An Evaluation of 10 Kits for Determination of Human Choriogonadotropin in Serum

Joan L. Raso,1 Sergio Farber, and Glenn D. Braunstein

We compared 10 commercially available radioimmunossay kits (American Diagnostics, Becton Dickinson, BioGenex, Clinical Assays, Hybritech, Leeco, Mallinckrodt, Microanlytic, Nuclear Medical Systems, and Radioassay Systems) for determination of human choriogonadotropin (hCG), using serum pools, hCG CR119 (NIH), 2nd I.S. (WHO), and 1st I.R.P. (WHO). Criteria were ease of performance, total assay time, sensitivity, potency, and parallelism as compared with reference standards and results for 15 serum pools. The Mallinckrodt kit exhibited the best overall performance, with good low-concentration sensitivity, parallelism with two of the three reference preparations, and good clinical correlation as compared with the reference kit from NIH. Because the antibodies used in the kits are occasionally changed by the manufacturers, these results are necessarily valid only for kits that include reagents identical to those in the kits that we tested.

Additional Keyphrases: pregnancy tests • radioimmunossay • "kit" methods

In early immunological pregnancy tests antibodies generated against intact human choriogonadotropin (hCG) were used, which generally did not discriminate between hCG and other structurally similar glycoprotein hormones (1). hCG and the other glycoproteins—luteinizing hormone (LH), thyrotropin, and follitropin—contain two non-covalently linked, dissimilar subunits, designated alpha and beta (2). Vaitukaitis et al. (3) found that a highly specific antibody for measuring hCG in serum could be produced by using beta-hCG as the immunogen. This antisera (SB-6) could detect intact hCG in serum without interference from even high physiological concentrations of hLH (4).

During the last five years, manufacturers of pregnancy-determination kits have been able to produce antibodies against hCG or beta-hCG that are more sensitive and specific, as a result of a better understanding of the biochemical and structural natures of the hCG molecule and development of monoclonal antibody techniques. Because these kits are used extensively for the diagnosis of pregnancy and pregnancy disorders and for monitoring the course of gestational and nongestational trophoblastic diseases, we have studied many of the commercially available kits that allegedly measure serum hCG accurately and specifically and have compared the antisera in these kits with the NIH reference antibody, SB-6. This report contains our evaluation of hCG kits from 10 manufacturers.

Materials and Methods

Materials

Commercial hCG radioimmunossay kits. The following kits were supplied gratis by the manufacturers for testing: "Beta-CG III" (American Diagnostics Corp., Newport Beach, CA 92663), "h-hCG RIA Kit" (Becton Dickinson Immunodiagnostics, Orangeburg, NY 10962), "RiaGen" (BioGenex Laboratories, Dublin, CA 94568), "GammaDab" (Clinical Assays, Cambridge, MA 02139), "Tandem" (Hybritech Inc., San Diego, CA 922121), "Concept-7-β-hCG" (Leeco Diagnostics, Inc., Southfield, MI 48034), "RIA-Quant" (Mallinckrodt, Inc., Immunoassay Systems, St. Louis, MO 63134), "hCG-Beta" (Microanalytic Research, Inc., Laguna Hills, CA 92653), "(hCG-β) RIA Kit" (Nuclear Medical Systems, Inc., Newport Beach, CA 92663), and "Quant-Preg" (Radioassay Systems Laboratories, Inc., Carson, CA 90746).

Reference method. The reference hCG "kit" consisted of an antiserum against beta-hCG (SB-6) and a highly purified urinary preparation of hCG (CR119) and pituitary hLH (LER 907), all of which were kindly provided by the Hormone Distribution Officer for the National Hormone and Pituitary Program, the Center for Population Research of the NICHD and NIADDKD, NIH, Baltimore, MD 21201.

hCG reference standards. The Second International Standard (2nd I.S., 61/6) and the First International Reference Preparation (1st I.R.P., 75/537) were supplied by the World Health Organization (WHO) through the Programme for the National Biological Standards Board, National Institute for Biological Standards and Control, London, England.

Other supplies. 125I (as sodium iodide, 100 Ci/L) was obtained from Amersham-Searle, Arlington Heights, IL 60005.

Polystyrene and polypropylene 12 × 75 mm test tubes purchased from VEM Scientific, Inc., Los Angeles, CA 90010, were used in the tests when indicated. Glass 12 × 75 mm test tubes purchased from American Scientific Products, Irvine, CA 92713, were also used.

Anti-rabbit gammaglobulin second antibody was obtained from Antibodies Inc., Davis, CA 95616.

All assay tubes were counted in a Model 4/200 gamma counter from Micromedic Systems, Horsham, PA 19044.

Methods

Kit methods. We followed the instructions provided by the manufacturers for the respective kits.

NIH kit. The procedure involving CR119 as standard and tracer and SB-6 as first antibody is described elsewhere (4–5).

Data reduction. For each kit except the one from Hybritech, all calculations were based on iterative, weighted, least-squares regression of the log-logit transformation of the dose–response data on a WANG 2200 system, using a program modeled after Rodbard (6). Since 100% binding
could not be determined for the Hybritech kit, all calculations were based on linear regression analysis of log-log transformations of the dose–response data. Log-log transformations of the Hybritech data corresponded well with data read off the linear graph supplied by the manufacturer (y = 0.856x − 8.426, r = 1.00). We did tests for parallelism with a DEC PDP-11, using the hypothesis that there is no difference between the slopes tested. Dose–response curves were considered parallel if the p values exceeded 0.05.

Results

Kit specifications. Table 1 contains procedural information for each of the 10 kits tested, as well as the quoted sensitivities and expected ranges. Most (80%) of the kits tested were calibrated by the manufacturer against the WHO 2nd International Standard (2nd I.S., 61/6). Primary incubation intervals for quantitative determinations varied from 90 min to overnight. Shorter incubation times (30–45 min), used for qualitative determinations, resulted in less sensitivity to hCG (25–40 int. units/L). Most of the kits required incubation intervals of at least 1.5 to 3 h for maximum sensitivity (2–5 int. units/L). In three of the kits—those from Becton Dickinson, Hybritech, and Mallinkrodt—an hCG-free human serum was used in the standards.

Parallelism studies. With all kits, dose–response lines for hCG (CR119), 2nd I.S., 1st I.R.P., and hCG (LER 907) were tested for parallelism with the kit standard dose–response line. Three of the kits (from Becton Dickinson, Leeco, and Radioassay Systems) demonstrated nonparallelism with hCG (CR119). Of the eight kits calibrated against the WHO 2nd I.S., only Mallinkrodt’s and Leeco’s exhibited nonparallelism with this standard. Of the two kits calibrated against the WHO 1st I.R.P., only Clinical Assays’ kit showed nonparallelism with the reference preparation. The kits from Hybritech, Microanalytic, and NMS demonstrated comparable slopes with all of the gonadotropin preparations. With the NIH kit, hCG (CR119) exhibited parallelism with both of the WHO preparations but not with hCG (LER 907). Five kits (those from Clinical Assays, Hybritech, Microanalytic, NMS, and Radioassay Systems) showed parallelism with hLG (LER 907).

Sensitivity and specificity. The equivalent dose at 50% displacement (ED50) for kit standards and reference preparations, 80% binding dose, and “minimum detection level” ([B0 – 2 SD]/B0% binding dose) for each kit are listed and compared in Table 2.

Cross reactivity with hLG in all the kits was tested, with use of a pooled specimen of postmenopausal sera (PMP) that contained 29.6 int. units of hLG per liter and a partly purified preparation provided by NIADDK (LER 907). Of the kits tested, five showed no inhibition for PMP at >95% B/B0 (American Diagnostics, BioGenex, Hybritech, Mallinkrodt, and Microanalytic). The rest, including the NIH kit, showed inhibition for PMP between 80 and 94% B/B0. This is consistent with the degree of inhibition of the purified hLG exhibited by these kits, as can be seen in Table 2. It should be noted that the NIH kit has been the standard for comparison for many years (4) and the <0.2% cross-reactivity with hLG of that antisera was considered acceptable. All of the kits tested with the partly purified hLG showed less cross-reactivity (ED50 > 1000 ng) than that seen with the NIH kit (ED50 = 606 ng).

To compare sensitivities at 80% binding dose for the kits tested, we calibrated all kit standards within their respective assays against hLG (CR119). As can be seen in Table 2, based on one kit tested per manufacturer, the lowest 80% binding dose was obtained with the kit from Nuclear Medical Systems; however, this kit also produced the flattest slope. The highest 80% binding dose was obtained with the kit from BioGenex, which also produced the steepest slope. Two kits, those from Microanalytic and Radioassay Systems, had similar ED50 for CR119 as compared with the NIH kit but had higher 80% binding doses, because in these assays 100 µL of serum sample is used, whereas the sample volume for the NIH kit was 200 µL. Clinical evaluation. Fifteen serum samples with hLG concentrations ranging from 1.4 to 573 µg/L were measured in duplicate by each of the test methods. In order that the results for the patients’ samples could be compared for all kits, we recalculated the data from each kit, using hLG.

Table 1. Kit Specifications as Supplied by the Manufacturers

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Calibration standard</th>
<th>Primary incubation</th>
<th>Sensitivity</th>
<th>Range (Int. units/L)</th>
<th>Specimen type and vol, µL</th>
<th>LH cross reactivity (standard)</th>
<th>Type of first antibody</th>
<th>Type of label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Am. Diag.</td>
<td>2nd I.S.</td>
<td>0.5 h, 37 °C 1 h, 25 °C</td>
<td>1.2</td>
<td>5–160</td>
<td>serum, 200</td>
<td>&lt;1% (LER 907)</td>
<td>Rabbit, xhLG-beta</td>
<td>12S-hLG</td>
</tr>
<tr>
<td>B.D. Immunod.</td>
<td>2nd I.S.</td>
<td>2 h, 25 °C</td>
<td>2.8</td>
<td>5–200</td>
<td>serum, plasma, 200</td>
<td>&lt;0.23 (NG)</td>
<td>Rabbit, xhLG-beta</td>
<td>12S-hLG</td>
</tr>
<tr>
<td>BioGenex Lab.</td>
<td>2nd I.S.</td>
<td>5–12 h, 25 °C</td>
<td>1–3</td>
<td>5–250</td>
<td>serum, 200</td>
<td>&lt;0.1 (NG)</td>
<td>Murine monoclonal xhLG-beta</td>
<td>12S-hLG</td>
</tr>
<tr>
<td>Clinical Assays</td>
<td>1st I.R.P.</td>
<td>1.5 h, 25 °C</td>
<td>3</td>
<td>5–200</td>
<td>serum, 200</td>
<td>1.2 (NG)</td>
<td>Rabbit xhLG-beta</td>
<td>12S-hLG</td>
</tr>
<tr>
<td>Hybritech</td>
<td>2nd I.R.P.</td>
<td>2–4 h, 37 °C</td>
<td>1.5</td>
<td>5–400</td>
<td>serum, 100</td>
<td>NG (NG)</td>
<td>Murine monoclonal xhLG-beta</td>
<td>12S-anti-xhLG</td>
</tr>
<tr>
<td>Leeco Diag.</td>
<td>1st I.R.P.</td>
<td>1–2 h, 37 °C</td>
<td>NG</td>
<td>2.5–100</td>
<td>serum, plasma, 200</td>
<td>&lt;0.5 (NG)</td>
<td>Rabbit xhLG-beta</td>
<td>12S-hLG</td>
</tr>
<tr>
<td>Mall. Imm. Sys.</td>
<td>2nd I.S.</td>
<td>3 h, 25 °C</td>
<td>1.5</td>
<td>1.5–100</td>
<td>serum, 200</td>
<td>0.11 (NG)</td>
<td>Rabbit xhLG-beta</td>
<td>12S-hLG-beta</td>
</tr>
<tr>
<td>Microa. Res.</td>
<td>2nd I.S.</td>
<td>3.5 h, 25 °C</td>
<td>1.0</td>
<td>3.12–100</td>
<td>serum, plasma, CSF, 100</td>
<td>NG (NG)</td>
<td>Rabbit xhLG-beta</td>
<td>12S-hLG</td>
</tr>
<tr>
<td>Nuc. Med. Sys.</td>
<td>2nd I.S.</td>
<td>2.5 h, 25 °C</td>
<td>1</td>
<td>3.12–100</td>
<td>serum, 100</td>
<td>&lt;0.7 (hLG IRP-1)</td>
<td>Rabbit xhLG-beta</td>
<td>12S-hLG</td>
</tr>
<tr>
<td>Rad. Sys. Lab.</td>
<td>1st I.R.P.</td>
<td>2 h, 37 °C</td>
<td>2.45</td>
<td>2.45–122.5</td>
<td>serum, plasma, 100</td>
<td>&lt;1.0 (NG)</td>
<td>Rabbit xhLG-beta</td>
<td>12S-hLG</td>
</tr>
</tbody>
</table>

NG, not given.
(CR119) as the calibration standard instead of the standard supplied with each kit. hCG (CR119) was chosen for calibration even though three of the kits had shown nonparallelism, because it is measured by mass rather than by international units, which are based on biological activity. As can be seen in Figure 1, after linear regression analysis, two kits demonstrated good agreement with the NIH kit. Those from Leeco and Mallinckrodt had the best regression slopes and the lowest y-intercepts.

When the clinical data for each kit were calculated in terms of the WHO 2nd I.S. and compared with CR119 in terms of the 2nd I.S. by linear regression analysis, 60% of the kits tested produced regression slopes of <0.80. Only two kits (American Diagnostics and Becton Dickinson) had regression slopes near 1.00. However, they also had y-intercepts >7.0.

**Discussion**

During 15 months (February 1982 through April 1983) we tested 10 commercial kits for the determination of hCG in serum. The criteria we kept in mind were sensitivity, specificity, speed, and simplicity. No attempt was made to check all the commercial kits for hCG that are currently on the market, but some attempt was made to at least evaluate those kits that are most widely used in other clinical laboratories in our area. Only one kit from each manufacturer was tested except for American Diagnostics (two), Leeco (two), Hybritech (four), and Nuclear Medical Systems (two) in this preliminary study. Additional kits were requested from American Diagnostics, Leeco, and Nuclear Medical Systems because the testing protocol had changed for the study after these kits had been tested initially. Technical difficulties with the performance of the Hybritech kits required more kits for testing.

![Table 2. Sensitivity and Specificity of the Kits](https://example.com/table.png)

**Fig. 1. Correlation between serum hCG values as determined with the NIH kit and serum hCG values as determined by each of the commercial kits**

Linear regression analysis: \( y = mx + b \). Regression coefficients \( r \) for all correlations >0.96

Other groups have published evaluations of various kits manufactured since 1980 (7–9). In these studies, the kits...
were intercompared but not with the NIH kit or the WHO calibration standards. In our study, we have used three reference standards for hCG diluted with normal male plasma. Most of the kits we tested were calibrated against the WHO 2nd I.S., the rest against the WHO 1st I.R.P. The WHO 1st I.R.P. was prepared from highly purified hCG (batch CR119) and has been assigned a value of 650 int. units per ampoule in comparison with WHO 2nd I.S. in various types of bioassays and receptor assays (10). Attempts to measure immunological potency of the 1st I.R.P. have been unsuccessful, owing to the impurity of the 2nd I.S. Until an acceptable standard for measuring hCG by immunoassays can be found, it has been recommended (10) that all calibrations continue to be based on the 2nd I.S., with use, if necessary, of a factor converting 1st I.R.P. to 2nd I.S. Because highly purified hCG is available from NIH and who, we recommend that all results be reported in “mass units” to obviate the difficulties inherent with the use of standards based on biological activity.

Table 3 summarizes all the criteria used in testing. Only one of the 10 kits tested scored high marks for all these. The kit from Mallinkrodt had good low-end sensitivity at the 80% binding dose (1.4 µg of hCG (CR119) per liter or 6 int. units/L according to the manufacturer of the kit), demonstrated parallelism with hCG CR119 and who 1st I.R.P., and demonstrated good clinical correlation against the NIH kit. The kit from Leeco also performed well in the clinical correlation category, but it failed the parallelism tests against all of the reference standards.

Before deciding on an appropriate kit for use in the laboratory, one should be aware that the manufacturers are continually making changes in the kit specifications as existing antibody supplies are depleted, “better” antibodies are produced, or the method of separating bound from free hormone is improved. One other note of caution concerns whether a result can be said to be indicative of pregnancy. Our group and others have demonstrated the presence of hCG in normal tissues, urine, and plasma from non-pregnant individuals (11); thus low titters of hCG activity (5–20 int. units/L in serum relative to the 2nd I.S.) could be present without a concomitant pregnancy. We routinely report results in the 5–20 int. units/L range as "borderline" and request a second specimen a week later. Samples with titers in this range are "positive" for hCG but are not necessarily "positive" for pregnancy. If the individual is pregnant, the subsequent sample should show a significantly increased titer (12).

We thank each of the manufacturers for donating the commercial kits for testing. We also thank the Hormone Distribution Officer, NIH, Bethesda, MD, and the National Institute for Biological Standards and Control, London, England, for the reference standards and preparation.

References