ORIGINAL ARTICLE

Evaluation of the Analytical Performance Beckman Coulter Access Sensitive Estradiol Assay

O. Çakir Madenci, Z. Yıldız, A. Orcun, O. Hurmeydan, I. Onem, L. Koroglu Dagdelen

Dr. Lütfi Kirdar Kartal Research and Training Hospital, Biochemistry Laboratory, Istanbul, Turkey

SUMMARY

Background: The study aimed to assess the analytical performance of the Access Sensitive Estradiol (SNSE2) Assay on a DxI800 (Beckman Coulter, Brea, CA, USA) and compared it with a Cobas E 601 (Roche Diagnostics, Penzberg, Germany).

Methods: SNSE2 was assessed for imprecision, accuracy, limit of blank (LoB), limit of detection (LoD), limit of quantitation (LoQ), linearity, interference, and carryover. Two hundred and fourteen samples were run on both instruments. Bland-Altman plots, Passing-Bablok regression, and concordance correlation coefficient (CCC) graphs were used for comparisons.

Results: Access SNSE2 showed appropriate assay performance characteristics in terms of imprecision, LoB, LoD, LoQ, linearity, and interference. The Bland-Altman analysis of DxI 800 yielded negative bias from Cobas E 601 and the deviations for E2 ≤ 150 pmol/L, 150 - 500 pmol/L, and ≥ 500 pmol/L were found as 0.8%, -15%, and -8.9%, respectively. DxI 800 and E170 systems showed poor agreement for E2 levels ≤ 150 pmol/L and 150 - 500 pmol/L with CCC values of 0.7404 and 0.8342. For E2 levels ≤ 150 pmol/L there was a significant amount of both proportional and constant error with the highest slope of 1.518 (1.269 to 1.761) and an intercept of -45.08 (-66.09 to 18.78, respectively, according to the Passing-Bablok regression analysis).

Conclusions: Analytical performance for SNSE2 assay was found appropriate. However, attempts to improve harmonization and standardization across assays do not seem to contribute much for E2 measurements. Results obtained with different systems cannot be used interchangeably and follow up of patients should be done with the same system.


KEY WORDS
estradiol, hormone, standardization, immunoassay, limits of detection

INTRODUCTION

Estradiol (E2) measurements are used for diagnosis and management of various conditions such as evaluation of female infertility and menopause, feminization in males, hypogonadism, disorders of puberty, monitoring follicular development in assisted reproduction and response in women receiving aromatase inhibitor (AI) therapy [1,2]. E2 concentrations vary over a wide range in menstruating women from 18 to 1,505 pmol/L while concentrations are often less than 184 pmol/L, in men, chil-
women receiving hormone replacement therapy were not included in the study. Results below LoD of the systems were not included. Fasting blood samples were drawn into 5 mL BD Vacutainer Serum Separating Tubes II Advance Tube (REF 366566) (Becton, Dickinson, and Company, BD Plymouth PL6 7BP, UK) and centrifuged at 2,000 x g for 10 minutes. We divided the samples into two aliquots and stored at -20°C for a maximum of three months. We analyzed the samples on both systems. The Ethical Committee of our institution approved the study with a decision number: 2020/514/170/8.

Methods
Analytical performance of newly developed SNSE2 assay was assessed on a UniCel DxI 800 (Beckman Coulter, Brea, CA, USA) in terms of imprecision, accuracy, LoB, LoD, LoQ, linearity, interference, and carryover studies. Method comparison was performed with the Cobas e 601 (Roche Diagnostics, Penzberg, Germany) platform. Access SNSE2 (REF B84493) is a two-step competitive immunoassay and it was traceable to the Joint Committee for Traceability in Laboratory Medicine - approved isotope dilution mass spectrometry (ID/GC/MS) reference method procedure. Performance characteristics provided by the manufacturer for the Beckman Coulter SNSE2 assay are given in Table 1. There was standardization against CRM 6004a via ID GC/MS for Cobas e 601 assays. The total coefficient of variations (CV) values for Cobas e 601 E2 assays (REF 06656021) provided by the manufacturer was 10.6% for 101 pmol/L and 2.1% for 1,497 pmol/L and the linearity was given as 18.4 - 11,010 pmol/L.

Assay performance studies
Imprecision
Serum pools at three different E2 concentrations; (81.5, 132, and 1,104 pmol/L) were used to determine imprecision (CV total). We measured each sample in duplicate twice a day with a minimum of 4 hours between each run and completed the study in 20 days. Total CVs were calculated and CV ≤ 10% was accepted as appropriate [18].

Accuracy
For evaluation of accuracy, Randox International Quality Assessment Scheme (RIQAS) external quality assessment monthly immunoassay program (code RQ-9130) was used. Since there is no certified reference material for E2, we used group target values calculated by RIQAS. The last three samples of Cycle 16 for the concentrations 150, 833, and 1,457 pmol/L were analyzed on a DXI 800. Percent difference was calculated from the reported target mean using the formula: ((result - mean)/mean) x 100. RIQAS defined acceptable accuracy as 16.9% for E2 measurements.

MATERIALS AND METHODS

Subjects
Samples of male and female outpatients between the ages of 15 - 75 years attending to Istanbul Dr. Lutfi Kirdar City Hospital between October and December 2018 were included in the study. This is an analytical method evaluation study. We included E2 samples with concentrations between the measurement range of both systems. All test requests for diagnosis, prognosis, and monitoring were included to gain a wide range of E2. All phases of the menstrual cycle in reproductive age, ovulation induction, or at pre- and postmenopausal state were included. Females were not currently pregnant, not using oral contraceptives, nor had had hormonal therapy in the three months before the study. Post-menopausal

dren, and postmenopausal women [2]. Serum E2 concentration can differ between 918 - 7,342 pmol/L in women going through ovulation induction. AI therapy causes a decrease in serum E2 concentration to < 18 pmol/L [3]. It is a problem for the laboratories to determine precise and accurate E2 measurements at both ends of this concentration spectrum [1,4]. For accurate quantitation at these different clinical conditions, highly sensitive and specific E2 assays are needed [1,5-8]. Moreover, interfering compounds in the serum may produce 10 times higher false-positive E2 results [9]. The gold standard for the measurement of E2 is isotope dilution/gas chromatography (GC) coupled with mass spectrometry which was too complex and slow for routine clinical use [1]. Gas-chromatography mass spectrometry (GC-MS) [10-12], liquid chromatography-tandem mass spectrometry (LC-MS/MS) [13-16], and immunoassays were the other methodologies used for E2 quantitation. LC-MS/MS assays have better specificity and sensitivity than immunoassays but their cost, slower turn around time, and requirement for technical skills limit their routine use [16].

Despite high speed and throughput in immunoassays, high variability across different assays has been observed at low E2 concentrations especially in men, postmenopausal women, and children. In this patient population, most automated direct immunoassays fail to meet current analytical performance needs relevant to research and patient care [6,17]. To reduce the variability in E2 testing and to persuade better patient care, standardization of E2 measurements was recommended with a single widely available standard in which all measurements can be traced [1,6].

In April 2018, Beckman Coulter introduced a new Access SNSE2 assay and claimed to be able to detect low concentrations of E2 with high precision and accuracy. In this study, we aimed to assess the analytical performance of Access SNSE2 immunoassay performance in terms of imprecision, accuracy, LoB, LoD, LoQ, linearity, interference, and carryover on a DxI800 and compare it with a Cobas e 601.

Subjects
Samples of male and female outpatients between the ages of 15 - 75 years attending to Istanbul Dr. Lutfi Kirdar City Hospital between October and December 2018 were included in the study. This is an analytical method evaluation study. We included E2 samples with concentrations between the measurement range of both systems. All test requests for diagnosis, prognosis, and monitoring were included to gain a wide range of E2. All phases of the menstrual cycle in reproductive age, ovulation induction, or at pre- and postmenopausal state were included. Females were not currently pregnant, not using oral contraceptives, nor had had hormonal therapy in the three months before the study. Post-menopausal
LoB, LoD, and LoQ
The manufacturer’s zero calibrators were analyzed 20 times and LoB was calculated using the following formula.

\[ \text{LoB} = \text{mean (blank)} + 1.645 \times \text{SD (blank)} \]

The LoD was determined using the lowest non-zero calibrator (38.2 pmol/L) which was diluted (1/2) for a final concentration of 19.1 pmol/L and 20 replicates were analyzed. LoD was calculated using the following formula. \[ \text{LoD} = \text{LoB} + 1.645 \times \text{SD (low-concentration sample)} \]

For the LoQ study, we measured six samples with concentrations ranging from 19.08 to 245 pmol/L for 10 days and determined the CVs. The point at which the fitted curve crosses the 20% CV line is defined as LoQ [19,20].

Linearity
By diluting the highest standard of the SNSE2 reagent kit with the zero calibrators, nine different concentrations in the range of 19.08 - 19,530 pmol/L was performed. The diluents were not equally distributed and three replicates of analyses were performed in a single run. A deviation less than ± 15% from the target concentration was accepted [21].

Interference studies
Hemoglobin, bilirubin, and triglycerides were tested as interferents. Averages of the duplicate measurements were determined and percent differences for each interferent were calculated using the following formula.

\[ \text{Percent difference} = \frac{(\text{spiked} - \text{nonspiked})/\text{nonspiked}}{} \times 100 \]

Percent difference more than 10% was significant. We drew blood from healthy volunteers into Plastic Whole Blood tubes spray-coated with K2EDTA (BD Vacutainer, lot 9175879) and centrifuged the samples at 5,000 x g for 5 minutes for the determination of hemoglobin interference. We omitted the plasma and washed the cell pocket three times with physiological serum. The supernatant was omitted again and an equal volume of distilled water was added to the erythrocyte pocket. To get erythrocytes hemolysed, the tube was placed at -40°C for 20 minutes. After 20 minutes it was centrifugated again and hemolysate was obtained from the supernatant. We measured hemoglobin concentrations on the LH780 hematology analyzer (Beckman Coulter, Brea, CA, USA). Seven different concentrations of hemoglobin at 0.25, 0.5, 1, 2, 3, 4, and 5 g/L were spiked in patient serum pools [21]. For the determination of bilirubin interference, we used the bilirubin standard (≥ 98%, Sigma Aldrich B4126, EmM/453 = 60) and dissolved it in chloroform (Merck, M102445.250). Four different bilirubin concentrations at 171, 342, 513, and 684 pmol/L were spiked into the serum pools.

Samples with 7 different triglyceride concentrations at 5.65, 8.6, 11.2, 21.9, 25.7, 32, and 36.2 mmol/L were prepared from the lipid standard Intralipid 20% (SigmaAldrich) and spiked into serum pools for evaluating triglyceride interference [22].

Carryover
We measured 3 replicates of a high-concentration sample (samples a1, a2, and a3) immediately followed by 3 replicates of a low-concentration sample (samples b1, b2, b3). We used the equation \((b1 - b3)/(a3 - b3)\) for carryover calculation. Results < 2% was accepted as insignificant [23].

Method comparison
E2 samples with concentrations within the measurement range of both systems were measured in a single batch, in duplicate, within the same freeze/thaw cycles [24].

Statistical analysis
Kolmogorov-Smirnov test was used to evaluate the distribution of data and results were expressed as median and interquartile range. For imprecision, LoB, LoD, LoQ, and linearity studies EP Evaluator Release 10 software (David G Rhoads Association, Kennett Square, PA, USA) was used. Bland-Altman plots, Passing-Bablok regression, and concordance correlation coefficient (CCC) were performed for method comparison studies. MedCalc Statistical Software (version 12. MedCalc Software, Mariakerke, Belgium) was used. If the 95% confidence intervals did not include 1.0 for slope (a proportional error) or 0 for the y-intercept (constant error) the systematic error was considered significant. The interpretation of CCC is as follows: > 0.99 excellent; 0.95 - 0.99 good; 0.90 - 0.94 moderate, and < 0.90 poor agreement.

RESULTS
Nineteen postmenopausal women with a median age of 68 years (range 61 - 75 years) were excluded since their results were below LoD of both systems. Thus, a total of 214 patient samples (males; 17.2%, females; 82.8%) with a median age of 35 years (range 15 - 60 years), with E2 concentrations in the range of 37.8 - 10,356 pmol/L, were included in the study. Assay performance characteristics in terms of imprecision, LoB, LoD, LoQ, linearity, and interference were found appropriate for the Access SNSE2 assay. Performance characteristic results of the study and the manufacturer claims are given in Table 1. For method comparison, the median (2.5 - 97.5 percentiles) values of 214 samples were 211 (55.9 - 8,235) pmol/L for DxI 800 and 253 (38.1 - 8,522) pmol/L for Cobas e 601.

In the visual assessment of the Bland-Altman plots, the distribution of the samples was not homogenous. So, we categorized the data into three groups according to the concentration ranges of DXI 800 results as: E2 concentrations ≤ 150 pmol/L, 150 - 500 pmol/L, and > 500 pmol/L. Bland-Altman plots obtained for E2 concentrations ≤ 150 pmol/L, 150 - 500 pmol/L, and > 500 pmol/L were shown in Figure 1 - 3.
Table 1. Performance characteristics of methods.

<table>
<thead>
<tr>
<th>Performance criteria</th>
<th>SNSE2 DXI 800</th>
<th>SNSE2 DXI 800 (manufacturer’s claim)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imprecision Total CV %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>first concentration</td>
<td>14.4%</td>
<td>81.5 pmol/L</td>
</tr>
<tr>
<td>second concentration</td>
<td>4.14 %</td>
<td>132 pmol/L</td>
</tr>
<tr>
<td>third concentration</td>
<td>3.74%</td>
<td>1,104 pmol/L</td>
</tr>
<tr>
<td>Accuracy %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>riqas 1</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>riqas 2</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>riqas 3</td>
<td>5.64</td>
<td></td>
</tr>
<tr>
<td>Interference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hemoglobin (g/L)</td>
<td>≥ 5</td>
<td>≥ 3</td>
</tr>
<tr>
<td>triglyceride (mmol/L)</td>
<td>≥ 28.2</td>
<td>≥ 37</td>
</tr>
<tr>
<td>bilirubin (μmol/L)</td>
<td>≥ 581</td>
<td>≥ 684</td>
</tr>
<tr>
<td>Carryover %</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>LoB pmol/L</td>
<td>18.9</td>
<td>36.7</td>
</tr>
<tr>
<td>LoD pmol/L</td>
<td>36.1</td>
<td>55.1</td>
</tr>
<tr>
<td>LoQ pmol/L</td>
<td>65.5</td>
<td>69.7</td>
</tr>
<tr>
<td>Linearity</td>
<td>36.1 - 16,600</td>
<td>55.1 - 19,089</td>
</tr>
</tbody>
</table>

Table 2. Method comparison results of DXI 800 and Cobas E 601 systems.

<table>
<thead>
<tr>
<th>Concentration range</th>
<th>Passing-Bablok regression analysis</th>
<th>Concordance correlation analysis</th>
<th>Bland-Altman analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slope (CI)</td>
<td>Intercept (CI)</td>
<td>CCC (CI)</td>
</tr>
<tr>
<td>≤ 150 pmol/L</td>
<td>1.518 * (1.269 - 1.761)</td>
<td>-45.08 † (-66.09 - 18.78)</td>
<td>0.7404 (0.648 - 0.811)</td>
</tr>
<tr>
<td>150 - 500 pmol/L</td>
<td>1.226 * (1.138 - 1.368)</td>
<td>-17.579 † (-48.9 - 2.159)</td>
<td>0.834 (0.770 - 0.881)</td>
</tr>
<tr>
<td>≥ 500 pmol/L</td>
<td>1.053 * (1.029 - 1.087)</td>
<td>22.818 † (-4.503 - 64.806)</td>
<td>0.993 (0.990 - 0.996)</td>
</tr>
</tbody>
</table>

* Slope CI's do not include 1.0 mean proportional error. † intercept CI's do not include 0 mean constant error.
CI - confidence interval, Cb - Bias correction factor (accuracy), CCC - Concordance correlation coefficient, P - Pearson’s correlation coefficient (precision).

In this study, we found that Access SNSE2 showed appropriate analytical performance in terms of precision, accuracy, LoB, LoD, LoQ, linearity, interference, carryover, and provided manufacturer claims. LoB, LoD and LoQ values were lower than the manufacturer claim. Using calibrators instead of patient samples might cause the difference. In our study, both systems failed to measure E2 concentrations accurately in postmenopausal women and they were all below LOD of the systems. Minimal allowable bias according to biological variabil-

DISCUSSION

500 pmol/L were shown in Table 2. For all three concentration ranges, there was a proportional error. For E2 concentration ≤ 150 pmol/L there was both proportional and constant error with the highest slope of 1.518 (1.269 - 1.761) and an intercept of -45.08 (-66.09 to -18.78). Results of the method comparison study are given in Table 2. In concordance correlation analysis Dx1 800 and Cobas e 601 systems showed poor agreement for E2 concentrations ≤ 150 pmol/L and 150 - 500 pmol/L with CCC values of 0.7404 and 0.8342. The almost perfect agreement was observed for E2 concentrations ≥ 500 pmol/L with a CCC value of 0.993.
Figure 1. Bland-Altman plots of method comparison for Estradiol measurements ≤ 150 pmol/L.
Mean (thick solid line) - percentage bias (means of paired differences). Dashed lines demonstrate the 95% limits of agreement (bias ± 1.96 standard deviation).

Figure 2. Bland-Altman plots of method comparison for Estradiol measurements 150 - 500 pmol/L.
Mean (thick solid line) - percentage bias (means of paired differences). Dashed lines demonstrate the 95% limits of agreement (bias ± 1.96 standard deviation).
ity was given as ± 12.5% [25,26] and the acceptable accuracy was defined as 16.9% for RIQAS. In our study bias % for the last three RIQAS samples was found acceptable according to both criteria and 1.11%, 0.65% and 5.64% for the concentrations 150 pmol/L, 833 pmol/L, and 1,457 pmol/L, respectively. According to the method comparison study, the Bland-Altman analysis of DxI 800 yielded negative bias from Cobas e 601 for E2 concentrations all throughout the study range. In the visual assessment of Bland-Altman plots, percent difference distribution was different mainly in 3 concentration intervals indicating the existence of systematic error for concentrations ≥ 150 pmol/L, maybe as a result of the difference in calibrators. For the E2 concentrations below 150 pmol/L, the Bland-Altman analysis yielded a small bias but with a wide range of confidence intervals indicating the existence of imprecision problem in the systems. According to Passing-Bablok regression analyses, DxI 800 showed proportional error for all three concentration ranges. There was both proportional and constant error at E2 concentrations ≤ 150 pmol/L.

In a study comparing 11 immunoassays and 6 mass spectrometry methods with the GC-MS reference method reported great biases between test and reference methods in samples from men, pre- and post-menopausal women [2]. In another study, bias increased with decreasing E2 concentrations and most immunoassays were not able to measure E2 concentrations below 36.7 pmol/L. Constant diagnosis of premature ovarian failure, monitoring of AI therapy, and accurate E2 concentrations in men cannot be achieved because of interassay variability, low analytical sensitivity, and specificity [6]. Recent College of American Pathologists survey data reported that 14 immunoassays overestimated the E2 concentration of 106 pmol/L with a great variability [3]. In another study evaluating the performance of 7 automated assays, the method-specific variability was very different and the precision for very low E2 concentrations, as observed in postmenopausal women, were inadequate for clinical assessments [8].

The limitation of our study was that we did not compare the results with the reference ID-GC-MS method. Another limitation was that we used calibrators instead of patient samples in LoD and LoQ analyses. Besides, our study group was a heterogeneous group consisting of a variety of clinical situations such as menstruating, pre and postmenopause, and both sexes because we wanted to include E2 concentrations on both ends of the clinical range.
Although the analytical performance was acceptable, we could not find comparable results between the two systems in the method comparison study. There was still a precision problem at the low concentrations of E2 seen in post menopausal women and children. Thus, we conclude that attempts to improve harmonization and standardization across assays do not seem to contribute much to E2 measurements. Results obtained with different systems cannot be used interchangeably and follow up of the patients should be done in the same laboratory with the same system.

**Declaration of Interest:**
The authors declare that there is no conflict of interest.

**References:**