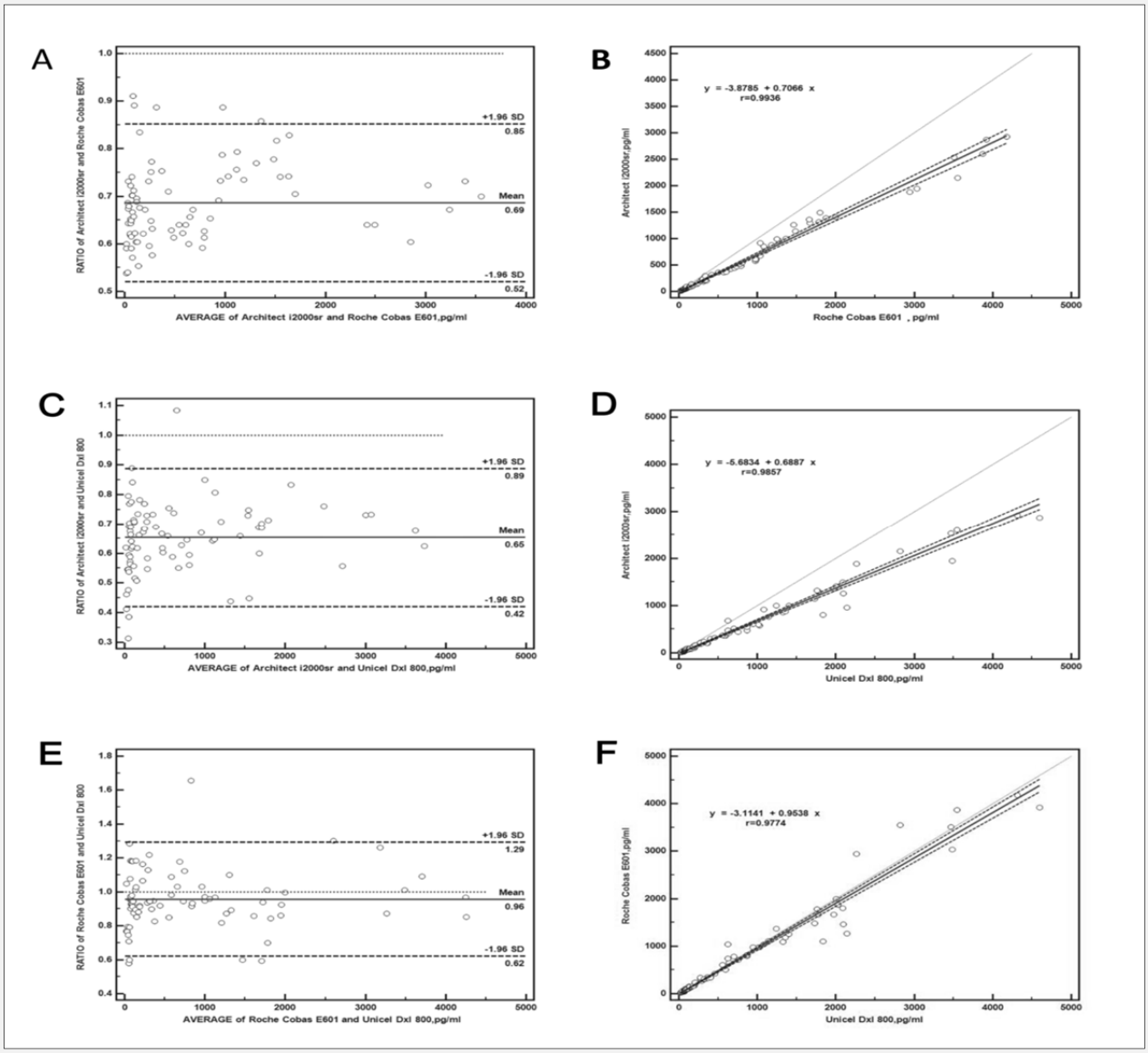

Figure 1. Comparison of all immunoassays against the mean serum estradiol by Passing-Bablok regression analysis (left panels) and Bland-Altman plots (right panels).

proved, estradiol measurement does not always achieve a high degree of accuracy and precision, especially at lower analyte concentrations, and intra-assay variability is generally higher. Ultimately, a combination of factors appears to cause inaccuracy of steroid hormone measurements, with nonuniform assay calibration and lack of specificity being two major contributors to assay variability [14]. Therefore, the lack of accurate and exogenous substance interference may be an enormous challenge for serum estradiol detection.

Some reasons can explain this discrepancy. First, the traceability of calibrators for each detection system is different. The Roche Cobas e601, with electrochemiluminescence immunoassay, has been standardized against CRM 6004a via ID-GC/MS (isotope dilution-gas chromatography/mass spectrometry). The Architect i2000SR, using chemiluminescent microparticle immunoassay, can be traced back to internal reference standard (-estradiol). The UniCel DxI 800, using paramagnetic particle chemiluminescent immunoassay, can be traced back to the ID-GC/MS Estradiol Reference Method. The value of calibration passing from the standard material to the manufacturer, and ultimately to the user, means there may be some deviation in the value of

the transfer process. Second, exogenous substances interfere with the test results. Serum is not a pure substance; there are many interfering substances that may affect the detection of estradiol. It is reported in the literature that human anti-animal antibodies (HAAA) and other substances will interfere with the reaction process [15]. There are related articles reporting that heterophilic antibodies and rheumatoid factor in human serum can react with reagent immunoglobulins, interfering with *in vitro* immunoassays [15,16].

There are some limitations in our research. First, the relatively small sample size is a deficiency of this study. Only 83 samples were used to compare the differences between the three detection systems. Second, only one medical inspection center was included in this study. We did not do combined research with the medical centers around us. In addition, we only compared the differences between the three detection systems for the detection of estradiol and did not extend it to other detection systems and testing programs. Despite these shortcomings, we also found that there are some differences between the three detection systems in the detection of estradiol. Multicenter joint studies and large sample data comparisons may be needed in the future.

CONCLUSION

In summary, the three detection systems have good correlation, but there is a large difference between them in the detection of serum estradiol concentrations. We hope that large sample, multicenter studies will be published in the future to reinforce our conclusion. We are even looking forward to the instrument manufacturers reaching agreement to ensure the consistency and standardization of test items.

Declaration of Interest:

The authors declare no conflicts of interest with respect to the authorship and/or publication of this article.

References:

1. Cauley JA. Estrogen and bone health in men and women. Steroids 2015;99(Pt A):11-15 (PMID: 25555470).

2. Tsilidis KK, Rohrmann S, McGlynn KA, et al. Association between endogenous sex steroid hormones and inflammatory biomarkers in US men. Andrology 2013;1(6):919-928 (PMID: 24124163).

3. Cohn BA, Cirillo PM, Hopper BR, Siiteri PK. Third Trimester Estrogens and Maternal Breast Cancer: Prospective Evidence. J Clin Endocrinol Metab 2017;102(10):3739-48 (PMID: 28973345).

4. Mongraw-Chaffin ML, Anderson CA, Allison MA, et al. Association between sex hormones and adiposity: qualitative differences in women and men in the multi-ethnic study of atherosclerosis. J Clin Endocrinol Metab 2015;100(4):E596-600 (PMID: 25636047).

5. Wang X, Xian T, Jia X, et al. A cross-sectional study on the associations of insulin resistance with sex hormone, abnormal lipid metabolism in T2DM and IGT patients. Medicine (Baltimore) 2017;96(26):e7378 (PMID: 28658166).

6. Wei M, Zhang XM, Gu FL, et al. The impact of LH, E2, and P level of HCG administration day on outcomes of *in vitro* fertilization in controlled ovarian hyperstimulation. Clin Exp Obstet Gynecol 2015;42(3):361-6 (PMID: 26152012).

7. Du L, Ji W, Zhang Y, et al. An ultrasensitive detection of 17beta-estradiol using a gold nanoparticle-based fluorescence immunoassay. Analyst 2015;140(6):2001-7 (PMID: 25672478).

8. Xiao L, Zhang Z, Wu C, Han L, Zhang H. Molecularly imprinted polymer grafted paper-based method for the detection of 17beta-estradiol. Food Chem 2017;221:82-86 (PMID: 27979279).

9. Lee C, Kim CH, Kim S, Cho SH. Simultaneous determination of bisphenol A and estrogens in hair samples by liquid chromatography-electrospray tandem mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 2017;1058:8-13 (PMID: 28521190).

10. Siren H, El Fellah S. Androgens, oestrogens, and progesterone concentrations in wastewater purification processes measured with capillary electrophoresis. Environ Sci Pollut Res Int 2017;24(20):16765-85 (PMID: 28567679).

11. Guan X, Sun Y, Zhang H, et al. Comparison of Architect i2000sr and Cobas e601 Systems for Determining Serum Human Chorionic Gonadotropin-Beta. Clin Lab 2016;62(9):1815-20 (PMID: 28164575).

12. Ispir E, Serdar MA, Ozgurtas T, et al. Comparison of four automated serum vitamin B12 assays. Clin Chem Lab Med 2015;53(8):1205-13 (PMID: 25720078).

13. Kocak FE, Ozturk B, Isiklar OO, Genc O, Unlu A, Altuntas I. A comparison between two different automated total 25-hydroxy-vitamin D immunoassay methods using liquid chromatography-tandem mass spectrometry. Biochem Med (Zagreb) 2015;25(3):430-8 (PMID: 26526462).

14. Vesper HW, Botelho JC, Wang Y. Challenges and improvements in testosterone and estradiol testing. Asian J Androl 2014;16(2):178-84 (PMID: 24407184).

15. Mebes I, Graf M, Kellner M, et al. High Estradiol Levels During Postmenopause - Pitfalls in Laboratory Analysis. Geburtshilfe Frauenheilkd 2015;75(9):941-4 (PMID: 26500371).

16. Holm BE, Sandhu N, Tronstrom J, et al. Species cross-reactivity of rheumatoid factors and implications for immunoassays. Scand J Clin Lab Invest 2015;75(1):51-63 (PMID: 25347362).