Three Commercial Methods for Serum Ferritin Compared and the High-Dose "Hook Effect" Eliminated

Ronald H. Ng,1 Becky A. Brown,2 and Roland Valdes, Jr.2,3

We evaluated four commercial kits for measuring serum ferritin, based on three techniques: immunoradiometric assay, radioimmunoassay, and enzyme immunoassay. The kits evaluated were those manufactured by Abbott Laboratories, Clinical Assays, Corning Medical, and Ramco. Two of the immunoradiometric kits showed satisfactory results with respect to sensitivity and precision; they should be useful in diagnosing individuals with uncomplicated iron-deficiency anemia. One of the immunoradiometric assay kits, however, showed a high-dose "hook effect," beginning at 10 mg of ferritin per liter. We modified this kit to eliminate this effect, at least to ferritin concentrations of 33 mg/L. (We observed a ferritin value as high as 47 mg/L in one patient.) Results with all these kits did not inter-compare well for ferritin concentrations >300 μg/L, a finding that casts further doubt on the controversial use of serum ferritin measurement in cases of iron overload.

Additional Keyphrases: iron overload • anemia • iron metabolism • "kit" methods • hemochromatosis • monitoring therapy with iron during hemodialysis • displacement radioimmunoassay • enzyme immunoassay • immunoradiometric assay

Recent advances in sensitive methods for measuring serum ferritin have aroused considerable interest, serum ferritin having been shown useful in assessing clinical disorders of iron metabolism (1, 2). Serum ferritin concentration may reflect the size of body iron stores (3, 4), and a low serum ferritin concentration has a high predictive value for the diagnosis of uncomplicated iron-deficiency anemia (5). However, data on serum ferritin concentration are of less value in conditions in which it is inappropriately increased with respect to reticuloendothelial iron stores, e.g., in anemia associated with inflammation, liver disease (6), leukemia, or hematological malignancies (7). In these conditions, a low ferritin concentration is indicative of iron-deficiency anemia, but a normal or above-normal concentration does not rule out this condition.

In addition, data on serum ferritin concentration may be useful in detecting iron overload. Although of possible value in early detection and monitoring treatment of patients with idiopathic hemochromatosis (1), a primary iron-overload disorder, the measurement of ferritin lacks clinical specificity in the presence of active hepatocellular disease, which may be a complication resulting from the iron overload (1).

Serum ferritin assays are also of value in monitoring therapy with iron in patients on maintenance hemodialysis, to avoid iron overload (8).

We have inter-compared four widely used assay kits, which together represent all of the three current techniques used in ferritin determination: displacement radioimmunoassay, enzyme immunoassay, and immunoradiometric assay. Because the finding of a low concentration of ferritin is useful in the diagnosis of iron-deficiency anemia, we evaluated the sensitivity and reproducibility of these kits for low ferritin concentrations. We also compared kits for ease of assay performance, precision, sensitivity, analytical recovery, sample-dilution response, and correlation of patient results. Furthermore, because a high-dose "hook effect," if present, would yield erroneously low results in cases of iron overload, we studied the effects of high concentrations of serum ferritin on the assay validity of each kit.

Materials and Methods

We evaluated the following kits: Gamma Dab (Clinical Assays, Cambridge, MA 02139), an RIA kit; Ferrizyme (Abbott Laboratories, North Chicago, IL 60064), an enzyme immunoassay; and Fer-iron (Ramco Labs Inc., Houston, TX 77098) and Immuno Phase (Corning Medical and Scientific, Medfield, MA 02052), both of which are two-site immunoradiometric ("sandwich") assays. We used the procedures as described by the manufacturers in the respective package inserts. The "Corning II" assay kit is our modification of the routine Corning kit (Corning I), as follows: (a) single 1-h incubation instead of two incubations (0.5 h and 2 h), (b) 10-μL sample volume instead of 50 μL, and (c) incubation at 37°C instead of room temperature.

For data reduction we used computerized analysis with the following curve-fitting routines: Abbott, linear point-to-point regression; Clinical Assays, weighted logit-log regression; Corning and Ramco, spline fitting routines. Radioactivity was counted with a Micromedic 4/200 automatic gamma counter. For the Abbott assay, absorbance at 492 nm was measured with a Quantum 1 spectrophotometer (Abbott Laboratories).

Results

Sensitivity. To determine assay sensitivity (the lowest concentration distinguishable from zero), we analyzed the zero calibrators from each kit 20 times within one assay run. Sensitivity was calculated as the mean ±2 SD of the number of counts or the absorbance, and the value was then extrapolated from the standard curve. The greatest sensitivity was seen in the low-range assay with the Clinical Assays kit (Table 1). The "low-range assay" is a procedure recommended by Clinical Assays for measuring ferritin concentrations below 5 μg/L, whereas the "high-range assay" is for measuring ferritin in concentrations of 5 to 500 μg/L. The Corning I and II assays were twice as sensitive as that of Abbott. Ramco and Clinical Assays high-range assays were intermediate in sensitivity.

---

1 Division of Laboratory Medicine, Departments of Pathology and Medicine, Washington University School of Medicine, St. Louis, MO 63110 (present address: Department of Laboratory Medicine, University Hospital, Boston University Medical Center, Boston, MA 02118).

2 Department of Pathology and Laboratory Medicine, The Jewish Hospital of St. Louis, St. Louis, MO 63110.

3 Address correspondence to this author.

Received Dec. 27, 1982; accepted Mar. 29, 1983.
Table 1. Sensitivity and Precision of the Kits Compared

<table>
<thead>
<tr>
<th>Kit supplier</th>
<th>Sensitivity, ( \mu g/L )</th>
<th>Within-assay Precision</th>
<th>Between-assay Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean, CV, ( % )</td>
<td>Mean, CV, ( % )</td>
<td>(n = 20)</td>
</tr>
<tr>
<td>Abbott</td>
<td>4.4</td>
<td>15.8, 10.4</td>
<td>11.7, 8.7</td>
</tr>
<tr>
<td>Clinical Assays</td>
<td>3.4&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6, 86.7</td>
<td>7.4, 12.4</td>
</tr>
<tr>
<td>Comming I</td>
<td>2.0</td>
<td>7.7, 7.7</td>
<td>8.1, 5.4</td>
</tr>
<tr>
<td>Comming II</td>
<td>1.9</td>
<td>11.5, 6.3</td>
<td>11.3, 10.8</td>
</tr>
<tr>
<td>Ramco</td>
<td>2.7</td>
<td>5.8, 16.4</td>
<td>14.2, 5.3</td>
</tr>
</tbody>
</table>

<sup>High-range method</sup>
<sup>Low-range method</sup>

**Precision.** We determined within-run precision (CV) with three concentrations of ferritin in pooled serum samples from patients, and between-run precision with the controls provided with each kit (Table 1). The Comming I and II assays showed the smallest overall imprecision, with CVs ranging from 3.3 to 7.7%; Abbott and Ramco had somewhat higher CVs. Clinical Assays, despite satisfactory precision at the high ferritin range, showed high variations at low concentrations of serum ferritin. The CV for this assay remained high even when we used the more sensitive low-range method. To estimate between-assay precision, we assayed in duplicate the respective control material from each kit on six different days. The five kits had comparable between-assay CVs, the Clinical Assays kit showing slightly more variability than the other four.

**Recovery.** In the analytical-recovery study, aqueous ferritin standards (20–600 \( \mu g/L \)) provided with the respective kits were added to pooled patients’ serum (ferritin concentration, 100–160 \( \mu g/L \)) and also to another aqueous standard from the same kit. As Table 2 shows, more than 90% of the ferritin added was accounted for in most cases.

**Dilution.** In the dilution study, we diluted serum and standards with the zero calibrator from the respective kits. Nonparallelism was observed only with the Abbott Kit, for which linear regression of the data for diluted standards showed that ferritin found \( = (1.01 \times \) ferritin expected \( + 15 \mu g/L \) for serum samples, ferritin found \( = (0.80 \times \) ferritin expected \( + 85 \mu g/L \). The other four assays had a much better dose–response relationship for dilutions of either serum sample or standard material (data not shown).

**Patient comparison.** Figure 1 shows comparisons of results for patients’ samples as measured with each assay. When only ferritin concentrations below 300 \( \mu g/L \) (n = 29) were compared, the correlations improved to the following: Abbott = 1.40 (Ramco) \( + 4.62, r = 0.9727; \) Clinical Assays = 1.27 (Ramco) \( - 3.98, r = 0.9894; \) Comming I = 0.95 (Ramco) \( + 1.08, r = 0.9710; \) and Comming II = 1.08 (Ramco) \( + 7.77, r = 0.9741. \)

**Hook effect.** We tested two separate serum samples from a patient who subsequently died of multiple complications after an acute myocardial infarct. This patient had laboratory test results consistent with liver disease: aspartate aminotransferase 2220 U/L, total bilirubin 42 mg/L. The ferritin concentrations were notably high as measured by all the assays, except that the Comming I assay gave falsely low results when the sample was analyzed without dilution (cf. serum samples A and B in Table 3). After examining the dose–response curve produced by serial dilution of this sample with zero calibrator from the same kit (Figure 2), we determined this was due to a high-dose hook effect (12). This effect appeared to begin at approximately 10 mg/L for this assay. Between 10 and 30 mg/L, the result was given as >1000 \( \mu g/L \), the radioactivity detected exceeding that of the most concentrated standard (1000 \( \mu g/L \); at ferritin concentrations >30 mg/L, results were falsely low. The Comming II procedure gave substantially different results, in that no hook effect was detected at ferritin concentrations up to at least 33 mg/L.

**Discussion.** The clinical use of serum ferritin in the diagnosis of uncomplicated iron-deficiency anemia has been well established by other studies (1, 2), a ferritin concentration of <10 \( \mu g/L \) usually being interpreted as indicative of iron-deficiency anemia (9). Also, the lower limits of reported normal reference intervals are as low as 10 \( \mu g/L \) (10) and 4 \( \mu g/L \) (11) for men and women, respectively. Therefore, the sensitivity and the precision of a method at low concentrations of ferritin (<20 \( \mu g/L \)) become very important in selecting an assay kit. Except for the low range of the Clinical Assay kit, the Corning kit (either procedure) showed the highest sensitivity (2.0 \( \mu g/L \)) and therefore had the widest working range. It also showed slightly better precision than the other kits. The sensitivity as determined in our laboratory corresponded with that reported by the respective manufacturers in their package inserts, except for the enzyme immunoassay kit: Abbott's claim was 0.75 \( \mu g/L \) vs our finding of 4.4 \( \mu g/L \). Imprecision (CV) at ferritin concentrations <10 \( \mu g/L \)
ferritin per liter was quite high with the Clinical Assays kit, even when the low-range method was used. We believe that this can be ascribed in part to the small difference in radioactivity detected between the zero and the 20 μg/L ferritin standard (i.e., a decrease of only 3000 cpm or 10%, from $B_0$) and in part to the heteroscedasticity inherent in displacement-type radioassays (16).

In general, analytical recovery was acceptable. With the
Clinical Assays and Corning kits, somewhat less ferritin in aqueous samples added to serum samples was accounted for than in the case of an aqueous sample added to another aqueous sample—evidently a matrix effect of the serum proteins (12).

Excepting the Abbott assay, the other four kits showed parallelism in the dilution study. Nonlinearity in dilution has been reported in the Corning I assay for standards and patients' sera (13). We also noticed a slight deviation from linearity, but only in one of the several patients' samples tested with the Corning I assay. No apparent deviation from linearity was seen for other serum samples with the Corning I or Corning II assays. These observations can probably be explained by the well-recognized finding that serum ferritin is not immunogenically homogeneous (7, 10).

In the comparison study with patients' samples, the Abbott assay produced much higher results than the Ramco assay. Clinical Assays was slightly higher than Ramco, and Corning I and II gave results comparable with those of Ramco. The correlations, in general, were much better if comparison was made up to only 300 μg/L, owing to the scatter of data at higher concentrations. We also found (data not shown) that the ferritin standards from the four kits could not be interchanged because they gave different results by the different assays. Thus, variation in standard calibration, in addition to antibody specificity and affinity, may account for the discrepancies at concentrations exceeding 300 μg/L. Serum ferritin is immunogenically similar to ferritin from liver and spleen (13), the two sources from which the ferritin used in the standards was derived and used in preparation of the antisera. However, these isoferritins are not identical (1) and small discrepancies in the affinity of antibodies to certain isoferritins in serum can account for unexpected findings in pathological conditions in which the concentration of serum ferritin is increased. Nevertheless, correlations were satisfactory for ferritin concentrations up to 300 μg/L, a range that includes concentrations found in both normal persons and those with uncomplicated iron-deficiency.

In considering overall assay validity, one must remember that various factors have been suggested (12, 14) as possibly affecting ferritin assays: isoferritin composition in sera, antibody specificity, character of the tracer (125I-labeled ferritin or 125I-labeled antibody), ferritin standards (tissue source and calibration), and interference by other serum proteins in the antigen–antibody reaction.

Although it is well known that two-site immunoradiometric assays can exhibit the high-dose hook effect (12), to our knowledge this is the first report that demonstrates this effect in one of these commercial ferritin kits and the first report of a serum ferritin concentration of 47 mg/L, which exceeds the value of 27.6 mg/L reported by Prieto et al. (15) in 1975.

We have subsequently used the Corning assays to measure ferritin in serum from 10 patients with extremely high values for aminotransferase activity and total bilirubin. All 10 had above-normal ferritin concentrations, and the two patients who had the highest concentrations showed the hook effect with the Corning I assay kit. Table 3 shows results for four of the 10 patients (sera C–F), including the two that exhibited a hook effect, and for the patient who first showed the hook effect (sera A and B). All three patients with the hook effect had aminotransferase activities >1000 U/L. Thus, in the Corning I assay, for concentrations exceeding 10 mg/L the result is unpredictable. On the other hand, the Corning II assay apparently is not susceptible to the hook effect at ferritin concentrations at least up to 33 mg/L.

In general, the high-dose hook effect is believed to be caused by the presence of antibodies of various avidities (12). Thus, the binding of antigen to low-affinity solid-phase antibody in the first reaction and subsequent exchange of the antigen from the solid phase into a soluble complex with 125I-labeled antibody during the second reaction may reduce the number of radioactive antibodies available for binding to insolubilized antigen.

Serum ferritin concentrations rarely exceed 10 mg/L in iron overload except in the presence of severe liver disease, such as massive liver necrosis or acute hepatitis (7). In such cases measuring concentrations of ferritin in serum to determine body iron stores is hardly useful, because the increase in ferritin is not particularly related to iron stores. The high-dose hook effect with the Corning I assay at ferritin concentrations greater than 10 mg/L could thus lead to potential misinterpretation of assay results if one is not aware of the limitations (e.g., liver disease) of the serum ferritin assay.

In conclusion, we found that the immunoradiometric assays (Ramco and Corning) produced acceptable results and should be useful in distinguishing individuals with uncomplicated iron-deficiency from normal subjects. The Corning kit, however, showed greater sensitivity and better precision, and the assay also was easier to perform. Results
with none of the four kits compared well with each other for ferritin concentrations >300 μg/L. Thus, the problem of monitoring iron overload with serum ferritin measurements may be further complicated by discrepancies among commercial kits.

References

Effects of Whole Blood Storage on Results for Glycosylated Hemoglobin as Measured by Ion-Exchange Chromatography, Affinity Chromatography, and Colorimetry
Randie R. Little, Jack D. England, Hsiao-Mei Wiedmeyer, David E. Goldstein

After storage of whole blood at either 4 or 20 °C, results for glycosylated hemoglobin by ion-exchange chromatography ("high-performance" liquid and mini-column chromatography), thiobarbituric acid colorimetry, and affinity chromatography were compared. At 4 °C, all methods gave acceptable results for samples stored for as long as a week. At 20 °C, the colorimetric and affinity methods also showed sample stability for a week or more. The ion-exchange methods were associated with a marked increase in values for glycosylated hemoglobin after a few days of storage. Evidently, care in details of sample collection and handling is especially important for ion-exchange methods, and the colorimetric and affinity methods have advantages over ion exchange in situations where long delays between sample collection and assay are unavoidable.

Additional Keyphrases: screening · diabetes

Glycosylated hemoglobin assay is a clinically useful index of long-term blood glucose control in diabetes mellitus (1--3). Several assay methods have been described, but little attention has been given to effects of sample handling and storage before processing and analysis, although these factors may be of great importance for the clinical laboratory (4, 5).

Here we present data on the stability of glycosylated hemoglobin in whole blood stored for as long as 15 days before hemolysate preparation and analysis for glycosylated hemoglobin. We compare results obtained by several different assay methods: ion-exchange chromatography ("high-performance" liquid (HPLC-IE) and mini-column chromatography), thiobarbituric acid (TBA) colorimetry, and affinity chromatography.

Materials and Methods

Blood samples. All specimens were obtained by venipuncture, from five Type I (insulin-dependent) diabetic patients and from five non-diabetic subjects. All diabetic patients were being treated with two daily injections of insulin. Blood specimens were collected into EDTA-containing evacuated blood-collection tubes and promptly refrigerated (4 °C). Within 2 h of collection, all specimens were divided and stored for various periods of time up to 15 days at either 4 or 20 °C in capped 12 × 75 mm tubes.

Preparation of hemolysates. At designated intervals, whole-blood samples were mixed thoroughly by inversion and aliquots removed for preparation of hemolysates as previously described (6). Briefly: for each specimen, plasma and leukocytes were removed after centrifugation (800 × g, 10 min). Erythrocytes were washed three times with five volumes of isotonic saline and incubated in 12 volumes of

Received Jan. 13, 1983; accepted Mar. 7, 1983.

1 Department of Pathology and 2 Department of Child Health, University of Missouri-Columbia, Health Sciences Center, Columbia, MO 65212.