Performance of Four Homogeneous Direct Methods for LDL-Cholesterol

W. Greg Miller,1* Parvin P. Waymack,2 F. Philip Anderson,1 Steven F. Ethridge,2 and Eduviges C. Jayne1

Background: Homogeneous LDL-cholesterol methods from Genzyme, Reference Diagnostics, Roche, and Sigma were evaluated for precision, accuracy, and specificity for LDL in the presence of abnormal lipoproteins.

Methods: Each homogeneous method was performed by a Roche/Hitachi 911 according to the vendors’ instructions, and the results were compared with the β-quantification reference method. We measured precision over 20 days using quality-control and frozen serum specimens. Sera from 100 study participants, including 60 with hyperlipidemias, were assayed by each method. Accuracy was evaluated from regression and total error analysis. Specificity was evaluated from the bias (as a percentage) vs concentration of triglycerides.

Results: The total CV was <2% for all methods. Regression slope and intercept (with 95% confidence intervals) were as follows: Genzyme, 0.955 (0.92 to 0.99) and 30.3 (12 to 73) mg/L; Reference Diagnostics, 0.975 (0.93 to 1.02) and −8 (−63 to 47) mg/L; Roche, 1.067 (1.02 to 1.11) and −101 (−161 to −42) mg/L; and Sigma, 0.964 (0.91 to 1.02) and 164 (89 to 239) mg/L. The percentages of individual results with >12% bias were as follows: Genzyme, 8.0%; Reference Diagnostics, 11.0%; Roche, 10.0%; and Sigma, 30.0%. Total error calculated from mean systematic bias and all-sources random bias was as follows: Genzyme, 12.6%; Reference Diagnostics, 16.5%; Roche, 41.6%; and Sigma, 38.3%. Slopes of bias (as a percentage) vs triglycerides were P < 0.001 for all methods except the Roche method, which was P = 0.094.

Conclusions: The evaluated methods show nonspecificity toward abnormal lipoproteins, thus compromising their ability to satisfy the National Cholesterol Education Program goal for a total error of <12%. These homogeneous LDL-cholesterol results do not improve on the performance of LDL-cholesterol calculated by the Friedewald equation at triglyceride concentrations <4000 mg/L.

© 2002 American Association for Clinical Chemistry

LDL-cholesterol (LDL-C)3 is the primary lipoprotein risk factor recommended by the National Cholesterol Education Program (NCEP) to determine the need for and to assess the efficacy of medical intervention to lower cholesterol and risk for cardiovascular disease (1, 2). The LDL fraction includes lipoproteins in the density range of 1.019–1.063 kg/L, which, in addition to LDL, include intermediate-density lipoprotein [(IDL); density, 1.006–1.019 kg/L] and lipoprotein(a) [(Lp(a)); density, 1.045–1.080 kg/L]. LDL-C values for the epidemiologic basis for the NCEP clinical recommendations were validated by the β-quantification reference method (RM) (3). This RM includes cholesterol from IDL and Lp(a) in the measured LDL-C value. The NCEP Working Group on Lipoprotein Measurement (3) has recommended that methods used to measure LDL-C give results equivalent to those used to establish the epidemiologic database that was used to develop the recommendations for risk evaluation. Consequently, routine methods for measurement of LDL-C should determine the range of lipoproteins included in the LDL-C fraction measured by the β-quantification RM.

New methods that use homogeneous liquid assay techniques have been introduced to directly measure the LDL-C fraction by routine automated chemistry instruments. The potential advantages of directly measuring LDL-C include the ability to measure LDL-C when the triglycerides (TGs) are >4000 mg/L, the ability to mea-

1 Department of Pathology, Virginia Commonwealth University Health System, Richmond, VA 23298.
2 CDC, National Center for Environmental Health, Division of Laboratory Sciences, Special Activities Branch, Atlanta, GA 30341.
*Address correspondence to this author at: Virginia Commonwealth University, 401 N. 13th St., Richmond, VA 23219. Fax 804-828-0353; e-mail millerg@hsc.vcu.edu.

Received September 25, 2001; accepted December 11, 2001.
sure LDL-C in nonfasting individuals, the ability to measure LDL-C without the need to make the three measurements needed for the calculated result, and the potential reduction of imprecision by a single measurement instead calculating the value from three measured results.

The homogeneous direct methods use various physicochemical combinations of surfactants, polymeric complexes, and specific binding molecules to selectively measure cholesterol from the LDL fraction. A difficult analytic challenge for these methods is to accurately measure the LDL components in the presence of widely varying proportions of the different lipoprotein molecules encountered in clinical practice. Individuals with dyslipoproteinemia may have altered proportions of IDL and Lp(a), as well as chylomicrons, β-VLDL, and other intermediate species from lipid transport and metabolism (4, 5).

This study was designed to evaluate each of the four different homogeneous chemical methodologies for LDL-C that were commercially available in the United States. Each specimen was analyzed nearly simultaneously with a separate channel of a single automated random access analyzer to ensure equivalent conditions. This study is the first study to compare all the available homogeneous methods with the NCEP recommended β-quantification RM (nonmodified). The only other studies that attempted to establish a link to the nonmodified RM were performed by Sugiuichi et al. (6) and Nauck et al. (7), who evaluated only the Roche method. Other reports, such as those from Esteban-Salán et al. (8), Rifai et al. (9), and others discussed below, have evaluated one or two methods in the same study and have used an accuracy base of another field method or a modified ultracentrifugation procedure. Use of modified versions of the β-quantification RM or other ultracentrifugation methods is a confounding factor for comparison of performance of homogeneous methods between different reports.

Patient selection included 60% with several types of increased lipoproteins. The choice of study participants can influence the apparent performance of a method because nonspecificity for the LDL fraction is a critical challenge for the homogeneous methods. Testing the same panel of specimens on all methods allowed direct comparison for method accuracy and specificity performance. This report addresses accuracy, precision, and specificity for LDL in the presence of abnormal lipoproteins, performance vs the Friedewald calculated LDL-C, the effect of nonfasting specimens on LDL-C, and limitations in the total error (TE) calculation used by the NCEP Working Group on Lipoprotein Measurement (3).

Materials and Methods

Methods

Homogeneous LDL-C methods were performed with a Roche/Hitachi 911 (Roche Diagnostics) at Virginia Commonwealth University (Richmond, VA). Methods were performed on the analyzer according to the instructions of the respective reagent vendors with the calibrators provided by each vendor and the verification materials recommended by each vendor.

Genzyme Diagnostics (N-Geneous LDL Cholesterol Reagent) and Equal Diagnostics (LDL Direct Liquid Select Cholesterol Reagent) distribute the Daiichi reagent method. This method uses a detergent polymer mixture (reagent 1) to release cholesterol from non-LDL lipoproteins that reacts with cholesterol esterase, cholesterol oxidase, peroxidase, and 4-aminoantipyrine to produce colorless products. A second detergent reagent (reagent 2) that releases cholesterol from LDL is added. The cholesterol reacts with components of reagent 1 plus N,N-bis(4-sulfobutyl)-m-toluidine, disodium in reagent 2 to form a colored product that is measured spectrophotometrically. Three Genzyme reagent lot sets (cat. nos. 80-4598-00 and 80-4601-00 for reagents 1 and 2, respectively) were used with target values supplied by the NCEP. Calibration was performed with a single lot from the Genzyme LDL Cholesterol Calibrator Kit (cat. no. 80-4610-02). Calibration was performed once per reagent lot. A single lot of Genzyme LDL Cholesterol Control Set (two concentrations; cat. no. 80-4658-00) was used to verify method calibration for each reagent lot.

Reference Diagnostics (RDI Direct LDL) distributes the Denka Seiken reagent method. This method uses a surfactant to stabilize LDL molecules, so the cholesterol is nonreactive. Cholesterol from non-LDL lipoproteins reacts with cholesterol esterase, cholesterol oxidase, catalase, and N-(2-hydroxy-3-sulfopropyl)-3,5-dimethoxyaniline to produce colorless products. A second reagent is added that contains surfactants that release cholesterol from LDL and sodium azide to inhibit catalase. The cholesterol reacts with components of reagent 1 plus peroxidase and 4-aminoantipyrine in reagent 2 to form a colored product that is measured spectrophotometrically. One reagent lot set (cat. nos. 4300-1 and 4300-2 for reagents 1 and 2, respectively) was used during the evaluation. Calibration was performed with a single lot of Reference Diagnostics Direct LDL Cholesterol Calibrator (cat. no. LC-5). Calibration was performed once per reagent lot. Bio-Rad Liquichek Lipids Control (two concentrations; cat. nos. 641 and 642 for concentrations 1 and 2, respectively) were used with target values supplied by Reference Diagnostics to verify method calibration.

Roche Diagnostics (LDL-C plus) distributes the Kyowa reagent method. This method uses a cyclodextrin sulfate, dextran sulfate, magnesium complex to stabilize VLDL, and chylomicron molecules. Reagent 1 also contains peroxidase and sodium N-(2-hydroxy-3-sulfopropyl)-3,5-dimethoxyaniline. A second reagent is added that contains a detergent that forms stable micelles with HDL and the previously stabilized VLDL and chylomicrons and inhibits their reaction with cholesterol esterase and oxidase. Cholesterol from LDL reacts with components of reagent 1 plus cholesterol esterase, cholesterol oxidase, and 4-aminoantipyrine in the second reagent to produce a colored
product, which is measured spectrophotometrically. Three Roche reagent lot sets (cat. no. 1985612) were used during the evaluation. Calibration was performed with a single lot of Calibrator for Automated Systems HDL/LDL-C plus (cat. no. 1972235). Calibration was performed once per reagent lot. A single lot of Roche Precinorm L Special Lipid Control Serum and Precipath HDL/LDL-C (cat. nos. 781827 and 1778552, respectively) were used to verify method calibration for each reagent lot.

Sigma Diagnostics (EZ LDL Cholesterol) distributes the Wako reagent method. This method uses an enzymatic complex to stabilize LDL molecules so the cholesterol is nonreactive. Cholesterol from non-LDL lipoproteins reacts with cholesterol esterase, cholesterol oxidase, catalase, and N-(2-hydroxy-3-sulfopropyl)-3,5-dimethoxyaniline to produce colorless products. A second reagent is added that contains sodium azide to inhibit catalase and a detergent that releases cholesterol from LDL. The cholesterol reacts with components of reagent 1 plus peroxidase and 4-aminoantipyrine in reagent 2 to form a colored product that is measured spectrophotometrically. Two Sigma reagent lot sets (cat. no. 450061) were used during the evaluation. Calibration was performed with a single lot of EZ LDL Calibrator (cat. no. 358–3). Calibration was performed once per reagent lot. A single lot of Sigma Cardiolipid Control 1 and 2 (cat. nos. 4H020 and 4H021, respectively) were used to verify method calibration for each reagent lot.

Total cholesterol [(TC); cat. no. 450061] and TGs with glycerol blank correction (cat. no. 450032) were measured with enzymatic reagents from Roche by a Roche/Hitachi 911 analyzer according to the specifications of the manufacturer. Method performance was verified by participation in the CDC National Heart Lung and Blood Institute Lipid Standardization Program Part III. VLDL-cholesterol (VLDL-C) was calculated as the TC minus the bottom fraction cholesterol from the β-quantification procedure described below.

**REFERENCE METHOD**
The β-quantification RM for LDL-C and HDL-cholesterol (HDL-C) was performed at the CDC. The CDC β-quantification RM for HDL-C is an extension of the CDC HDL-C RM (10). VLDL-C was removed by ultracentrifugation of 5.00-mL serum samples for 18.5 h at 40 000 rpm in a Beckman type 50.3 rotor at 18 °C. The VLDL-C was removed with a tube slicer, and the remaining HDL-C and LDL-C were quantitatively transferred to a 5-mL volumetric flask, and the volume was increased to 5.00 mL with 0.15 mol/L sodium chloride. This bottom fraction cholesterol was analyzed in duplicate by the CDC Abell-Kendall cholesterol RM, and another aliquot was precipitated with heparin-manganese followed by duplicate Abell-Kendall analysis of the HDL-C in the supernatant. LDL-C was obtained by subtracting the HDL-C from the bottom fraction cholesterol.

**SPECIMEN COLLECTION**
Adult ambulatory patients (n = 100; 37 male, 63 female) were recruited to provide specimens that covered a wide range of cholesterol and TG concentrations. Each volunteer signed an informed consent document approved by the Institutional Review Board for Human Subjects Research of both organizations. The evaluation spanned 6 months, with most volunteers recruited during the first 3 months. Patients were instructed to report for phlebotomy after a 12-h fast. Four 15-mL draw SST vacuum tubes (cat. no. 366432; Becton Dickinson) of blood were collected from each participant. The blood was allowed to clot for 30–90 min and was then centrifuged at 2700g for 10 min at room temperature. Serum from each volunteer was transferred to a plastic vial and pooled. Aliquots were placed in glass tubes at 4 °C overnight to evaluate the presence of turbidity and chylomicrons. Aliquots (12 mL) in plastic tubes were shipped as liquid serum on cold packs by overnight express to the CDC. The 12 mL completely filled the tubes to avoid any agitation or frothing during transportation. The serum was assayed by field methods and RMs for TC, TGs, HDL-C, and LDL-C ~24 h (in some cases 48 h) after collection. After the initial blood collection, 25 of the volunteers consumed a breakfast containing a minimum fat content of 24 g and had their blood recollected 2–3 h after the meal. The blood was processed as described above, and the LDL-C assay was performed on the nonfasting serum specimens in the same analytical run as the fasting specimens.

**PRECISION**
Precision was evaluated according to NCCLS protocol (11) with duplicate assays in two analytical runs per day on 20 days with a new bottle of control material or new aliquot of frozen serum each day. During the time of data collection for precision estimation, one lot of reagents was used for the Reference Diagnostics and Sigma methods, and two lots of reagent were used for the Genzyme and Roche methods. Four materials were used to measure precision: Bio-Rad Liquichek Lipids Control Level 1 and 2; and aliquots of freshly collected off-the-clot serum stored at ~70 °C obtained from two donors at two LDL-C concentrations (Soloman Park Research Laboratories).

**ACCURACY**
Accuracy was evaluated by comparison of LDL-C results for single-donor, off-the-clot serum specimens between the routine methods and the β-quantification method performed at the CDC. All assays were performed in duplicate by routine methods and RMs, and the mean value was used for statistical analysis. The difference between routine method and RM results was calculated and expressed as the mean difference (bias). Regression analyses used the Deming calculation and assumed the same imprecision for each method (12). Three outlier values were excluded from the regression analysis as discussed in Results.
TOTAL ERROR

TE was evaluated by three procedures. TE (method 1) was calculated from the equation suggested by the NCEP TE = absolute value [mean bias (%)] + 1.96 (CVT), which uses the mean bias (as a percentage) from patient comparison results and the total CV from replication imprecision results (13). The CVT was the mean imprecision value of the four materials tested for each method. TE (method 2) was calculated nonparametrically as the percentage of routine method percentage differences from the RM value that exceeded the NCEP guideline of 12%. In this case, only the first observation of the duplicate measurements for each routine method was used for comparison with the RM value. TE (method 3) was calculated as TE = [SE + RE]/(RM mean value)] × 100, where SE is the absolute value of mean bias from individual results for the homogeneous methods vs the RM, and RE (random error) is the overall SD for the distribution of concentration differences for each individual homogeneous method result vs the RM (14). In this case, the random error includes all sources of imprecision, including replication and random specimen-specific effects, and assumes the combined sources of imprecision have a gaussian error distribution. The TE is expressed as a percentage of the RM mean LDL-C for all specimens.

INTERFERENCES (NONSPECIFICITY)

Interference from other lipoproteins, or nonspecificity for LDL-C, was evaluated by recruiting volunteers with increased concentrations of lipoproteins and comparing performance of the β-quantification RM to the routine methods. Concentration difference plots between the routine and reference LDL-C methods were constructed as a function of TG or HDL-C concentrations. The slopes of these plots were determined by regression analysis and evaluated for significance with a Student t-test. Three outlier values were excluded from the regression analysis as discussed below. Interferences from hemoglobin, bilirubin, and other substances have been reported in the product inserts of the manufacturers and were not evaluated further in this report.

Results

VOLUNTEERS

The 100 volunteers reported here included 22 with TC, HDL-C, LDL-C, and TGs within the NCEP Adult Treatment Panel III low risk cutoffs (1), and an additional 18 had at least one value within the NCEP borderline risk cutoffs. Sixty study participants had at least one lipid value in a high risk classification. Three volunteers were Fredrickson type III, 7 were type IIa, 11 were type IIb, 38 were type IV, and 1 had increased TC of 2530 mg/L, which was substantially attributable to 1020 mg/L HDL-C.

PRECISION

Precision of the methods is shown in Table 1. The total CV for all materials and all methods was ≤1.9% which is within the nominal 4% CV recommended by the NCEP (3, 13). LDL-C concentrations within the reference interval and those that were moderately increased had similar precision as did commercial control materials and frozen off-the-clot human sera. Other evaluations of homogeneous methods have also reported precision in the 1–3% CV range (6–9, 15).

ACCURACY

Three volunteers had highly discrepant results among the homogeneous methods. The specimens from each had results that were obvious outliers on visual inspection, suggesting the presence of sample-specific effects. These three study participants had Fredrickson type III lipid abnormalities as determined from increased TG (2770, 4190, and 5340 mg/L), TC (2460, 3370, and 4310 mg/L), and VLDL-C (1132, 1917, and 3204 mg/L) with a VLDL-C/TG ratio >0.3 and low LDL-C concentrations (832, 1109, and 752 mg/L), respectively. The percentage difference for LDL-C vs RM across methods ranged from −34.6 to 86.7% for Subject 1, −26.0 to 91.6% for Subject 2, and −13.6 to 278% for Subject 3. Another sample from the first volunteer was collected and assayed with similar results. The second and third volunteers were not available for retesting. Results for these three study participants were not included in regression analysis, but were included in the other evaluation parameters. Another report (15) evaluated the Roche and Wako (Sigma) meth-
methods and concluded they were not valid for dys-β-lipoproteinemic (type III) samples.

Figs. 1–4 present regression plots for each method vs the CDC RM. The points for the discrepant samples of the three volunteers described above are shown on the plots but are not included in the regression analysis. For the Roche and Sigma methods, the results of the type III patients are clearly outliers, whereas they appear to be more contiguous with other observations for the Genzyme and Reference Diagnostics methods. All data from these three volunteers were excluded from regression statistics to maintain uniform method-comparison parameters.

Table 2 presents summary statistics for the accuracy of the individual results of volunteers assayed by each method when results from all study participants were included. The mean biases for the Genzyme, Reference Diagnostics, and Roche methods were within the nominal 4% recommended by the NCEP. The Sigma mean bias was high, and the regression plot shows a large positive intercept consistent with a constant bias in this method. Because of the nonspecificity of the methods, statistics for accuracy were recalculated when only results for 40 volunteers with <2400 mg/L TC, <160 mg/L LDL-C, >40 mg/L HDL-C, and <1500 mg/L TG were included. In this subset, the slopes were 0.996, 1.017, 1.038, and 1.021 with intercepts of −42, −70, −86, and 24 mg/L, and mean biases were −4.0, −4.4, −3.9, and 4.4% for Genzyme, Reference Diagnostics, Roche, and Sigma, respectively. On the basis of normolipidemic results, each of the four methods had acceptable calibration traceability to the RM.
Difference plots for the sensitivity of the homogeneous LDL-C methods to various compositions of lipoproteins reflected by their TG concentrations are presented in Figs. 5–8. In addition, the difference when the LDL-C was calculated from the Friedewald equation is shown. The plots are qualitatively similar, showing that the LDL-C results become more positive as TG increases. The Genzyme and Sigma methods had slopes of 10.3 and 14.3, respectively, which were different from zero ($P < 0.001$) and had a pronounced systematic dependency on TG concentration. The Reference Diagnostics method had a slope of 4.5, which was different from zero ($P < 0.001$), but showed a smaller effect over the TG range and was usable to concentrations of at least 6000 mg/L. The Roche method had a nonsignificant slope of 1.7 ($P = 0.094$) but had substantial variability at higher TG values, making the mean slope a poor indicator of specificity performance. The Roche and Sigma methods both had substantial nonspecificity for three type III dyslipidemic specimens; biases were 86–278% for Roche and 72–162% for Sigma.

When HDL-C was used as a marker for abnormal lipoproteins, only the Roche method had a nonsignificant slope; all other methods had significant ($P < 0.001$) sensitivity to changes in HDL-C concentrations in the specimens. Both the Genzyme and Sigma methods had a noticeable increase in bias at lower HDL-C concentrations, which is consistent with their more pronounced dependency on TG concentration and the inverse correlation between high TG and low HDL-C concentrations.

The slopes of plots of LDL-C difference from RM vs TG values for the subset of normolipidemic subjects ranged from 4.6 to 5.9 with $P < 0.001$ for the Genzyme, Reference Diagnostics, and Sigma methods. These results show a pronounced sensitivity to lipoprotein composition despite the overall acceptable calibration accuracy statistics in this limited range of concentrations. The Roche method for this subset of normolipidemic volunteers had a nonsignificant slope of 0.3 ($P = 0.74$).

### Table 2. Accuracy parameters for patient specimens vs CDC $\beta$-quantification RM for four homogeneous LDL-C methods.

<table>
<thead>
<tr>
<th>Accuracy parameter</th>
<th>Genzyme N-Geneous LDL</th>
<th>Reference Diagnostics LDL</th>
<th>Roche LDL-C plus</th>
<th>Sigma EZ-LDL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deming slope ($95% \text{ CI}$)</td>
<td>0.955 (0.92 to 0.99)</td>
<td>0.975 (0.93 to 1.02)</td>
<td>1.067 (1.02 to 1.11)</td>
<td>0.964 (0.91 to 1.02)</td>
</tr>
<tr>
<td>$y$ intercept, mg/L ($95% \text{ CI}$)</td>
<td>30.3 (−12 to 73)</td>
<td>−7 (−63 to 47)</td>
<td>−101 (−161 to −42)</td>
<td>164 (89 to 239)</td>
</tr>
<tr>
<td>$S_y$, mg/L</td>
<td>58.4</td>
<td>74.9</td>
<td>81.3</td>
<td>100</td>
</tr>
<tr>
<td>Mean bias, % ($95% \text{ CI}$)</td>
<td>−1.4 (−2.7 to 0)</td>
<td>−3.5 (−5.0 to −2.1)</td>
<td>2.8 (−3.4 to 8.9)</td>
<td>14 (9.6 to 18)</td>
</tr>
<tr>
<td>TE as percentage of individual biases $&gt;12%$</td>
<td>8.0</td>
<td>11.0</td>
<td>10.0</td>
<td>30.0</td>
</tr>
<tr>
<td>TE from mean bias and imprecision CV, %</td>
<td>3.8</td>
<td>5.9</td>
<td>5.8</td>
<td>15.5</td>
</tr>
<tr>
<td>TE as SE + RE for 95% range of differences vs RM, %</td>
<td>12.6</td>
<td>16.5</td>
<td>41.6</td>
<td>38.3</td>
</tr>
</tbody>
</table>

* Regression statistics were calculated for 97 subjects after exclusion of three type III individuals as explained in the text. All other statistics were calculated from results for 100 individuals with no exclusions.
* CI, confidence interval.
* The first observation for each volunteer was used for this parameter.
* Calculated as $\text{TE} = \text{absolute value} \left[ \text{mean bias (β)} + 1.96 (\text{CV}) \right]$.
* From Kringle reference (13).
Others. For Genzyme, mean biases of the analytes reported here is consistent with that observed by the Friedewald method. The mean accuracy traceability for a subset of 40 normolipidemic clinical specimens, which are selected for the purpose of validating the calibration traceability of a field method to the RM. We found that all four methods had a TE of approximately 1% for Roche LDL plus and Friedewald-calculated LDL-C vs TG concentration. The mean accuracy traceability for a subset of 40 normolipidemic clinical specimens, which are selected for the purpose of validating the calibration traceability of a field method to the RM. We found that all four methods had a TE of approximately 1% for Roche LDL plus and Friedewald-calculated LDL-C vs TG concentration.

Table 3 shows the ability of the homogeneous LDL-C methods vs the RM to classify individual volunteers into treatment groups according to NCEP cutoffs (1). For the Friedewald method, only individuals with TG concentrations <4000 mg/L were included.

Nonfasting Specimens
The effect of postprandial lipoproteins on the four homogeneous assays and the RM is shown in Fig. 9. A positive t-test was found for the nonfasting vs fasting values for each method. The postprandial changes were similar for the RM and the homogeneous methods. There was postprandial variability among patients, with 16% having nonfasting values that were >5% different from their fasting values. Some individuals had large enough changes that nonfasting specimens could have produced erroneous conclusions regarding LDL-C status.

Discussion

Accuracy Traceability
Traditional figures of merit for accuracy such as mean bias or slope and intercept of a regression plot of results vs a RM are determined with data averaging statistics and are most useful to evaluate how successfully a method has been calibrated by the manufacturer to agree with the RM. Calculating the means of the data can mask the impact of measurement nonspecificity. The mean accuracy reported here is consistent with that observed by others. For Genzyme, mean biases of -2.2 and -5.9% and slopes of 0.91 and 0.99 with intercepts of 71 and -46 mg/L have been reported (9, 15). For Roche, mean biases of -11.0% to 0.7% and slopes of 1.01 to 1.11 with intercepts of 0 to -186 mg/L have been reported (6–8, 15). For Sigma, a mean bias of approximately -1.8% and a slope of 0.91 with an intercept of 104 mg/L have been reported (8).

The NCEP Working Group on laboratory measurement has recommended that the TE for an individual LDL-C measurement should not exceed 12% from the result by the RM (3, 13). This TE goal specifies acceptable analytical performance for a single measurement that would produce a clinically acceptable result related to the within-individual biological variability of LDL-C (3). The NCEP calculation of TE was established as the mean bias + (1.96 × CVT), where CVT was from replicate imprecision measurements and accounted for the sources of error pertinent to validating the mean calibration status of a method. We found the TE for three of four methods was acceptable, using the NCEP equation, because the mean systematic biases were small and the replication imprecision was very small. The Sigma method exceeded the 12% specification primarily because of a large systematic bias. Other reports have used the NCEP equation for TE and reported values of 6.8–11.4% for Genzyme (9, 15), 4.1–15.0% for Roche (7, 8, 15), and 2.6–5.6% for Sigma (8). One report (15) for the Roche method had TEs of 15.1% and 12.5% at 1300 and 1600 mg/L LDL-C respectively. The results in most cases support acceptable method calibration by the manufacturers but suggest some improvement in consistency may be desirable.

When method nonspecificity is an issue, the NCEP TE equation is best applied to a subset of normolipidemic clinical specimens, which are selected for the purpose of validating the calibration traceability of a field method to the RM. We found that all four methods had a TE of 6–7% for a subset of 40 normolipidemic specimens, with LDL-C values ranging from 430 to 1600 mg/L, which was consistent with acceptable calibration by the product calibrators of the manufacturers.

The Cholesterol Reference Method Laboratory Network certification protocol for manufacturers specifies distributions for LDL-C and TG values in the fresh split-sample comparison of the field method with the RM to establish accuracy traceability for their product calibrators. It is clear from Figs. 5–8 that the distribution of TG values in the specimens used for accuracy transfer can affect the concentration at which minimum bias is observed. Because of method nonspecificity for LDL fraction
molecules, the apparent accuracy of a method can vary with the lipoprotein composition of the specimens used for accuracy transfer and value assignment vs the RM. Another factor to be considered when evaluating the bias vs the RM is a contribution from RM performance with dyslipoproteinemic samples. The NCEP definition of LDL-C includes LDL, IDL, and Lp(a), which are contained in the bottom fraction of the β-quantification method. The clear zone at the cutting point of the ultracentrifugation tube becomes problematic in dyslipoproteinemic samples when populations of lipoproteins with atypical density present a continuum of densities at this critical point.

**TOTAL ERROR**

A limitation of mean bias, regression parameters, and TE with the NCEP equation is that these attributes represent statistical averages of performance and do not reflect the magnitude of individual result biases nor the number of individual results that have clinically unacceptable bias. To be clinically useful, the TE must include calibration bias, measurement imprecision, and other sources of imprecision, especially random specimen effects attributable to method nonspecificity. Examination of Figs. 5–8 shows a large number of individual results that exceed the 12% TE specification. Because this source of imprecision is not included in the NCEP TE equation, the NCEP TE underestimated the actual TE and overstated the clinical suitability of the methods.

The clinical suitability of a method is best described when all sources of bias and imprecision are included in the estimate of TE. We examined alternate methods to express TE that included all sources of imprecision including specimen nonspecificity. A nonparametric approach determined the percentage of individual results that exceeded the clinically important threshold of 12% difference from the RM. This simple nonparametric calculation includes all components of error, including those that may not be recognized and may not conform to a gaussian error distribution. Because 95% of measurements made on individual clinical specimens should fall within the TE specification, any value >5% represents nonconformance to the TE goal. Another nonparametric procedure has been suggested by Krouwer (16) to estimate TE range as the central 95% of ranked concentration differences between the field and RMs. This approach gave observed difference ranges of −13.3 to 16.2%, −25.2 to 12.2%, −13.5 to 88.8%, and 0.4 to 89.2% for Genzyme, Reference Diagnostics, Roche, and Sigma, respectively.

The parametric TE estimation procedure described by Kringle et al. (14) accounts for all sources of imprecision including random measurement and specimen nonspecificity effects. This procedure provides a numeric estimate of the magnitude of the TE as an aggregate percentage very similar to that used by the NCEP calculation.

None of the LDL-C homogeneous methods had acceptable clinical performance according to TE with nonparametric or parametric procedures, which included all sources of variability. The current NCEP approach to estimating TE can produce erroneous conclusions regarding clinical performance of a method primarily attributable to failure to account for random imprecision from nonspecificity in individual specimens.

**NONSPECIFICITY**

The TE results suggest that the homogeneous methods suffer nonspecificity for the LDL-C fraction. Nonspecificity was also evaluated from the slope of plots of the concentration difference between the routine and RM results vs TG or HDL-C concentrations, which provide a surrogate measure of the presence of altered lipoprotein proportions or the presence of abnormal lipoprotein molecules. The current generation of homogeneous LDL-C methods does not perform as well as desired over the range of lipoprotein composition encountered in clinical practice.

<table>
<thead>
<tr>
<th>Classification group (no. in group by RM)</th>
<th>Genzyme N-Geneous LDL</th>
<th>Reference Diagnostics LDL</th>
<th>Roche LDL-C plus</th>
<th>Sigma EZ-LDL</th>
<th>Friedewald LDL-C*</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1300 mg/L (59)</td>
<td>98</td>
<td>100</td>
<td>92</td>
<td>76</td>
<td>96</td>
</tr>
<tr>
<td>1300–1600 mg/L (24)</td>
<td>69</td>
<td>67</td>
<td>75</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>&gt;1600 mg/L (17)</td>
<td>91</td>
<td>91</td>
<td>94</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

*a* Calculated for specimens with TG concentrations <4000 mg/L.

The limits of the upper and lower bars represent the maximum and minimum values, the upper and lower box lines represent the 75th and 25th percentile values, and the line in the box represents the median value.
We can speculate that mechanisms involving differential recoveries of altered proportions of normolipidemic proteins and the presence of abnormal lipoproteins may account for the incorrect results. Gómez et al. (17) reported positive bias of the Denka Seiken (Reference Diagnostics) method in patients with liver cirrhosis in which analysis of apolipoprotein B-rich nascent VLDL as LDL was attributed to lecithin:cholesterol acyl transferase deficiency. Fei et al. (18) reported that the Roche method recovered 52% and the Genzyme method recovered 31% of purified IDL-cholesterol (IDL-C) added to a serum sample, whereas 0% and 51%, respectively, of added lipoprotein-X was recovered by these two methods. In addition, the methods reacted with 18% and 8%, respectively, of apolipoprotein E-rich HDL-C from cholestatic patients. Ordoñez-Llanos et al. (19) reported that the Roche method was reactive with 7% of VLDL-C, 64% of IDL-C, 75% of LDL-C, 37% of Lp(a), and 0% of HDL-C in various serum pools containing these fractions. Sugiuchi et al. (6) reported 78% recovery of IDL-C in an early version of the Roche method. Sakaue et al. (20) reported recovery of purified fractions of VLDL-C, IDL-C, and LDL-C, respectively, of 19%, 47%, and 87% for Genzyme; 10%, 31%, and 95% for Reference Diagnostics; and 16%, 64%, and 86% for Roche. Ragland et al. (21) found no interference with the Genzyme method in the presence of abnormal lipoproteins in diabetic patients.

**Comparison to Friedewald-Calculated LDL-C**

Historically, the Friedewald equation \[ \text{LDL-C} = \text{TC} - \text{HDL-C} - \left( \frac{\text{TG}}{5} \right) \] has been used to calculate LDL-C from measured values of TC, HDL-C, and TG (22). For well standardized measurements, the calculated LDL-C agrees with the \( \beta \)-quantification RM to TG concentrations of 2000 mg/L and gives clinically useful values (errors, 7–10%) up to TG concentrations of 4000 mg/L (3, 23). The calculation is not recommended for specimens with >4000 mg/L TG, nor in patients with chylomicrons or type III hyperlipidemia, because of unreliable estimation of VLDL-C by the TG/5 factor.

One of the principle reasons to perform a direct homogeneous assay of LDL-C is to overcome the limitations of a Friedewald-calculated value. Examination of Figs. 5–8 suggest that these homogeneous methods are no better than the Friedewald-calculated value at <3000 to 4000 mg/L TG because the agreement and imprecision vs the RM are similar. The Genzyme and Reference Diagnostics methods had less variability in LDL-C results than did the Friedewald-calculated values at 3000–6000 mg/L TG. Above 6000 mg/L TG, there were too few observations for reliable conclusions. The Roche method also gave acceptable values between 3000 to 6000 mg/L TG, but had several obvious outlier values. The Sigma method had the greatest sensitivity to TG and did not produce acceptable results at TG concentrations >2000 mg/L.

The homogeneous methods offered no improvement compared with Friedewald-calculated LDL-C for classifying individuals into treatment groups on the basis of cutoffs recommended by the NCEP. Other investigators have also reported that homogeneous LDL-C methods do not offer better performance than Friedewald-calculated values. Nauck and Rafai (15) evaluated the Genzyme and Roche methods vs a modified RM and found no better classification performance (71–87% correct) than Friedewald (75–84% correct) for LDL-C and that Roche was actually poorer (57% correct) at TG concentrations <2000 mg/L. They did report that classification by the homogeneous methods at 4000–6000 mg/L TG (79–84% correct) was approximately the same as at lower TG values. Similarly, Esteban-Salán et al. (8) reported that classification by the Roche and Sigma methods vs a modified RM at TG concentrations <2000 mg/L agreed with the RM in 98–99% of cases vs 95% for Friedewald, at TG concentrations of 2000–4000 mg/L in 90–92% of cases vs 76–89% for Friedewald, and at TG concentrations of 4000–10 000 mg/L in 74–86% of cases, which was similar to Friedewald at the lower TG values. Rifai et al. (9) reported that the Genzyme method vs a modified RM had poorer sensitivity (77%) than Friedewald (91%) for LDL-C values ≥1300 mg/L, whereas the methods had approximately the same specificity (92% and 90%, respectively).

The data suggest the homogeneous methods perform no better than the Friedewald calculation at low TG concentrations, but can produce clinically useful results at higher TG concentrations where the Friedewald calculation is not valid.

In summary, all four homogeneous LDL-C methods had assay precision within the NCEP performance guideline. The mean bias and regression statistics vs the \( \beta \)-quantification RM for the Genzyme, Reference Diagnostics, and Roche methods supported acceptable method calibration by the manufacturers. The Sigma method bias was not acceptable. Both the Roche and Sigma methods had intercepts consistent with constant bias and results for several specimens that had large differences from the RM. All four methods had excessive TE for individual specimens. The principal limitation of these homogeneous methods for LDL-C was nonspecificity for the LDL fraction over the range of lipoprotein composition encountered in clinical practice. The homogeneous methods offered no advantage over a Friedewald-calculated value at <4000 mg/L TG. The practical utility of these homogeneous methods was the ability to perform LDL-C measurements for specimens with moderately increased TG for which the Friedewald calculation was not acceptable. However, all the methods had progressively poorer performance as the TG concentration of the specimens increased. Measured LDL-C in nonfasting specimens is not recommended because of variable interferences among individuals.

We are grateful to each of the vendors who supplied the reagents used in this evaluation. We also thank Jean
Coaker, Helen Cosby, Francine Jones, Dr. Lampros Karganoros, Dr. James McKenney, Juanita Mojica, Sylvia Phillips, Dr. Jamie Plata, Victoria Smith, Mary Jo Taylor, Linda Tiller, and Criptal Tucker who assisted with recruitment of volunteers and collection of specimens.

References


