ORIGINAL RESEARCH

The Application of the EP9-A Protocol in the Analysis of the Performance of the Immunofluorescence Assay for HCG Detection

Rong Huang[1]
Shui Fu[1]
Yin-Jiang Lv[1]
Liang Zhang[1]
Li-Feng Wu[1]
Qi-Lei Hu[1]
Juan-Fei Qi[2]

[1]Department of Clinical Laboratory, The First People's Hospital of Yuhang District, Hangzhou, Hangzhou, 311100, People's Republic of China; [2]Department of Clinical Laboratory, Xinchang Hospital of Traditional Chinese Medicine, Shaoxing, 312500, People's Republic of China

Objective: The present study aims to evaluate the comparability of the results of two methodologies for detecting human chorionic gonadotropin (HCG) to assess whether the immunofluorescence method for detecting HCG is adequate for clinical applications.

Methods: Referring to the protocol requirements of the American Clinical Laboratory Standards Institute (CLSI) EP9-A2 (methodological matching and bias assessment with patient samples), we collected 40 fresh serum specimens from our outpatients and inpatients, including 20 specimens with abnormal HCG concentrations (eight samples with different concentration ranges were selected daily and HCG was measured simultaneously with the two testing systems for five consecutive days). The assays were performed on a Dxl 800 fully automated immunoassay analyzer from Beckman Coulter Inc., USA, as a comparative method and on a Jet-iStar 3000 immunoassay analyzer from Zhonghan Shengtai Inc. as an experimental method. Methodological comparison and bias assessment of the results of the two methods for HCG detection were performed. The OLR regression model was used for calculating bias and regression analysis, and Spearman's rank correlation coefficient was used for correlation analysis. The correlation and comparability of the two systems were calculated based on the results of the analysis.

Results: A good correlation in HCG results in the range of 5–50,000 U/mL was obtained from the two assay systems (r = 0.998) with the regression equation of y = 1.020x + 12.96. The estimated deviation was within the permissible deviation and acceptable.

Conclusion: The results of HCG measurement by the two different assay systems were well correlated and comparable.

Keywords: human chorionic gonadotropin, assay system, EP9-A2, correlation

Introduction

With the development of testing technology and the widespread use of immunoassay analyzers, it has become common to have different systems testing the same specimens in the same clinical laboratory.[1,2] It is necessary to perform a periodic evaluation of varying assay systems to ensure good homogeneity and accuracy of results between different assay systems. Human chorionic gonadotropin (HCG), a glycoprotein hormone, has been widely used to diagnose early pregnancy, ectopic pregnancy, differential diagnosis of acute abdomen, and diagnose some kinds of malignancy. Therefore, many hospitals have adopted HCG as an emergency item to serve clinical practices and patients. The Jet-iStar 3000 immunoassay analyzer has the characteristics of automation, informationization, and

Correspondence: Juan-Fei Qi
Department of Clinical Laboratory, Xinchang Hospital of Traditional Chinese Medicine, Shaoxing, 312500, People's Republic of China
Tel/Fax +86 575-86502591
Email qijuanfei_dr@163.com

integration, and the assaying speed is fast, which meets the needs of domestic primary care hospitals. However, there are relatively few studies on its accuracy and consistency. Currently, the accuracy and stability of the Beckman Dxl 800 immunoassay for HCG have been confirmed by numerous authoritative studies.[3,4] Therefore, in the present study, with the comparison method developed according to the EP9-A2 protocol, the Dxl 800 immunoassay analyzer was used as the reference method to analyze the performance of the Jet-iStar 3000 immunoassay analyzer.

Materials and Methods
Sample Preparation
According to the data distribution requirements of the method comparison test in the EP9-A2 protocol, 40 samples were collected,[5] of which 20 had abnormal HCG concentrations (samples in the abnormal range accounting for 50%), ensuring a wide concentration distribution of samples.

Apparatus and Reagents
With the adoption of the Beckman Coulter Dxl 800 immunoassay analyzer and the Jet-iStar 3000 immunoassay analyzer, together with the respective original reagents and quality control products, the assays were tested according to the instrument operation instructions.

Assay System
The Reference Assay System
The Beckman Coulter Dxl 800 immunoassay analyzer was used as the reference system. This analyzer was used in the provincial and Ministry of Health inter-laboratory quality assessment with excellent results.

The Assay System for Evaluation
The Jet-iStar 3000 immunoassay analyzer was the system for evaluation.

Methods
We collected 5 mL of venous blood and mixed it upside down five to eight times, placed it at room temperature for 30 min, centrifuged it at 4000 r/min (equivalent to 2200 g centrifugal force) for 10 min, and then analyzed it, making sure the instrument was in proper condition before measurement. Eight specimens were collected each day and tested simultaneously on the two apparatuses, and each specimen was measured in double, in the order of

1-8, 8-1, for five days. A total of 40 samples were analyzed.

Statistical Analysis
Outlier checks for double determinations within methods and visual checks for inter-method outliers were performed according to the CLSI EP9-A2 document requirements. First, the outliers were determined. If there was only one outlier, it could be deleted. If there were more than two outliers, the experimental data should be added after analyzing the reason and examined again. The percentage deviation of the measurement results between the two systems was calculated. The OLR regression model was used for calculating bias and regression analysis, and Spearman's rank correlation coefficient was used for correlation analysis. The results were regressed, and the linear and regression equations were calculated as $y = bx + a$.

Judgment of the Acceptability in the Performance Results
Since the permissible error of CLIA'88 HCG is ± 3 s of the target value (positive or negative), which is challenging to fit into the present study, we considered ≤12.5% as the clinically acceptable standard according to half of the inter-room quality evaluation standard (25%) of the clinical test center of the Health Care Commission. HCG has no authoritative medical decision level. The cut-off point (5 mIU/mL and 25 mIU/mL) combined with literature reports (400 mIU/mL and 10,000 mIU/mL) were adopted as the systematic error to determine whether the inter-system error was acceptable.

Results
Test of the Outlier in the Experimental Data
The intra-method check did not find any outliers, while the inter-method check found one outlier. According to the EP9-A2, the outlier was deleted, and then the reason was analyzed further.

The Scatter Plots and Regression Lines of the Measurement results of the Comparison Instrument and Experimental Instrument
The mean of the two measurement results by the comparison method was used as the axis X, and the mean
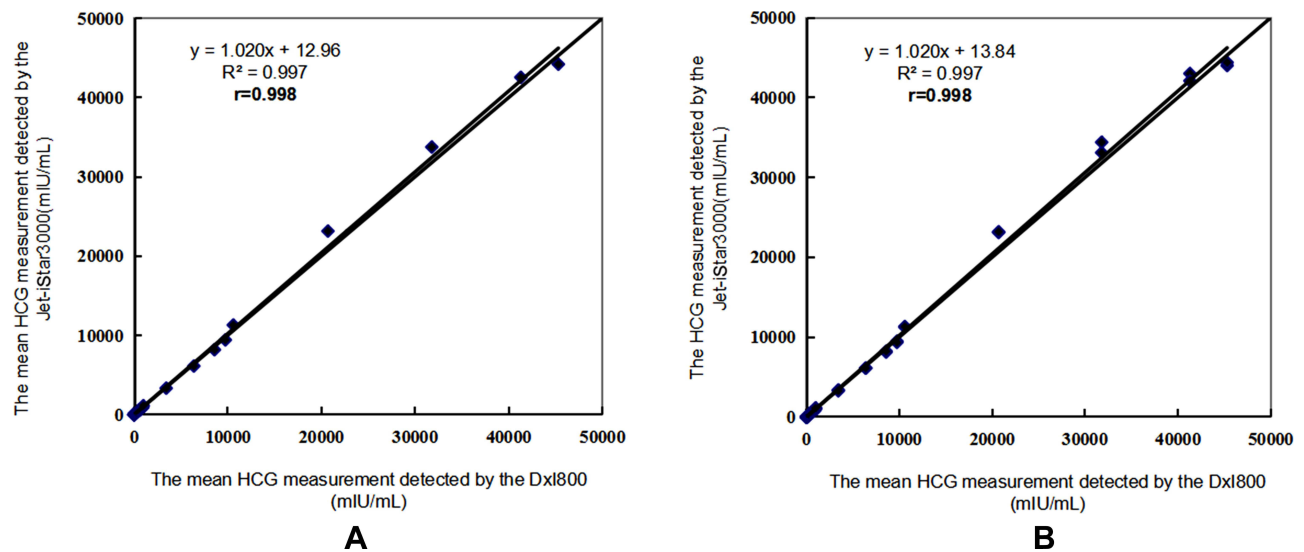
Figure 1 The results of reproducibility and correlation by the DxI800 and Jet-iStar3000. (A) The mean of the two measurement results by the comparison method was used as the axis X, and the mean of the two measurement results by the experimental method was used as the axis Y, plotted with a slope of 1.0 across the origin. (B) The mean of the two measurement results by the comparison method was used as the axis X, and the results of the experimental method were used as the axis Y, plotted with a slope of 1.0 across the origin.

of the two measurement results by the experimental method was used as the axis Y, plotted with a slope of 1.0 across the origin (Figure 1A). Then, the mean of the two measurement results by the comparison method was used as the axis X, and the results of the experimental method were used as the axis Y, plotted with a slope of 1.0 across the origin (Figure 1B). The reproducibility and correlation between the two systems were checked within the detection range. The results showed good reproducibility and correlation for both methods.

The Bias Results of the Measurement Results of the Comparison Instrument and Experimental Instrument

Figures were plotted with the adoption of (the mean of the experimental method − the mean of the comparison method) versus (the mean of the experimental method + the mean of the comparison method)/2 (Figure 2A) and (the observed results of the individual experimental method − the observed results of the individual comparison method) versus (the mean of the experimental method
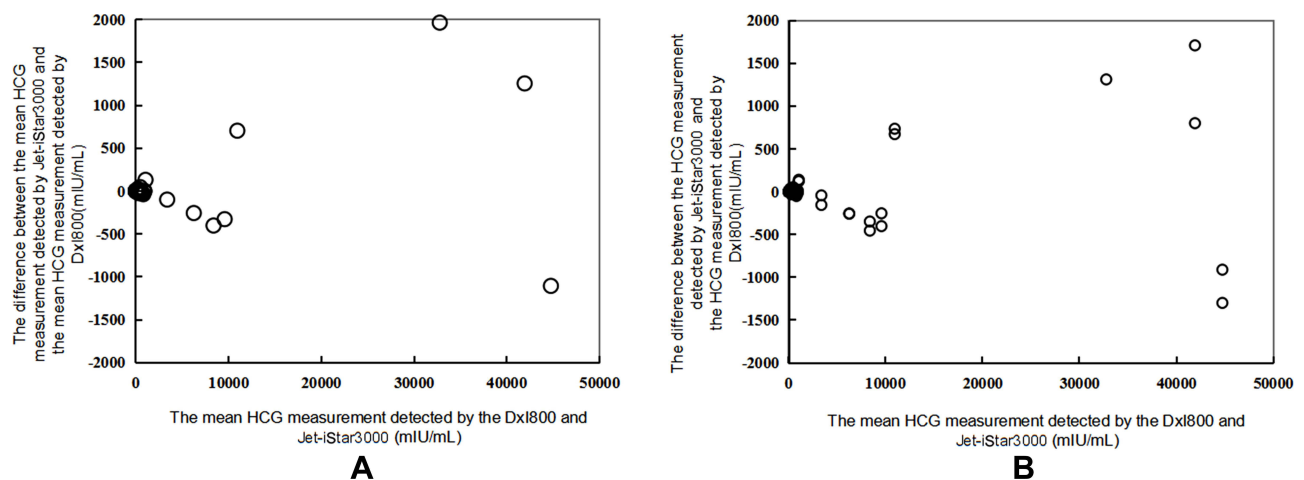

Figure 2 The scatter plot of measurement deviation by the DxI800 and Jet-iStar3000 (A) (the mean of the experimental method − the mean of the comparison method) versus (the mean of the experimental method + the mean of the comparison method)/2. (B) (the observed results of the individual experimental method − the observed results of the individual comparison method) versus (the mean of the experimental method + the mean of the comparison method)/2.

Table 1 Evaluation of the Clinical Acceptability of the HCG Results by Different Assay System (mIU/mL)

Medical Determination Level	Estimated Bias	Acceptable Bias
5	0.426	0.625
25	2.962	3.125
400	3.893	5
10,000	98.175	125

+ the mean of the comparison method)/2 (Figure 2B) to observe the bias. The results showed a relatively small difference.

The Linear Regression Equation

Using the Dxl 800 as the target detection system, the estimated correlation and regression analysis of the Jet-iStar 3000 detection system were conducted, and the regression equation was obtained as y = 1.020x + 12.96, r = 0.998.

Judgment of the Clinical Acceptability in Performance

The expected deviations Bc of HCG at different medical decision levels were calculated separately and compared with the determined systematic errors (SE%) ≤12.5%. These showed that the expected deviations Bc of the assay results were within the acceptable range, and the experimental method was accepted (Table 1).

Discussion

HCG is a glycoprotein dimer composed of subunits α and β. The subunit α is common for the anterior pituitary hormones, while the subunit β is specific to HCG and can accurately reflect HCG levels in humans. In some rare cases the mutation in subunits α or ß, and Heterophilic antibodies produced different false-negative results on serum hormones by immunoassay methods. The complete HCG is produced entirely by the syncytial trophoblast of the placental chorionic villi.[6,7] HCG is the most commonly used "pregnancy test" hormone in clinical practice and is important for early pregnancy diagnosis. It is valuable for the diagnosis, differentiation, and observation of the course of pregnancy-related diseases, trophoblastic tumors, and other diseases. With the deepening of the clinical application, rapid detection of HCG is an important clinical demand. An immunofluorescence analyzer (the POCT method) can rapidly detect HCG to meet

the needs of the clinical practice and patients. However, the detection performance needs to be validated regularly. Otherwise, the widely varying results will bring many problems to the clinical application.[8,9] It is necessary to conduct a regular comparative analysis of the results between different systems to ensure the accuracy of the assay and the comparability of the results between different assay systems. There are many comparison methods between different assay systems, but EP9-A2 is the classic and reliable comparison method.[10,11] Therefore, in the present study, the EP9-A2 was used to compare with the Jet-iStar 3000 immunoassay analyzer to provide reliable, accurate, and consistent results for clinical practice.

According to the EP9-A2 protocol, the two HCG assay systems in the present department were compared and analyzed together to evaluate the bias.[10] The Beckman Dxl 800 participated in the national and provincial quality assessments and achieved excellent results. Thus, it was used as the comparison method and the "standard HCG detection system" in our department. The Jet-iStar 3000 immunoassay analyzer was used as the experimental method. The bias between the measurement results of the two systems was within a controllable range. There was a good correlation through the linear regression analysis. The average value of HCG detected by the Jet-iStar 3000 was slightly higher than the detection level by the Dxl 800, and the reproducibility of the Jet-iStar 3000 detection was also slightly inferior to the Dxl 800. The reason may be correlated with the respective detection principles and instrument performance. Therefore, when the individual specialties have higher requirements for HCG, an independent reference range could be established for the Dxl 800 and Jet-iStar 3000, respectively. Although the accuracy of the results by the Jet-iStar 3000 assay was slightly lower than that by the Dxl 800, the Jet-iStar 3000 assay was simple and fast, which was suitable for emergency or POCT assays. The Dxl 800 was suitable for daily specimen assays. In the present comparison, the sample selection was strictly implemented following EP9-A2 protocol, and fresh samples were selected instead of quality control products to avoid the matrix effects. The Jet-iStar 3000 immunofluorescence analyzer and Dxl 800 had different upper detection limits, so we chose the Jet-iStar 3000 immunofluorescence analyzer with a lower detection limit (50,000 mIU/mL). When the concentration of the specimen is >50,000 mIU/mL, manual dilution is required, which will cause greater errors in the results. Therefore, the upper limit of the concentration in the present study

was set at 50,000 mIU/mL. In daily work, for samples with a concentration >50,000 mIU/mL, it could be first tested on the Jet-iStar 3000 immunofluorescence analyzer and reported orally to the clinic for emergency treatment. It could then be tested on the Dxl 800 to issue a formal test report.

The CLSI EP9 document has been revised in several versions and has been an important guidance document for clinical laboratories performing methodological comparisons and bias assessments. The EP9-A2 in 2002 is the classic version, with features such as being easy to operate and generating reliable results. In 2013, CLSI EP9-A2 was updated to CLSI EP9-A3, which mainly specifies the classification in the number of samples and introduces ESD for outlier testing, with more scientific results and diversified statistical methods. Although the application is broader and the protocol design is more reasonable, the operation is complicated, and the statistical difficulty increases. It makes it more difficult to promote it in the majority of primary hospitals. Therefore, in this paper, the EP9-A2, which is more commonly used in primary hospitals, was used to compare the HCG results of the two different instruments. Although the results are slightly inferior to those of the EP9-A3, the comparison results are sufficient to meet the requirements of clinical laboratories and provide a sample for the standardized comparison in primary hospitals.

The results of the present study were similar to those of Xu et al,[12] who adopted EP9-A to evaluate the Abbott i2000 chemiluminescence immunoassay analyzer and the Merieux VIDAS fluorescence enzyme-labeled immunoassay analyzer to determine the serum subunit β of HCG. There was a certain deviation in HCG measurement by different detection systems, but the results were all within the allowable range. The results of the present study were consistent with and comparable to the results of Yuan Xiaohua,[13] who investigated the immunofluorescence dry quantitative POCT method and electrochemiluminescence method (ECLIA) parallel detection of blood concentration of HCG. Many studies have shown that the emergency detection of HCG had a good correlation with reference methods since it had a relatively narrow linear range and greater fluctuations in results compared with reference methods. However, emergency detection of the HCG system could only be used as a clinical qualitative standard, which was not suitable for routine monitoring.

In summary, the Jet-iStar 3000 and Dxl 800 had good consistency in the performance of the HCG assay, and the

results were highly comparable. The instruments should be maintained and serviced in strict accordance with the relevant requirements to achieve standardized operation. The bias of HCG detection by both systems was within the acceptable range, and the test results could be used in clinical practice. The comparison between systems may be necessary to ensure the accuracy and consistency of the assay results in the same laboratory and better serve clinical practice.

Ethics Approval and Consent to Participate

The study was conducted in accordance with the Declaration of Helsinki (as was revised in 2013). The study was approved by Ethics Committee of the Xinchang Hospital of Traditional Chinese Medicine. We confirm that the patients were informed about the purpose of the study. A written consent was obtained from each patient.

Disclosure

The authors declare that they have no competing interests.

References

1. Schug S, Baunacke A, Goeckenjan M, et al. Endometrial human chorionic gonadotropin (hCG) expression is a marker for adequate secretory transformation of the endometrium. *Arch Gynecol Obstet*. 2019;299(6):1727–1736. doi:10.1007/s00404-019-05130-y
2. Rahimkhani M, Khavari-Daneshvar H, Sharifian R. Asymptomatic bacteriuria and pyuria in pregnancy. *Acta Med Iran*. 2008;46 (5):409–412.
3. Wang Z, Gao Y, Zhang D, Li Y, Luo L, Xu Y. Predictive value of serum β-human chorionic gonadotropin for early pregnancy outcomes. *Arch Gynecol Obstet*. 2020;301(1):295–302. doi:10.1007/s00404-019-05388-2
4. Nwabuobi C, Arlier S, Schatz F, Guzeloglu-Kayisli O, Lockwood CJ, Kayisli UA. hCG: biological functions and clinical applications. *Int J Mol Sci*. 2017;18(10):2037. doi:10.3390/ijms18102037
5. The National Committee for Clinical Laboratory Standards. *Method Comparison and Bias Estimation Using Patient Samples; Approved Guideline*. Second ed. EP9-A2; 2002.
6. Hadinata IE, Doyle LW, Thompson D, Reti L. Serum β-hCG levels post-treatment of ectopic pregnancy with a single dose of intramuscular methotrexate. *Aust N Z J Obstet Gynaecol*. 2015;55(2):181–184. doi:10.1111/ajo.12316

7. Thompson BW, Zerden ML, Morse JE. Abnormally elevated serum hCG in a patient with end-stage renal disease seeking abortion: to be expected or a cause for concern? A case report. *Contraception*. 2015;92(1):84–86. doi:10.1016/j.contraception.2015.03.014

8. Lu WB. Predictive value of serum hCG, PROG, E2 level in early pregnancy for fetal suspension. *Gansu Med J*. 2018;37(10):889–890. (Article in China).

9. Jiang SH. Evaluation of the combined detection of serum HCG, β-HCG and progesterone in the diagnosis of ectopic pregnancy. *Chin Community Doctors*. 2018;34(3):112, 114. (Article in China).

10. National Committee for Clinical Laboratory Stardards. *EP9 A2 Method Comparison and Bias Estimation Using Patient Samples: Approved Guide-Line*. Wayne, PA, USA: NCCLS; 2004.

11. Rahimkhani M, Kazemian K, Daryasari RR. Falsely undetectable TSH in a euthyroid patient. *Thyroid Res*. 2020;13(1):2. doi:10.1186/s13044-020-0076-y

12. Xu YW, Wang LJ, Zhu J. Comparison and bias assessment of serum β-HCG results measured by different assay systems. *Chin J Rural Med Pharm*. 2018;25(4):55–56. (Article in China).

13. Yuan XH, Yang CJ, Hu HP, et al. The performance analysis of immunofluorescence POCT method to detect the blood HCG. *J Clin Hematol*. 2015;28(5):875–877. (Article in China).