Performance Characteristics of Eight Estradiol Immunoassays

David T. Yang, MD,[1] William E. Owen, MT(ASCP),[2] Carol S. Ramsay,[3] Hui Xie, MS,[3] and William L. Roberts, MD, PhD,[1,2]

Key Words: Gas chromatography; Mass spectrometry; Accuracy; Imprecision

DOI: 10.1309/5N2R4HT4GM0AGPBY

Abstract

Measurement of estradiol is useful in assisted reproduction, evaluation of infertility, menopause, and male feminization. The analytic performance of 8 estradiol immunoassays was evaluated. The imprecision and accuracy of the Access, ADVIA Centaur, ARCHITECT i2000, AutoDELFIA, Elecsys 2010, IMMULITE 2000, and Vitros ECi estradiol assays (see text for proprietary information) were evaluated by using an isotope dilution–gas chromatography–mass spectrometry (ID-GC-MS) reference method.

The coefficient of variation (CV) ranged from 6.9% on the Elecsys 2010 to 42.6% on the ADVIA Centaur at an estradiol concentration of 18 pg/mL (66 pmol/L), with the ARCHITECT i2000 assay in development and the Vitros ECi having a CV below 10% at this estradiol concentration. Agreement between the automated assays and ID-GC-MS was variable, with slopes ranging from 0.87 to 1.20. The Access, ARCHITECT i2000 in development, and the IMMULITE 2000 were the most accurate, with slopes of 0.99, 0.98, and 1.03, respectively. These findings indicate that the ARCHITECT i2000 estradiol assay in development had the best precision and accuracy of the assays evaluated for measurement of serum estradiol concentrations.

Estradiol, a steroid hormone, is the most potent of the naturally occurring estrogens, and in the nonpregnant female, it is derived almost exclusively from the ovaries. Estradiol is responsible for the development and maintenance of the female sex organs and secondary sex characteristics. It also has important roles in regulation of the menstrual cycle and in maintenance of pregnancy. Estradiol measurements frequently are used during evaluation of female infertility and menopause and feminization in males and in assisted reproduction to monitor follicular development.[1] Physiologic levels are low, often less than 50 pg/mL (184 pmol/L), in men, children, and postmenopausal women and can range widely, from 5 to 410 pg/mL (18-1,505 pmol/L), in menstruating women.[2] The wide reference intervals and the clinical applications of this assay dictate the need for high sensitivity, excellent accuracy and precision through a wide analytical range, and a prompt turnaround time. Concern regarding these requirements, accuracy in particular, led to efforts toward standardization of serum estradiol measurements through a reference serum panel with estradiol concentrations determined by isotope dilution–gas chromatography–mass spectrometry (ID-GC-MS).[3,4]

Recently, estradiol assays have become available on a number of fully automated immunoassay analyzers. While these systems provide rapid turnaround times and high throughput, there presently are only limited reports regarding the analytic performance of these assays, and most do not assess accuracy with ID-GC-MS as the reference method.[5-8] The aim of our study was to assess the imprecision and accuracy of 8 automated estradiol assays with ID-GC-MS as the reference method.

Materials and Methods

The following automated methods and instruments were evaluated: the Access analyzer (Beckman Coulter, Brea, CA), the ADVIA Centaur analyzer (Bayer Diagnostics, Tarrytown, NY), the ARCHITECT i2000 analyzer using current reagents and reagents for a method that is in development (Abbott Diagnostics, Abbott Park, IL), the AutoDELFIA automatic immunoassay system (Wallac Oy, Turku, Finland), the Elecsys 2010 analyzer (Roche Diagnostics, Indianapolis, IN), the IMMULITE 2000 analyzer (Diagnostics Products, Los Angeles, CA), and the Vitros ECi analyzer (Ortho-Clinical Diagnostics, Raritan, NJ). All of these immunoassay methods use a competitive format. All methods use rabbit antiestradiol antibodies except for the Vitros ECi assay, which uses both rabbit and sheep antiestradiol antibodies. The ARCHITECT i2000 in development assay requires 150 µL of sample compared with 100 µL for the current assay and has a comparable resistance to interference by hemolysis, icterus, and lipemia. The Elecsys 2010 assay adds mesterolone to release estradiol from the sample.

Assay imprecision was assessed by assaying aliquots of 5 serum pools with estradiol concentrations of 18, 31, 93, 387, and 846 pg/mL (66, 114, 341, 1,420, and 3,105 pmol/L, respectively) in 6 replicates each on 2 different days for a total of 12 measurements. The estradiol concentration of these samples was determined by ID-GC-MS as previously described.[9] *Functional sensitivity,* defined as the estradiol concentration that would produce a coefficient of variation (CV) of 10%, was determined from a power curve fit of the CV plotted against the estradiol concentration determined by ID-GC-MS.

A method comparison study between each of the 8 estradiol assays and the ID-GC-MS reference method was performed using 61 samples with estradiol concentrations ranging from 0 to 1,101 pg/mL (0-4,041 pmol/L). EP Evaluator Release 5 software (David G. Rhoads Associates, Kennett Square, PA) was used for simple precision calculations and Deming regression analysis. For Deming regression analysis, the same SD was used for each method.

Results

A summary of data from the imprecision study is shown in **Table 1**. CVs ranged from 1.2% to 42.6% for estradiol concentrations from 18 to 846 pg/mL (66-3,105 pmol/L). At 18 pg/mL (66 pmol/L), the range of CV observed for each method was broad, ranging from 6.9% on the Elecsys 2010 to 42.6% on the ADVIA Centaur. At 93 pg/mL, CVs became more uniform and ranged from 2.5% on the ARCHITECT i2000 in development to 9.4% on the IMMULITE 2000. The functional sensitivities for a cutoff CV of 10% were determined from the data and are listed in **Table 2**. The ARCHITECT i2000 in development, Elecsys 2010, and Vitros ECi are capable of determining estradiol concentrations as low as 18 pg/mL (66 pmol/L) while maintaining CVs of less than 10%. The remaining instruments had functional sensitivities ranging from 35 to 80 pg/mL (128-294 pmol/L). It is noteworthy that the mean estradiol concentrations obtained by some immunoassay methods for the sample with the lowest estradiol concentration were substantially higher than the ID-GC-MS reference method.

Comparison of results obtained for samples collected from individual subjects using the 8 automated estradiol immunoassays and the reference ID-GC-MS method by Deming regression analysis is shown **Figure 1**. There were varying degrees of assay accuracy apparent, with slopes ranging from 0.87 to 1.20. The Access, ARCHITECT i2000 in development, and IMMULITE 2000 assays had the best accuracy with slopes of 0.99, 0.98, and 1.03 respectively. The AutoDELFIA and Elecsys 2010 assays showed the poorest accuracy with slopes of 0.87 and 1.20, respectively. Difference plots for the same data are shown **Figure 2**. The dispersion of data around the regression line, as reflected in the standard error of the estimate ($S_{y/x}$) and the correlation coefficient (r) and represented graphically in the difference plot, showed that the ARCHITECT i2000 assay in development had the best agreement with the reference method with an $S_{y/x}$ of 23 pg/mL (84 pmol/L), $r = 0.997$, followed by the Elecsys 2010 method with an $S_{y/x}$ of 28 pg/mL (103 pmol/L) and $r = 0.997$. The ADVIA Centaur had the largest amount of dispersion with an $S_{y/x}$ of 66 pg/mL (242 pmol/L) and $r = 0.980$.

Discussion

Our protocol for assessing imprecision was abbreviated owing to limited quantities of serum assayed by the ID-GC-MS reference method. Therefore, the true overall imprecision of the estradiol immunoassays during multiple days of analysis and multiple reagent lots is likely to be substantially higher than our data would indicate. Nevertheless, our results reflect the relative imprecision of each method.

Of the 8 estradiol immunoassays evaluated, the ARCHITECT i2000 assay in development, Elecsys 2010, and Vitros ECi methods were the most precise at low estradiol concentrations, being able to quantify estradiol concentrations as low as 18 pg/mL (66 pmol/L) while maintaining CVs of less than 10%. The Access, ARCHITECT i2000 assay in development, and IMMULITE 2000 had the best average results compared with the reference method, with

DOI: 10.1309/5N2R4HT4GM0AGPBY

Table 1
Summary of Imprecision Data*

Method†	Serum Pool	Estradiol Concentration (pg/mL)		Coefficient of Variation (%)
		ID-GC-MS	Method Mean	
Access	1	18	38.7	14.7
	2	31	52.5	8.0
	3	93	98.2	9.3
	4	387	422.0	3.3
	5	846	824.9	3.9
ADVIA Centaur	1	18	40.5	42.6
	2	31	57.4	19.4
	3	93	106.2	6.5
	4	387	443.7	6.8
	5	846	1,065.5	7.6
ARCHITECT i2000	1	18	58.6	12.0
	2	31	60.3	15.7
	3	93	119.8	4.8
	4	387	458.3	2.0
	5	846	>1,000	Not determined
ARCHITECT i2000 in development	1	18	25.1	9.9
	2	31	30.2	5.1
	3	93	75.9	2.5
	4	387	256.4	1.2
	5	846	893.3	2.5
AutoDELFIA	1	18	22.9	31.4
	2	31	32.7	20.3
	3	93	90.9	5.8
	4	387	361.2	3.7
	5	846	718.9	4.3
Elecsys 2010	1	18	28.1	6.9
	2	31	35.5	8.9
	3	93	91.6	4.9
	4	387	520.0	3.6
	5	846	1,023.1	3.6
IMMULITE 2000	1	18	28.4	14.7
	2	31	20.9	21.0
	3	93	67.9	9.4
	4	387	368.6	4.7
	5	846	901.6	4.4
Vitros ECi	1	18	34.9	8.9
	2	31	38.1	6.6
	3	93	73.7	6.1
	4	387	338.0	2.5
	5	846	673.3	3.1

ID-GC-MS, isotope dilution–gas chromatography–mass spectrometry.
* Values are given in conventional units; for conversion to Système International units (pmol/L), multiply by 3.67.
† For proprietary information, see the text.

Table 2
Estimated Functional Sensitivity*

Assay	Functional Sensitivity (pg/mL)
Access	37
ADVIA Centaur	116 (61)†
ARCHITECT i2000	35
ARCHITECT i2000 in development	<18
AutoDELFIA	74 (57)†
Elecsys 2010	<18
IMMULITE 2000	80
Vitros ECi	<18

* The functional sensitivity was defined as the estradiol concentration that would produce a coefficient of variation (CV) of 10%. It was estimated from a power curve fit of the data in Table 1 as shown in Figure 1. For proprietary information, see the text.
† Inspection of the power curves in Figure 1 revealed that the concentration of estradiol corresponding to a CV of 10% was overestimated for the ADVIA Centaur and AutoDELFIA methods. The imprecision data for the 18, 31, and 93 pg/mL samples were refitted with a power curve, and the result shown in parentheses provides a better estimate of the true functional sensitivity of these assays.

slopes very close to 1.0. The remainder of the instruments showed varying degrees of bias and possibly could benefit from improved calibration. This is especially true for the Elecsys 2010, which demonstrated excellent precision but had a positive proportional bias.

Ovarian stimulation protocols require serial monitoring of estradiol levels to direct dosage adjustments of gonadotropins for optimization of follicular growth and prevention of ovarian overstimulation.[10,11] Most of these decision points are between estradiol concentrations of 200 and 2,000 pg/mL (734-7,340 pmol/L),[10] and all 8 instruments have adequate precision within this range with CVs of less than 10%. Before initiating an ovarian stimulation protocol, differentiating populations with poor stimulation

DOI: 10.1309/5N2R4HT4GM0AGPBY

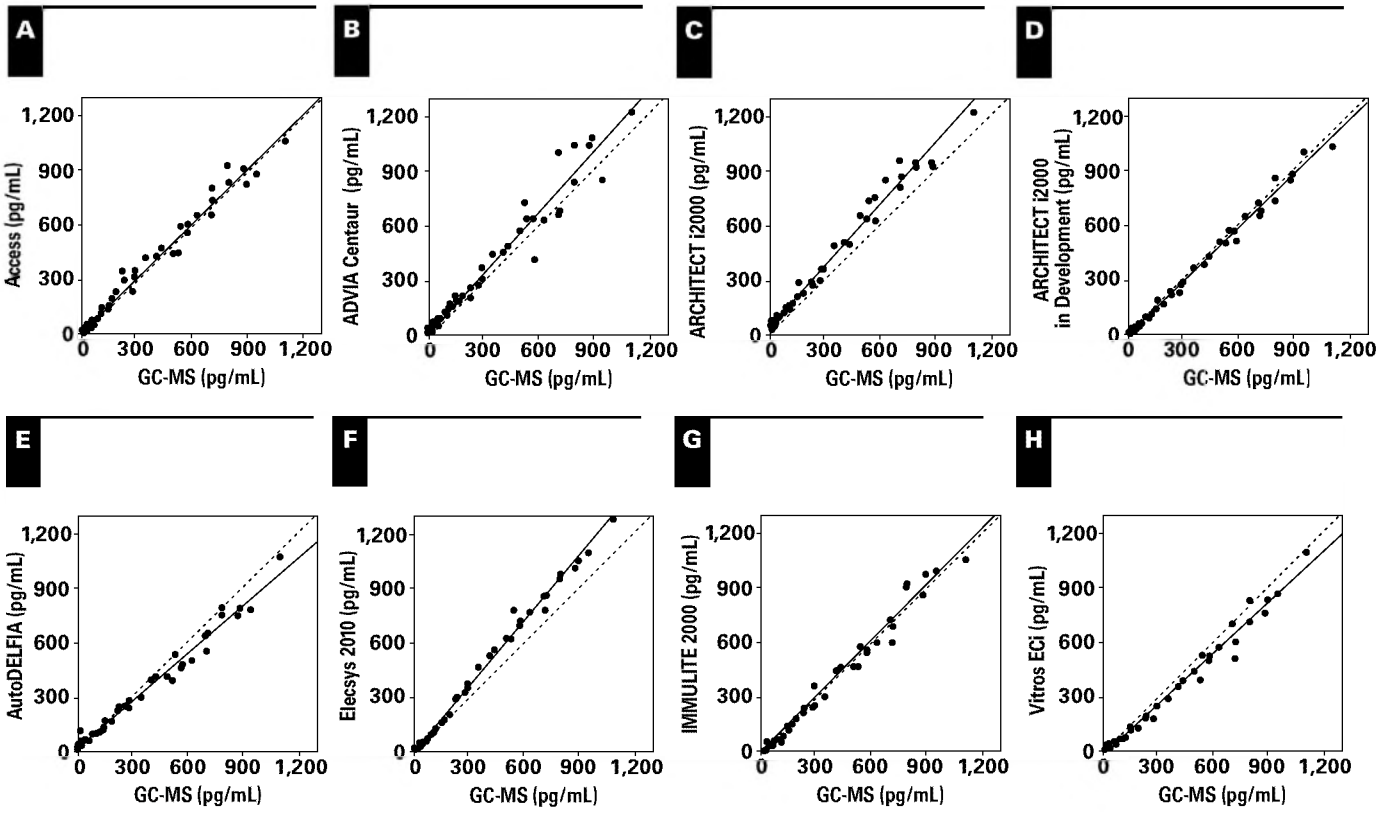

Figure 1 Comparison of estradiol immunoassay and isotope dilution–gas chromatography–mass spectrometry (ID-GC-MS) results using Deming regression. The solid lines are from Deming regression analysis, and the dashed lines are x = y. Data are given for the assay compared with the ID-GC-MS reference method. **A**, Access: slope, 0.99; intercept, 21 pg/mL; standard error of the estimate ($S_{y/x}$), 38 pg/mL; $r = 0.992$. **B**, ADVIA Centaur: slope, 1.10; intercept, 9 pg/mL; $S_{y/x}$, 66 pg/mL; $r = 0.980$. **C**, ARCHITECT i2000: slope, 1.12; intercept, 41 pg/mL; $S_{y/x}$, 40 pg/mL; $r = 0.993$. **D**, ARCHITECT i2000 in development: slope, 0.98; intercept, 1 pg/mL; $S_{y/x}$, 23 pg/mL; $r = 0.997$. **E**, AutoDELFIA: slope, 0.87; intercept, 17 pg/mL; $S_{y/x}$, 32 pg/mL; $r = 0.993$. **F**, Elecsys 2010: slope, 1.20; intercept, –1 pg/mL; $S_{y/x}$, 28 pg/mL; $r = 0.997$. **G**, IMMULITE 2000: slope, 1.03; intercept, –7 pg/mL; $S_{y/x}$, 36 pg/mL; $r = 0.993$. **H**, Vitros ECi: slope, 0.92; intercept, –9 pg/mL; $S_{y/x}$, 32 pg/mL; $r = 0.993$. GC-MS, gas chromatography–mass spectrometry. Intercept and $S_{y/x}$ values are given in conventional units; for conversion to Système International units (pmol/L), multiply by 3.67. For proprietary information, see the text.

characteristics from those with good stimulation characteristics is a prognostic necessity to select the appropriate stimulation protocol. This can be done by measuring basal levels of estradiol and follicle stimulating hormone on day 3 of the menstrual cycle, and the determination point is an estradiol concentration of 60 pg/mL (220 pmol/L).[12] All of the assays except the IMMULITE 2000 demonstrated what we consider to be adequate precision at low estradiol concentrations for this task (CV <10%).

While the Access, ARCHITECT i2000 assay in development, and IMMULITE 2000 assays demonstrated good concordance with the reference method, additional calibration standardization of the ADVIA Centaur, AutoDELFIA, Elecsys 2010, and Vitros ECi assays might improve their accuracy and could lead to better harmonization of estradiol results. Additional studies to examine interfering substances, particularly endogenous and exogenous estrogens, need to be performed because the presence of some potential cross-reacting substances could contribute to the bias and scatter observed for some of the methods.

The clinical applications of estradiol assays can be diverse. This study shows that the ARCHITECT i2000 assay in development, Elecsys 2010, and Vitros ECi assays have adequate low-end precision for the measurement of estradiol levels in men and postmenopausal women, but the accuracy of the Elecsys 2010 method needs some improvement. The rapid analysis time and low technical demands of automated immunoassays have made these instruments popular in the field of assisted reproduction. Within the range of estradiol concentrations typically encountered in this application, the Access, ARCHITECT i2000 in development, and Vitros ECi assays seem to have adequate precision and accuracy. The

DOI: 10.1309/5N2R4HT4GM0AGPBY

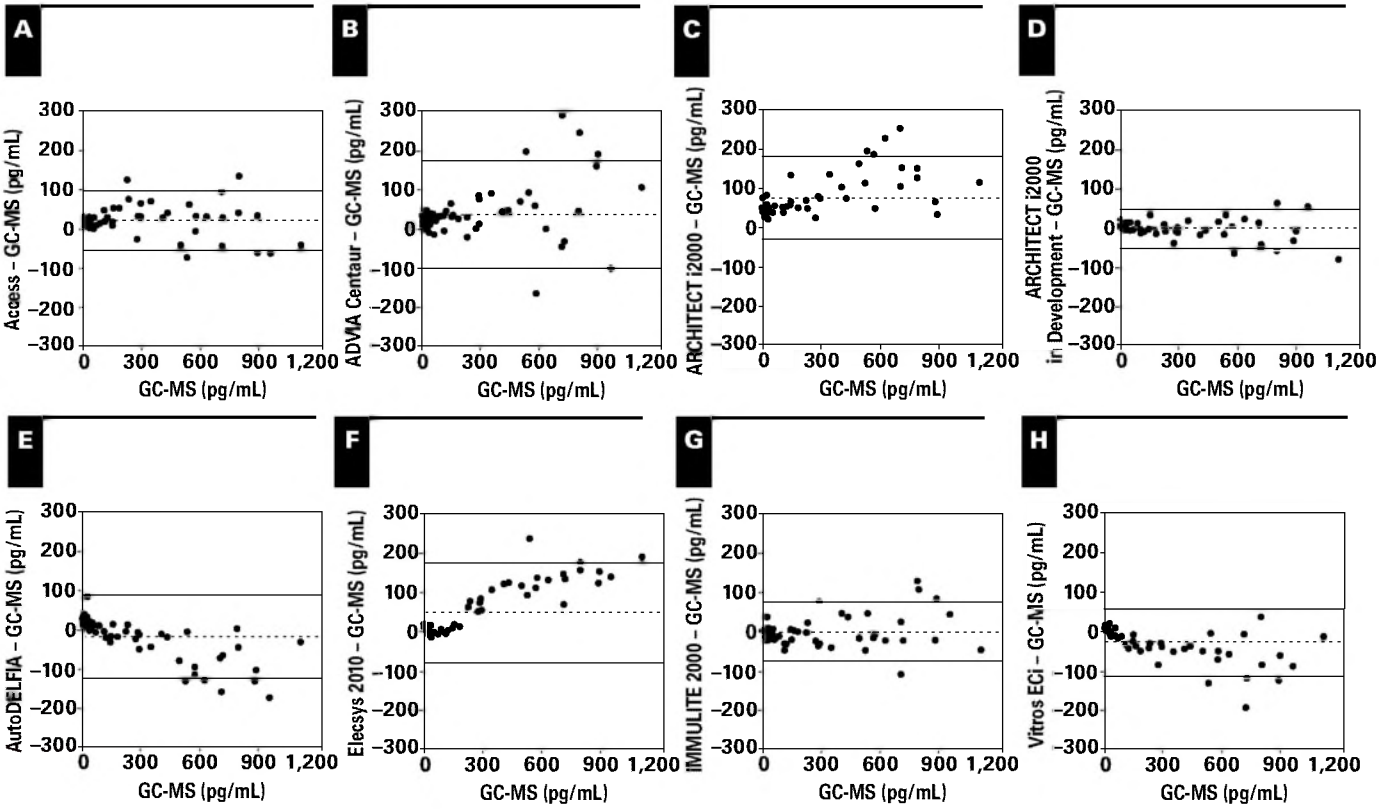

Figure 2 Comparison of estradiol immunoassays with isotope dilution–gas chromatography–mass spectrometry by difference plots. The solid lines are the ± 2 SD limits of agreement, and the dashed lines are the mean bias. All values are given in conventional units (pg/mL); for conversion to Système International units (pmol/L), multiply by 3.67. **A**, Access: mean bias, +20; limits of agreement, –55 and +95. **B**, ADVIA Centaur: mean bias, +34; limits of agreement, –104 and +172. **C**, ARCHITECT i2000: mean bias, +70; limits of agreement, –32 and +173. **D**, ARCHITECT i2000 in development: mean bias, –5; limits of agreement, –54 and +44. **E**, AutoDELFIA: mean bias, –17; limits of agreement, –122 and +89. **F**, Elecsys 2010: mean bias, 48; limits of agreement, –79 and +176. **G**, IMMULITE 2000: mean bias, 0; limits of agreement, –72 and +72. **H**, Vitros ECi: mean bias, –29; limits of agreement, –113 and +54. GC-MS, gas chromatography–mass spectrometry. For proprietary information, see the text.

changes made to the reagents for the ARCHITECT i2000 in development have resulted in an improvement in assay performance compared with the current ARCHITECT i2000 reagents. The IMMULITE 2000, while accurate, lacked adequate precision at low estradiol concentrations for this application, and, conversely, the Elecsys 2010, while precise, lacked the requisite accuracy for this application.

From the [1]Department of Pathology, University of Utah Health Sciences Center, Salt Lake City; [2]ARUP Institute for Clinical and Experimental Pathology, Salt Lake City; and [3]Abbott Diagnostics Division, Abbott Park, IL.

Supported by the Diagnostics Division of Abbott Laboratories and the ARUP Institute for Clinical and Experimental Pathology.

Address reprint requests to Dr Roberts: ARUP Laboratories, 500 Chipeta Way, Salt Lake City, UT 84108.

References

1. Gronowski AM, Landau-Levine M. Reproductive endocrine function. In: Burtis CA, Ashwood ER, eds. *Tietz Textbook of Clinical Chemistry*. 3rd ed. Philadelphia, PA: Saunders; 1999:1601-1641.
2. Painter PC, Cope JY, Smith JL. Reference information for the clinical laboratory. In: Burtis CA, Ashwood ER, eds. *Tietz Textbook of Clinical Chemistry*. 3rd ed. Philadelphia, PA: Saunders; 1999:1788-1846.
3. Thienpont L. Meeting report: first and second estradiol international workshops [letter]. *Clin Chem*. 1996;42:1122-1124.
4. Thienpont LM, De Leenheer AP. Efforts by industry toward standardization of serum estradiol-17β measurements. *Clin Chem*. 1998;44:671-674.
5. Levesque A, Letellier M, Dillon P, et al. Analytical performance of Bayer Immuno 1 estradiol and progesterone assays. *Clin Chem*. 1997;43:1601-1609.

DOI: 10.1309/5N2R4HT4GM0AGPBY

6. Dancoine F, Couplet G, Buvat J, et al. Analytical and clinical evaluation of the Immulite estradiol assay in serum from patients undergoing in vitro fertilization: estradiol increase in mature follicles. *Clin Chem*. 1997;43:1165-1171.
7. Rodriguez-Espinoza J, Otal-Entraigas C, Gascon-Roche N, et al. Analytical and clinical performance of an automated immunoassay system (Immulite) for estradiol in serum. *Clin Chem Lab Med*. 1998;36:969-974.
8. Reinsberg J, Jost E. Analytical performance of the fully automated AxSYM estradiol assay. *Clin Chem Lab Med*. 2000;38:51-55.
9. Wu H, Ramsay C, Ozaeta P, et al. Serum estradiol quantified by isotope dilution–gas chromatography/mass spectrometry. *Clin Chem*. 2002;48:364-366.
10. Maclin VM. Ovarian stimulation and ovulation induction. In: Keel BA, May JV, De Jonge CJ, eds. *Handbook of the Assisted Reproduction Laboratory*. Boca Raton, FL: CRC Press; 2000:139-178.
11. Navot D, Bergh PA, Laufer N. Ovarian hyperstimulation syndrome in novel reproductive technologies: prevention and treatment. *Fertil Steril*. 1992;58:249-261.
12. Evers JL, Slaats P, Land JA, et al. Elevated levels of basal estradiol-17β predict poor response in patients with normal basal levels of follicle-stimulating hormone undergoing in vitro fertilization. *Fertil Steril*. 1998;69:1010-1014.

DOI: 10.1309/5N2R4HT4GM0AGPBY