Three Commercial Immunoradiometric "Kit" Assays for Serum Ferritin Evaluated

C. D. R. Dunn1,2 and D. J. Boden1

We evaluated three commercially available immunoradiometric assays for serum ferritin, with particular emphasis on the statistical validity of the results. The data show that it was unusual (a) for dose/response relationships of the standards to be linear over the whole concentration range suggested by the manufacturers and (b) for dose/response relationships of test sera to parallel those of the standards. These findings cast doubt on the ability of any of the assays reliably to discriminate small differences in ferritin concentrations. Nevertheless, all three methods gave reproducible results, but this reflects technical expertise rather than accuracy of the results per se. The method based on liver ferritin as tracer detected less of the iron-binding protein than those based on spleen ferritin, despite significant cross-reactivity of the two antibodies.

Additional Keyphrases: "kit" methods - "sandwich" immunoradiometric methods

Ferritin was first recognized as a normal, physiological constituent of human serum in 1972 (1). Since then it has come to be recognized that concentrations of ferritin in serum accurately reflect body-iron stores in several clinical situations, particularly those involving changes in iron metabolism (see, for example, reference 2 for a recent review). To satisfy the demand for what is perceived as an informative clinical test, several commercial assays for ferritin are now available. The object of this study was to evaluate three immunoradiometric assays (IRMA) for serum ferritin, with particular emphasis on statistical performance. The data show that none of the three assays was entirely satisfactory: the relation of log concentration to response frequently deviated significantly from linearity, and test sera generally produced log concentration/response relationships with slopes that differed significantly from those of the standards. These findings cast doubt on the ability of these assays accurately to discriminate differences in ferritin concentrations between samples. The reproducibility of all the assays, as determined from repetitive measurements at a standard dilution, was within acceptable limits.

Materials and Methods

We used IRMA kits from three manufacturers: Fer-Iron® (Ramco Labs Inc., Houston, TX 77008), Immuno Phase Fer® (Corning Medical and Scientific, Medfield, MA 02052), and Ventre/Sep® (Ventrex Labs Inc., Portland, ME 04103). The methodologies recommended by the manufacturers in the package inserts were closely followed throughout. All three assays are based, either explicitly or implicitly, on the two-site "sandwich" IRMA described by Addison et al. (1). However, there are some potentially important technical differences, which are summarized in Table 1. The assays from Ramco and Corning both use human spleen ferritin as a tracer, whereas Ventrex utilizes human liver ferritin. The sources of the antibodies differ from manufacturer to manufacturer, as do the methods for separating free from bound antibody. Recommended incubation intervals are generally comparable, but longer alternative periods are detailed for the Ramco assays.

Human sera from 20 subjects, none with malignant disease, were used. When assayed according to the manufacturer's recommendations, they had ferritin concentrations of 4.2–860 (Ramco), 1.6–700 (Corning), and 5.3–860 µg/L (Ventrex). These sera were selected because preliminary assay results suggested their ferritin titers were on the linear part of the log concentration/response relationship of the standard. All dilutions of test sera were made with the buffer provided (when the sera were to be assayed with the Ramco or Ventrex methods) or with isotonic saline (for assay with the Corning technique). Complete standard curves were constructed in each assay run according to the manufacturer's specifications.

Standard statistical methods were used to analyze differences between mean results (two-way analysis of variance) and to evaluate correlation between assays. One-way analyses of variance were used to define the quality of standard and test sera log concentration/response relationships, both individually and by comparison (3, 4). Tests for deviations from linearity were based on at least three consecutive concentrations (or dilutions) (4). Results differing at p ≤ 0.05 were considered significant.

Results

Preliminary investigations were directed at determining the reproducibility of results obtained when a sample was run repeatedly in the same assay. A serum selected without conscious bias was assayed 10 times in all assays, both with and without a 10-fold prior dilution. The results are summarized in Table 2. Coefficients of variation (CV) varied between 2% (Ramco) and 5.9% (Ventrex) for measurements made on prediluted serum, or between 3.6% (Ramco) and 9.1% (Ventrex) on the same serum assayed without prior dilution. Two-way analysis of variance of the results summarized in Table 2 followed by a post-test comparison by use of Tukey's Honestly Significant Different method (5) demonstrated that the mean ferritin titers showed significant differences both between assays and between dilutions. Furthermore, the presence of a significant interaction indicated that the differences between titers in prediluted and undiluted sera were not constant across manufacturers. Thus, the apparent ferritin concentrations were a function of the particular assay and of the dilution used in that assay. The quality-control samples consistently gave results within the ranges predicted by the manufacturers.

Figure 1 shows representative log concentration/response relationships for the standards in each assay. Results from a serum (different from that used for the results in Table 2, but representative of all 20 investigated) studied in all three assays after a range of predilutions are also documented. Note that

---

1 Life Sciences Laboratory, Northrop Services Inc., P.O. Box 34416, Houston, TX 77034. (Use this address, to C. D. R. D., in correspondence.)
2 Division of Experimental Biology, Baylor College of Medicine, 1200 Moursund Avenue, Houston, TX 77030.

Received March 17, 1981; accepted April 20, 1981.
the results are graphically represented according to the manufacturer's suggestions, i.e., log concentration/logit response for the Ramco and Ventrex assays and log concentration/linear response for the results obtained with the Corning system.

With the Ramco assay (Figure 1A), the log concentration/response relationships for the standard showed significant deviations from linearity when the results of all six recommended concentrations were employed in the analysis. However, no significant deviations from linearity were observed over the concentration range 5-500 μg/L. The slopes of the standard relationships over this concentration range varied between 2.50 and 2.65. Test sera, with concentrations greater than approximately 10 μg/L (when assayed on the one-dilution basis recommended by the manufacturer), produced log concentration/response relationships which did not demonstrate significant deviations from linearity over at least a three-dilution range. However, the log concentration/response relationships of the test sera always showed significant deviations from parallelism to that of the standard.

When the results of the Corning assay (Figure 1B) were expressed as a linear response all possible ranges of three consecutive standard concentrations, or of test sera dilutions, produced log concentration/response relationships which showed significant deviations from linearity. Deviations from linearity were reduced, but were still significant, if the data were recalculated on the basis of log concentration/logit response.

Only over the range 50-1000 μg/L—the four highest concentrations used in constructing the standard curve—did the log concentration/response relationship to the ferritin standard in the Ventrex assay (Figure 1C) not show significant deviations from linearity. The slope of the log concentration/response relationship over this range varied between 1.50 and 1.77. Test sera with ferritin concentrations exceeding approximately 25 μg/L (one-dilution assay) produced log concentration/response relationships, over at least a three-dilution range, not differing significantly from linearity but which showed significant deviations from parallelism to that of the standard.

The results from the 20 test sera, when analyzed in all three assays according to the one-dilution method described by the manufacturers, showed highly significant correlations between methods. Correlation coefficients (r) varied between 0.9788 (Corning vs Ventrex) and 0.9554 (Ramco vs Ventrex). Relative efficiencies of the three assays were 1.00/0.96/0.90 for Ramco/Corning/Ventrex, respectively. To investigate the difference in efficiency in detecting serum ferritin between the Ramco and Ventrex assays, "crossover" studies were performed whereby the Ventrex tracer (human liver ferritin) was reacted with the Ramco antibody and the Ramco tracer (human spleen ferritin) was reacted with the Ventrex antibody. Considerable cross reactivity was observed, with 90% binding of the Ramco antibody to 1000 μg/L Ventrex antigen and 70% binding of the Ventrex antibody to 2000 μg/L Ramco antigen. The Ventrex antigen/Ramco antibody relationship did not show significant deviations from linearity over the ferritin concentration range 10-5000 μg/L and had a slope of 3.06. In contrast, the Ramco antigen/Ventrex antibody relationship did not show significant deviations from linearity.

Table 1. Some Differences in the Kit Manufacturers' Claimed Specifications for the Immunoradiometric Assay of Serum Ferritin

<table>
<thead>
<tr>
<th>Tracer source</th>
<th>Ramco</th>
<th>Corning</th>
<th>Ventrex</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Human spleen</td>
<td>Human spleen</td>
<td>Human liver</td>
</tr>
<tr>
<td>Separation of bound and free antibody</td>
<td>No centrifugation</td>
<td>Centrifugation</td>
<td>No centrifugation</td>
</tr>
<tr>
<td>Phase I antibody</td>
<td>Rabbit anti-human spleen</td>
<td>Goat anti-human?</td>
<td>Goat anti-?</td>
</tr>
<tr>
<td>Phase II antibody</td>
<td>Rabbit anti-human spleen</td>
<td>Rabbit anti-?</td>
<td>Rabbit anti-?</td>
</tr>
<tr>
<td>Controls</td>
<td>Three sera provided</td>
<td>Two provided (not sera)</td>
<td>Two sera provided</td>
</tr>
<tr>
<td>Phase I incubation</td>
<td>2.0 h/room temp.</td>
<td>0.5 h/room temp.</td>
<td>1.0 h/room temp.</td>
</tr>
<tr>
<td>Phase II incubation</td>
<td>2.0 h/room temp.</td>
<td>2.0 h/room temp.</td>
<td>1.0 h/room temp.</td>
</tr>
<tr>
<td>Claimed intraspecimen CV (%)</td>
<td>7.0-7.8</td>
<td>3.6-6.6</td>
<td>4.5-6.8</td>
</tr>
<tr>
<td>Claimed interspecimen CV (%)</td>
<td>3.7-10.6</td>
<td>4.2-6.6</td>
<td>6.1-9.3</td>
</tr>
<tr>
<td>Claimed minimum detectable dose, μg/L</td>
<td>1.4</td>
<td>1.6</td>
<td>4.8*</td>
</tr>
<tr>
<td>Recommended correction for nonspecific binding</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Graphic representation of results</td>
<td>Log concn/ logit response</td>
<td>Log concn/ linear response</td>
<td>Log concn/ logit response</td>
</tr>
</tbody>
</table>

* The Ventrex-Sep assay allows for a modification of the basic technique in which the minimum detectable concentration is claimed to be 1.0 μg/L.

Table 2. Ferritin Concentration in One Serum, Diluted or Undiluted

<table>
<thead>
<tr>
<th>Ferritin concn., μg/L (and CV, %)</th>
<th>Ramco</th>
<th>Corning</th>
<th>Ventrex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undiluted</td>
<td>290.7 ± 3.3</td>
<td>326.3 ± 7.7</td>
<td>234.2 ± 6.7</td>
</tr>
<tr>
<td></td>
<td>(3.6)</td>
<td>(7.5)</td>
<td>(9.1)</td>
</tr>
<tr>
<td>10-fold dil.</td>
<td>327.0 ± 2.1</td>
<td>373.0 ± 6.2</td>
<td>188.0 ± 3.5</td>
</tr>
<tr>
<td></td>
<td>(2.0)</td>
<td>(5.2)</td>
<td>(5.9)</td>
</tr>
</tbody>
</table>

* Concentration is reported as mean of 10 assays ± standard error of the mean. Results for the diluted serum are corrected for the dilution factor. For comparison between assays and between dilutions, Tukey's Honestly Significantly Different Coefficient = 22.15 for p ≤0.05; i.e., any mean differing from another by >22.15 is significant at p ≤0.05.
over the ferritin concentration of 50–1000 μg/L and had a slope of 1.39. The slopes of these two relationships showed significant deviations from parallelism, both from one another and from the non-"crossover" standard curves.

**Discussion**

A basic principle in the measurement of any material that cannot be quantitated on the basis of a physical property (e.g., weight or volume) is that dose/response relationships to a generally acceptable standard and to the test substance not show significant deviations from linearity or from parallelism to one another (4, 6). (Indeed, it has been stated that the standard in such assays should be especially selected to produce dose/response relationships parallel to those of the test samples). Only if these criteria are met (together with others related to the slopes of the relationships, which were easily fulfilled in the current evaluations) can the potency of the test solution be stated with any degree of certainty and its biological (or immunological) similarity to the standard material confirmed. In practice, dose/response relationships to both standard and test materials are seldom constructed in a routine clinical environment, for a variety of reasons. Nevertheless, these basic fundamentals of bio- or immunoassays are assumed in the one-dose "tests" (4) commonly used. The studies reported herein emphasize the care with which the results of one-dose "tests" should be interpreted in the absence of details regarding dose/response relationships.

In all of the assays evaluated the standard curves generally showed significant deviations from linearity over the concentration range recommended by the manufacturers. This problem could be overcome, at least in the Ramco and Ventrex assays, by omitting the lowest and/or highest concentrations from the calculations (Figure 1). This has the result that accurate data will be obtained only over a narrower concentration range than implied by the manufacturers. Test sera frequently produced dose/response relationships which did not differ significantly from linearity. However, parallelism of these relationships to those of the standards was such a rare occurrence that it would appear that our findings, while based on a relatively small sample size (n = 20), are more generally applicable. One possible explanation for these deviations from parallelism is that human serum ferritin is antigenically distinct from the human spleen or liver ferritins used as tracers in the assays. The demonstration of several iso-ferritins in human serum makes this hypothesis plausible (7, 8). Alternative explanations involve the interference of other serum proteins in the antigen–antibody reaction. In our experience the combination of significant deviations from linearity and from parallelism result in an error of approximately ±50 μg/L (for a one-dose “test”) in sera with ferritin concentrations normalized to 300 μg/L (i.e., ±17%; see also Table 2). With 95% confidence limits typically ranging from 80 to 125% of the expected value (100%) (4, 10) and technical reproducibility on the order of 5% (Table 2), it can be calculated that it is improbable that ferritin concentrations in two samples differing by less than two-fold could be considered statistically significant. This may not be particularly critical, given the wide range of "normal" serum ferritin concentrations (10), although data of doubtful statistical validity from the one-dose "tests" may have contributed to this apparent wide range. However, it casts doubt on the ability of the commercially-available assays to discriminate small differences in ferritin concentrations. This conclusion may be particularly important clinically in distinguishing two sera each of which have low ferritin titers. As titers approach the sensitivity threshold of the assay, confidence limits of the measurements are generally larger. Thus a greater difference between two measured titers would be necessary for statistical certainty. In this context the reproducibility of the measurements (Table 2) and acceptable data from the assay of the quality-control samples reflects technical expertise rather than accuracy of the results per se. The generally smaller CV with the Ramco as compared with the Corning or Ventrex assays is probably a result of the steep dose/relationships observed with this kit.

In these studies the so-called high-dose "hook" effect (11) was not demonstrated. However, it is another potential problem in many bio- (4) and immunoassay systems (11). Because it is not possible to distinguish whether the one-dose "test" response is on the ascending or descending limb of the

---

dose/response relationship, errors in apparent potency of 10-fold or greater could be encountered (4, 11).

Despite the potential variability, a good correlation was demonstrated between the results of the three assays (current study) and between the Ramco (9) and the Hoechst (12) assays and standard laboratory methods. However, both previous studies (9, 12) demonstrated that the commercially-available assays detected significantly less ferritin than standard laboratory procedures. The present results (Table 2) confirm previous observations (12, 13) that assays based on liver ferritin as a tracer (Ventrex) give lower concentrations for serum ferritin than those in which spleen ferritin is used as the tracer (Ramco and Corning). The reasons for this are unclear, particularly as human spleen ferritin antibody strongly cross-reacts with human liver ferritin, and vice versa. This observation supports the manufacturer's claim and has been independently demonstrated previously (13). In the present studies, however, the "crossover" dose/response relationships seemed to retain the characteristics of the antibody (particularly in terms of concentration ranges and slopes) much more so than of the antigen.

Supported by contracts NAS 9-16180 and NAS 9-15425 with the National Aeronautics & Space Administration. We are grateful to Dr. P. R. Swank for assistance with the statistical analyses.

References