Regional Interlaboratory Standardization of Determinations of Cholesterol, High-Density Lipoprotein Cholesterol, and Triglycerides

Gregory A. Tetrault,1,3 W. Greg Miller,1,4 Vernon M. Chinchilli,2 and others

The Clinical Chemistry Forum of Central Virginia initiated a lipid standardization program to help ensure that its members meet the current National Cholesterol Education Program guidelines for cholesterol testing, and to standardize assays of high-density lipoprotein (HDL) cholesterol and triglycerides so as to provide accurate lipid profiles. We found that freshly collected, never-frozen human sera must be used to assess interlaboratory accuracy for cholesterol, HDL cholesterol, and triglycerides assays, and that at least 23 samples are required to detect a 3% bias with 90% power when the between-laboratory imprecision (CV) is 3%. After recalibration, all 12 laboratories had a mean cholesterol bias ≤5%, nine of 10 laboratories had a mean HDL cholesterol bias ≤40 mg/L for samples with values ≤570 mg/L, and 10 of 12 laboratories had a mean triglycerides bias ≤10% for fresh human sera split between participants and the Centers for Disease Control. Pools of frozen human serum were shown to have matrix biases >3% for cholesterol in seven of 11 laboratories, and >40 mg/L for HDL cholesterol in six of nine laboratories.

In 1987 the National Cholesterol Education Program (NCEP) Adult Treatment Panel released guidelines for concentrations of serum cholesterol and correlation of these values with coronary heart disease (1). The standardized epidemiological studies upon which the current NCEP guidelines are based were possible because of the development of the National Reference System for Cholesterol, which includes the Definitive Method at the National Institute of Standards and Technology and the Reference Method at the Centers for Disease Control (CDC). To apply the NCEP risk-evaluation cut points, laboratories must obtain assay results standardized to the accuracy base maintained by the National Reference System for Cholesterol. Lipid reference intervals determined by individual laboratories, and therefore peculiar to individual methods, are no longer acceptable. Accurate, precise measurements of cholesterol, HDL cholesterol, and triglycerides are necessary if consistent estimates of LDL cholesterol are to be obtained with the Friedewald formula (2).

The NCEP Laboratory Standardization Panel has issued guidelines for analytical performance of individual laboratories, which state that cholesterol bias must be <5% from the national Reference Method maintained by the CDC and that imprecision (CV) be <5% (3). These allowances are to be decreased to <3% by 1992. Accuracy and imprecision goals for HDL cholesterol and triglycerides have not yet been issued. For this investigation, we set accuracy goals of bias ≤40 mg/L for HDL cholesterol and <10% for triglycerides, and imprecision goals of ±30 mg/L (SD) for HDL cholesterol and ±5% (CV) for triglycerides.

The Clinical Chemistry Forum of Central Virginia is an organization of clinical laboratories whose purpose is to enhance the practice of clinical chemistry in the community. To meet the goals of the NCEP and to provide consistent, reliable lipid testing in the community, the Forum established a lipid-standardization program that included community hospitals ranging in size from 150 to 500 beds, a 700-bed Veterans Administration hospital, a 1000-bed university hospital, and an out-of-state 900-bed teaching hospital.

We report our findings here.

Materials and Methods

Frozen human serum pools, two lots of lipid proficiency-monitoring material (LPMM1 and LPMM2), and one concentration of secondary standard (CAL1) were prepared by collecting blood from healthy donors (into plastic blood-donor bags for LPMM1; into glass bottles for LPMM2, CAL1, and CAL2), allowing it to clot, centrifuging, and pooling the serum. The other secondary standard (CAL2) was prepared by mixing pooled serum (as described above) with a low-and very-low-density lipoprotein fraction obtained by ultracentrifuging pooled, outdated plasma obtained from a blood bank. We over laid 18 mL of plasma with 12 mL of a 1.20 kg/L NaBr solution (final solution density 1.063 kg/L) in 30-mL sealed tubes, ultracentrifuged for 18 h at 100 000 × g at 16 °C, removed the upper 5 to 6 mL from each tube, and dialyzed it against 150 mmol/L NaCl for 24 to 48 h. The dialyzed material (345 mL total volume) was added to 870 mL of pooled normal serum. We dispensed 1-mL aliquots of each pool into polypropylene cryogenic vials and stored them at −70 °C. These products were distributed packed in solid carbon dioxide at six-month intervals, and most laboratories stored them at −20 °C. Each frozen pool was assayed by the CDC with its Reference Methods for cholesterol (4–6), HDL cholesterol (6–8), and triglycerides (9, 10).
Each round of fresh, never-frozen, human-donor serum samples was prepared by sampling, from eight apparently healthy donors, five to nine 10-mL evacuated tubes of blood with no anticoagulant. Some, but not all, donors had fasted overnight. After clot formation, the tubes were centrifuged at 1500 × g for 10 min. The serum from each individual donor tube was mixed, and 1- to 1.5-mL aliquots of each donor’s serum were prepared and distributed at 4 °C to all participating laboratories within 30 h after collection. The samples were stored at 4 °C and assayed by individual laboratories within 48 h of receipt.

Twelve laboratories6 participated in the lipid standardization program, using the instruments and methods listed in Table 1. All cholesterol and triglycerides assays were enzymatic. Only one laboratory (no. 13) used a free-glycerol blank in its triglycerides assay.

Results

Imprecision was measured by assaying the frozen pool LPMM1 for five months before calibration adjustment. The median cholesterol CV was 2.3% (range 1.4 to 4.2%), with 11 of 12 laboratories ≤2.5%. The median HDL cholesterol CV was 4.3% (range 2.4 to 7.2%), with eight of nine laboratories ≤5.6%. The median triglycerides CV was 3.1% (range 1.3 to 9.2%), with 10 of 12 laboratories ≤4.7%.

The initial calibration status of each laboratory was assessed by measuring CAL1 and CAL2 in duplicate on six separate days and comparing the mean values with the values assigned by the CDC. Cholesterol biases were −2.8% to 9.8%, with six of 11 laboratories >3%, and three of 11 laboratories >5%. HDL cholesterol biases were −98 to −4 mg/L, with six of eight laboratories <−40 mg/L. Laboratories with unacceptable biases recalibrated their cholesterol and HDL cholesterol methods by use of the frozen pools, and all achieved results within the goals of the program. Triglycerides biases for the frozen calibrators were 5.8% to 47%, with 10 of 11 laboratories >10% owing to high concentrations of free glycerol in both secondary standards. The blood to prepare CAL1, CAL2, and LPMM2 was collected into glass containers that were found to have glycerol lubricated stoppers. Only one laboratory used a method with free-glycerol blank correction.

During the six months after recalibration of cholesterol and HDL cholesterol based on frozen pools, five rounds of freshly collected, never-frozen human serum samples were distributed to participating laboratories and several CDC-certified laboratories, to verify laboratory calibration. Three laboratories had, in at least two split-sample rounds, mean cholesterol biases <−5%, and one had a mean HDL cholesterol bias >40 mg/L. Because of this difference in results between fresh sera and frozen pooled sera, and because of the free-glycerol limitations of the frozen materials, all laboratories readjusted their methods to give correct results with the fresh serum samples.

After the developmental phase of the program was complete, two rounds of fresh, never-frozen human serum samples, spaced one week apart, were split among all participants and the CDC. Results from these two rounds were combined to determine the mean percent biases for cholesterol, HDL cholesterol, and triglycerides. Figure 1 shows that all 12 laboratories achieved mean cholesterol accuracy within ±5% of the CDC value, and eight of 12 were within ±3%. Nine of 10 laboratories achieved mean HDL cholesterol accuracy within ±40 mg/L of the CDC value for samples with values <570 mg/L. However, only six of 10 laboratories achieved mean HDL accuracy within 40 mg/L of the CDC when values >570 mg/L were included. HDL cholesterol bias plots (Figure 2) show that four of the five laboratories that used dextran sulfate precipitation methods had progressively larger negative biases at higher HDL cholesterol concentrations. Laboratory 7 had a constant positive HDL cholesterol bias of about 50 mg/L, which it has corrected by calibration adjustment.

Figure 1 also shows that 10 of the 12 laboratories achieved mean triglycerides values within ±10% of the CDC values. Two laboratories, using the Astra system, had constant positive biases of 12% to 19%, which they have since corrected by calibration adjustment. Standardization of triglycerides results is complicated by the lack of sample-blank correction for free glycerol in most methods. Some fresh donors’ sera had results for triglycerides with biases as high as 23% in laboratories with mean biases <5%.

The principal technical difficulty identified in this calibration program was the existence of a matrix bias in the frozen human serum pools when compared with fresh, never-frozen human serum samples. Figure 3 shows, for each laboratory,

<table>
<thead>
<tr>
<th>Lab. no.</th>
<th>Cholesterol</th>
<th>HDL cholesterol</th>
<th>Ptn method</th>
<th>Triglycerides</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Spectrumb</td>
<td>Spectrum/Abbott</td>
<td>DS500</td>
<td>Spectrum</td>
</tr>
<tr>
<td>5</td>
<td>Ektachem 400c</td>
<td>E-400/Kodak</td>
<td>DS50</td>
<td>E-400</td>
</tr>
<tr>
<td>6</td>
<td>Astrap</td>
<td>—</td>
<td>—</td>
<td>Astra</td>
</tr>
<tr>
<td>7</td>
<td>Hitachi 737*</td>
<td>Hitachi 737/Gilford'</td>
<td>DS500</td>
<td>Hitachi 737</td>
</tr>
<tr>
<td>8</td>
<td>Dimensiong</td>
<td>aca Vg</td>
<td>PW</td>
<td>Dimension</td>
</tr>
<tr>
<td>9</td>
<td>Parmaxh</td>
<td>aca IIIg</td>
<td>PW</td>
<td>Parmax</td>
</tr>
<tr>
<td>11</td>
<td>aca IVp</td>
<td>—</td>
<td>aca IV</td>
<td>—</td>
</tr>
<tr>
<td>12</td>
<td>Dimension</td>
<td>Dimension</td>
<td>PW</td>
<td>Dimension</td>
</tr>
<tr>
<td>13</td>
<td>Cobas Bio/BMD</td>
<td>Cobas Bio/BMD</td>
<td>HM</td>
<td>Cobas Bio/Behrungk</td>
</tr>
<tr>
<td>15</td>
<td>Astra</td>
<td>TDXp</td>
<td>DS50</td>
<td>Astra</td>
</tr>
<tr>
<td>18</td>
<td>Hitachi 705*</td>
<td>aca IV</td>
<td>PW</td>
<td>aca IV</td>
</tr>
<tr>
<td>19</td>
<td>Ektachem 700c</td>
<td>E-700</td>
<td>DS50</td>
<td>E-700</td>
</tr>
</tbody>
</table>

* Abbreviations: DS50, dextran sulfate, 4, 80,000; DS600, dextran sulfate, M, 500,000; HM, heparin–manganese; PW, phosphotungstate. b Abbott Laboratories, Irving, TX. c Eastman Kodak, Rochester, NY. d Beckman Instruments, Brea, CA. e Boehringer Mannheim/Hitachi, Indianapolis, IN. f HDL-Cholesterol Reagent; Gilford Systems, Oberlin, OH. g DuPont Co., Clinical Systems Division, Wilmingtom, DE. h Baxter, Paramax Systems Division, Santa Ana, CA. i Roche Diagnostic Systems, Nutley, N.J. j High Performance K" cholesterol reagents; Boehringer Mannheim Diagnostics, Indianapolis, IN. k Stat-Pack" Enzymatic Triglycerides Test; Behring Diagnostics, Somerville, NJ.
The mean percent biases are for two rounds of eight samples, n = 16. The error bars indicate the standard error of the mean. HDL results were separated into two groups: ≤570 mg/L (n = 6) and >570 mg/L (n = 10).

The bias of frozen samples expressed as the difference between the mean cholesterol percent bias or HDL cholesterol bias determined with frozen pools of human serum vs that determined with fresh donor sera. The frozen pools were assayed by each laboratory during the same time interval as the two rounds of fresh serum samples split with the CDC. The target values assigned to the frozen pools were confirmed by the CDC during this time, and the mean biases for the frozen products and the fresh serum samples were determined vs the CDC-assigned values. Seven of 11 laboratories had cholesterol bias differences ≥3%, and four of 11 had bias differences ≥5% between frozen pools and fresh, never-frozen serum samples. Both Astra systems, which used the same lot of reagents, showed a negative bias for two frozen pools and a positive bias for the other frozen pool with an apparent total difference of 5% to 6%. The two Ektachem and two Dimension systems used the same lot of reagents, respectively, but each instrument showed somewhat different performance with frozen materials. The two Hitachi analyzers used different reagent lots but showed similar performance with frozen materials.

Six of nine laboratories had HDL cholesterol bias differences ≥40 mg/L between some frozen pools and fresh serum samples. The two Ektachem instruments used the same lot of cholesterol reagents, and the two aca instruments used different lots of cholesterol reagents. The matrix bias could not be evaluated for triglycerides due to the high concentration of free glycerol in the frozen pools, and the fact that most methods do not correct for free glycerol in the sample.

Discussion

Our results demonstrate that a group of laboratories that use a variety of instrument and reagent systems can achieve nearly equivalent results for cholesterol, HDL cholesterol, and triglycerides. Two critical elements are needed to achieve accuracy among such a group of laboratories. The chemical procedures used must specifically and completely convert all the analyte in the sample to the product that is actually measured. Calibration materials must also be available that have the same chemical reactivity with all instrument and reagent systems, and thus can be used to transfer accuracy between routine methods and the Reference Method. Fresh, never-frozen human sera must be used to establish and confirm accuracy for cholesterol, HDL cholesterol, and triglycerides in routine clinical laboratories.

The matrix biases seen in frozen human serum pools have important implications for implementing standardization programs on a large scale. It is logistically impossible to split a fresh, never-frozen human serum sample between more than a few individual laboratories. In addition, donors with a range of lipid concentrations must be available to provide the serum samples. No frozen or lyophilized material has been developed that is suitable for use with a variety of analytical systems. Lyophilized pooled serum materials exhibit variable and unpredictable matrix effects for cholesterol and other analytes (11, 12). Boersma et al. (13) have shown that use of human-serum-based frozen pools can improve interlaboratory standardization of cholesterol, but that there are commutability problems with some methods. Koch et al. (14) demonstrated that
where $\sigma^2$ is the greater variance of the two methods at the mean concentration value, $\Delta x$ is the difference in mean values, $t_{N-1, 1-(a/2)}$ is the $t$ value for a two-sided test at $\alpha$ level of significance, and $t_{N-1, 1-\beta}$ is the $t$ value for a one-sided test at $\beta$ level of significance. Because $t_{N-1, 1-(a/2)}$ and $t_{N-1, 1-\beta}$ are a function of $N$, $z_{1-(a/2)}$ and $z_{1-\beta}$ from a normal-curve table are used first to get an estimate of $N$. Then $t_{N-1, 1-(a/2)}$ and $t_{N-1, 1-\beta}$ are used to get the second estimate of $N$. This process is continued until convergence is attained. At a cholesterol concentration of 2000 to 2400 mg/L, precision (CV) of 3%, 0.05 significance level, and 90% power, 10 samples are required for one to detect a 5% bias, and 23 samples for one to detect a 3% bias. In our protocol, a set of eight split samples can detect a cholesterol bias of 5% with 82% power. Combining two sets of eight split samples enables detection of a 3.7% bias with 90% power. The number of samples required to detect a given bias would decrease if the CV was <3%. For example, at CV = 2%, 12 samples are required to detect a 3% bias with 90% power. There are logistical and cost limitations to the number of samples—whether fresh, frozen, or lyophilized—that can be distributed among laboratories. The sample sizes required to detect a 3% bias may make it difficult to assess whether that accuracy goal for cholesterol testing can be met by routine laboratories.

The current protocol for our standardization program relies on fresh, never-frozen human serum samples to monitor and transfer accuracy between laboratories. We are currently monitoring method calibration based on two rounds, one to two weeks apart, of eight fresh serum samples split between participants and a CDC-certified laboratory. Bias plots are used to report the accuracy of individual laboratories, and a cumulative report summarizes performance each quarter. Once a method is correctly calibrated, each laboratory assays frozen pooled sera used as secondary standards and assigns instrument/reagent-specific target values to them. These secondary standards are then used within an individual laboratory to assist with method calibration when reagent or instrument calibrator lots are changed.

We offer the following recommendations for establishing and maintaining lipid standardization programs in different operating environments. A group of laboratories using several different instrument/reagent systems can be standardized as described in this report. The number of participants will be limited by the above-mentioned logistics. Triglyceride methods that do not correct for free glycerol will be limited by the requirements to account for a "typical" free-glycerol value when the instrument calibration is adjusted. Standardization of triglycerides would be simplified if manufacturers would incorporate sample-blank correction technology into future generations of instrumentation.

Individual laboratories can assess and confirm accuracy by sending aliquots of fresh donor specimens to a CDC-certified or network laboratory. Three rounds of eight samples each, selected to cover the clinically important concentration range, is logistically manageable and statistically adequate to detect a 3% bias, assuming a CV of 3%. A plot of individual sample bias vs concentration will reveal any constant or proportional bias. Multiple laboratories using identical instrument/reagent systems can be standardized by choosing one site as a standardization laboratory for the group. This laboratory can confirm its accuracy by splitting fresh donor specimens with a CDC-certified laboratory as described above, or it

---

**Fig. 3.** Cholesterol bias (top) and HDL cholesterol bias (bottom) of frozen samples plotted as the difference between the mean percent bias determined with frozen human serum pools vs that determined with fresh, never-frozen donor sera.

The frozen pools were assayed in six runs during the time of the two rounds of eight fresh sera split with the CDC. All biases were calculated vs the CDC reference laboratory results. For cholesterol, the frozen serum pools were CAL1, 1830 mg/L; LPM2, 1840 mg/L; and CAL2, 2820 mg/L. The error bars indicate the standard error of the mean differences. For HDL cholesterol, frozen serum pool CAL1 (740 mg/L) was compared with fresh sera values >570 mg/L; and pools LPM2 (410 mg/L) and CAL2 (310 mg/L) were compared with fresh sera values >570 mg/L.

Human-serum-based lyophilized pools, intended for use as calibration materials, have substantial matrix effects with several cholesterol methods, and that fresh human sera must be used if one is to evaluate cholesterol accuracy reliably. The number of laboratories in our investigation is too small for us to generalize the observations of frozen pool bias between instrument/reagent systems or possible reagent-lot-specific effects. However, use of information obtained with frozen or lyophilized materials clearly is inappropriate to assess the accuracy of methods for cholesterol or HDL cholesterol.

The number of samples (N) required for each of two laboratories to detect an accuracy bias between them can be estimated by analyzing the data with a paired $t$-test as follows (15):

$$N = \frac{(2\sigma^2/(\Delta x)^2) \cdot (t_{N-1, 1-(a/2)} + t_{N-1, 1-\beta})^2}{(t_{N-1, 1-(a/2)} - t_{N-1, 1-\beta})^2}$$

---

148 CLINICAL CHEMISTRY, Vol. 36, No. 1, 1990
may qualify to become a CDC-certified laboratory. In principle, because all laboratories are using the same instrument/reagent system, any matrix bias in frozen or lyophilized materials should be the same for all participants. Thus, the standardization laboratory can assign target values to frozen or lyophilized secondary standard materials for use by participating laboratories to assess their calibration. The principal limitation in such an approach will be the variable response to frozen or lyophilized materials owing to different lots of reagents and to inter-instrument hardware differences. Reagent-lot-specific target values for the secondary standards can be determined by having each independent lot of reagent tested by the standardization laboratory. The standardization laboratory can assay fresh serum samples and the secondary standards using its calibrated lot of reagents and the independent lot of reagents. The bias in fresh serum results between the calibrated and independent reagents will provide a lot-specific correction factor to apply to the apparent values obtained for the secondary standards with the independent lot of reagents. This system would be less complicated and less costly if the laboratories could all use the same lot of reagents with as long a shelf life as possible.

We thank the clinical chemistry technologists at the Medical College of Virginia for their assistance with donor sera collection, and we thank the technologists in all of our laboratories for their assistance.

References