Clinical impact of direct HDLc and LDLc method bias in hypertriglyceridemia. A simulation study of the EAS-EFLM Collaborative Project Group

Michel R. Langlois, Olivier S. Descamps, Arnoud van der Laarse, Cas Weykamp, Hannsjörg Baum, Kari Pulkkki, Arnold von Eckardstein, Dirk De Bacquer, Jan Borén, Olov Wiklund, Paivi Laitinen, Wytze P. Oosterhuis, Christa Cobbaert, for the EAS-EFLM Collaborative Project

European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) Working Group (WG) Cardiac Markers
WG Guidelines, EFLM
Department of Laboratory Medicine, AZ St-Jan Bruges and Ghent University, B-8000 Bruges, Belgium
Centre de Recherche Médicale de Jolimont, Hôpital de Jolimont, 7100 Haine Saint-Paul, Belgium
Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands
Queen Beatris Hospital, Winterswijk, The Netherlands
Institut für Laboratoriumsmedizin, Mikrobiologie, Blutdepot und Krankenhaushygiene, Regionale Kliniken Holding RKH GmbH, Posilipostr. 4, D-71640 Ludwigsburg, Germany
Department of Clinical Chemistry, University of Eastern Finland and Eastern Finland Laboratory Centre, P.O.B. 1700 FIN-70211 Kuopio, Finland
Institute of Clinical Chemistry, University Hospital of Zurich, Rämistrasse 100, 8091 Zürich, Switzerland
Department of Public Health, Ghent University, B-9000 Ghent, Belgium
Wallenberg Laboratory, Sahlgrenska University Hospital, SE-413 45 Gothenburg, Sweden
Department of Clinical Chemistry and Laboratory Medicine, Leiden University Medical Center, Leiden, The Netherlands

Article history:
Received 8 August 2013
Received in revised form 18 November 2013
Accepted 2 December 2013
Available online 4 January 2014

Keywords:
HDL-cholesterol
LDL-cholesterol
Hypertriglyceridemia
SCORE
Risk classification
Treatment goals

Background: Despite international standardization programs for LDLc and HDLc measurements, results vary significantly with methods from different manufacturers. We aimed to simulate the impact of analytical error and hypertriglyceridemia on HDLc- and LDLc-based cardiovascular risk classification.

Methods: From the Dutch National EQA-2012 external quality assessment of 200 clinical laboratories, we examined data from normotriglyceridemic (≤1 mmol/l) and hypertriglyceridemic (≥7 mmol/l) serum pools with lipid target values assigned by the Lipid Reference Laboratory in Rotterdam. HDLc and LDLc were measured using direct methods of Abbott, Beckman, Siemens, Roche, Olympus, or Ortho Clinical Diagnostics. We simulated risk reclassification using HDL- and sex-specific SCORE multipliers considering two fictitious moderate-risk patients with initial SCORE 4% (man) and 3% (woman). Classification into high-risk treatment groups (LDLc >2.50 mmol/l) was compared between calculated LDLc and direct LDLc methods.

Results: Overall HDLc measurements in hypertriglyceridemic serum showed negative mean bias of −15%. HDL-multipliers falsely reclassified 70% of women and 43% of men to a high-risk (SCORE ≥5%) in hypertriglyceridemic serum (P < 0.0001 vs. normotriglyceridemic serum) with method-dependent risk reclassifications. Direct LDLc in hypertriglyceridemic serum showed positive mean bias with Abbott (+16%) and Beckman (+14%) and negative mean bias with Roche (−7%). In hypertriglyceridemic serum, 57% of direct LDLc measurements were above high-risk treatment goal (2.50 mmol/l) vs. 29% of direct LDLc (33% of calculated LDLc) in normotriglyceridemic sera.

Conclusion: LDLc and HDLc measurements are unreliable in severe hypertriglyceridemia, and should be applied with caution in SCORE risk classification and therapeutic strategies.
1. Introduction

The EAS/ESC guidelines recommend that asymptomatic individuals at high cardiovascular disease (CVD) mortality risk should be identified for statin therapy [1]. For this purpose, risk assessment is performed using the SCORE (Systematic Coronary Risk Evaluation) prediction model estimating 10-year risk of CVD mortality, based on gender, age, total cholesterol, systolic blood pressure and smoking status. Recently, the 2011 ESC/EAS guidelines on the management of dyslipidemias have considered the additional impact of high-density lipoprotein cholesterol (HDLc) on CVD risk [2]. The effects of differing HDLc levels may also be calculated from the classical SCORE using HDL- and sex-specific multipliers according to Descamps et al. [2].

In patients with dyslipidemia, prevention strategies with either lifestyle changes or lipid-lowering agents are primarily targeted by low-density lipoprotein cholesterol (LDLc). The higher the predicted risk, the lower is the recommended LDLc goal and hence the need to initiate statin therapy. The recommended LDLc therapeutic goal is <2.59 mmol/l in high risk individuals (SCORE 5–9%) and <1.80 mmol/l or a 50% reduction in LDLc in very high risk individuals (SCORE >10%) [1].

There is a direct relationship between serum LDLc and incidence of CVD. Similarly, there is a strong inverse association between HDLc and CVD, although recent Mendelian randomization studies found no causal relationship between genetically decreased or increased HDLc and the risk of myocardial infarction [3,4]. However, our concern about including HDLc and LDLc in risk estimation models relates to the potential for analytical error due to imprecision and bias of the lipid measurements. Despite the widespread belief that the calculation or measurement of LDLc or HDLc is standardized and reproducible, results can vary significantly with methods from different manufacturers. In the previous century, the earliest measurements involved ultracentrifugation and precipitation for isolation of LDL and HDL [5]. In the late 1990s, “homogeneous” or “direct” LDLc and HDLc methods have been introduced in the clinical laboratories and largely replaced the older assays [6–8]. Direct LDLc and HDLc methods are commercially available as ready-to-use reagents, enabling full automation of the measurements, however their bias (deviation from “true” value) is a major point of concern. Discrepant results have been reported among the various direct methods, particularly in hypertriglyceridemic and dyslipidemic samples [9–13]. This is also evident from large-scale accuracy-based quality surveys organized across different laboratories [14]. Problems with direct HDLc assays also raise concerns about the reliability of calculated LDLc and non-HDLc treatment goals [12]. Poor reliability of these methods relate to the heterogeneity of both LDL and HDL particles [11,12].

In this study, we aimed to illustrate the potential impact of analytical errors in current LDLc and HDLc measurements on making clinical decisions. A simulation is used here to explore potential CVD risk misclassifications as defined by the SCORE model. Misclassification may occur if a true lipid concentration is within a desirable range, but the reported lipid value is in a high-risk range, or if a true lipid concentration is in a high-risk range but the reported lipid value is in a desirable range [15]. These misclassifications reflect a clinically relevant issue because they indicate that the practical difficulty situation with treatment options: to avoid unnecessary treatment of a patient whose lipid concentration is in a desirable risk category, or to treat a patient whose lipid concentration is in a high-risk category, and to distinguish between ‘moderate’ and ‘high-risk’ categories when lipid values are near a cutpoint [15]. Misclassification as defined here is of greatest concern because of its potential impact on the patient and healthcare economics. Using data of the Dutch National EQA-2012 external quality assessment of clinical laboratories, representing all LDLc and HDLc reagent systems used in The Netherlands, we simulated the effects of analytical error and hypertriglyceridemia on HDL-adjusted SCORES and concordance of treatment goals.

2. Materials and methods

2.1. Samples

The Dutch external quality assessment (EQA) organizer, the Stichting Kwaliteitsbewaking Medische Laboratoriumundiagnostiek (SKML), runs an accuracy-based EQA scheme for clinical chemistry analytes including lipids and apolipoproteins. Quality of the Dutch EQA program has been described previously [16,17]. Briefly, serum pools are prepared in an ISO 13485:2003 certified production facility according to CLSI C21-A protocol [18] and value-assigned for total cholesterol, LDLc and HDLc with CDC Reference Methods in the Lipid Reference Laboratory in Rotterdam, an international member of the CDC Cholesterol Reference Method Laboratory Network (CRMLN). Serum pools are assigned for apolipoprotein B (apoB) at Leiden University Medical Center (LUMC) and are traceable to the WHO-IFCC reference material SP5-08 [19]. The regular EQA-scheme encompasses the analysis of 24 fresh frozen commutable samples per year, i.e., one sample has to be analyzed per two week intervals. About 200 Dutch clinical chemistry laboratories participate in the EQA survey for serum lipids. EQA-test results are electronically submitted through Qbase to the SKML.

In this paper we consider national EQA-results from three normotriglyceridemic (NTG) serum pools (2012.1F, 2012.2D and 2012.2F) and two hypertriglyceridemic (HTG) serum pools (2012.1D and 2012.2E), prepared to study the effect of hypertriglyceridemia. For this purpose, original HTG pools from a selected single donor with triglycerides (TG) ~11 mmol/l were mixed with a NTG serum pool in order to end up with a total TG of ~7 mmol/l. HTG serum pool 2012.1D is a fresh frozen pool stored at ~84 °C, whereas 2012.2E is a fresh pool. After aliquoting, NTG and HTG (2012.1D) frozen serum pools were transported to the participating laboratories on dry ice and kept frozen at ~84 °C until analysis. Lipid and apolipoprotein measurements, including the target value assignments, of the fresh HTG aliquots (2012.2E) were performed within 2 days upon storage at 4 °C. The participating labs were instructed by SKML to store and process the specimen correctly.

2.2. Biochemical measurements

Direct HDLc measurements were performed in 2012 with state-of-the-art homogeneous methods using Accelerator Selective Detergent (Abbott Diagnostics Division, Beckman Coulter, Siemens Healthcare Diagnostics), PEG-modified enzymatic reagent (Roche Diagnostics), immunoinhibition (Olympus), or Vitros reflectometry slide technology (Ortho Clinical Diagnostics) on automated instruments from the same manufacturers. In 123 laboratories, LDLc was calculated with the Friedewald equation LDLc = cholesterol – HDLc – TG/2.22 (in mmol/l) [20], using direct HDLc from each manufacturer and cholesterol and TG measurements from the same manufacturer in the calculation. In other laboratories (n = 49), direct LDLc measurements were performed using α-cyclodextrin sulfate–dextran sulfate Mg2+ (Roche) or selective detergent methods (Abbott, Beckman, Ortho, Siemens) on the same instruments as for HDLc. In some but not all laboratories (n = 38), apoB was measured with immunonephelometry (Beckman, Siemens) or immunoturbidimetry (Abbott, Roche) instruments. Non-HDLc was not reported by the laboratories but calculated in this study by subtracting HDLc from total cholesterol values in the EQA
database. Fig. 1 shows the collected data for HDLc in HTG serum pool 2012.1D and LDLc, non-HDLc, and apoB in HTG pool 2012.2E.

2.3. Simulation of SCORES and treatment goals

From the EQA database, we selected those samples with mean concentrations close to the high-risk cutpoints for HDLc (men 1.0 mmol/l, women 1.2 mmol/l) and LDLc (2.5 mmol/l) in order to evaluate risk misclassifications due to measurement variability (Table 1). To study the effect of hypertriglyceridemia, we compared serum pools 2012.1D vs. 2012.1F for HDLc, 2012.2D vs. 2012.2E for LDLc, and 2012.2E vs. 2012.2F for apoB (cutpoint 1.00 g/l). To examine the concordance of SCORE classification by the various HDLc measurements, we simulated HDL-adjusted risk SCORE calculations by applying HDL multipliers according to Descamps et al. [2] considering a fictitious male and female patient with an initial SCORE of 4% and 3%, respectively (moderate risk). To examine the concordance of classification into treatment groups, the relative numbers of LDLc concentrations above the “high-risk” goal (2.5 mmol/l) were compared among the different direct LDLc methods and the calculated LDLc values, using the direct HDLc method from each manufacturer and cholesterol and TG from the same manufacturer in the calculation.

2.4. Statistics

Laboratory results are summarized by groups that use the same reagent. Method groups with small data sets (<5 were not separately studied but the data are included in the overall results of the HDLc and LDLc surveys. Lipid and apolipoprotein data are presented as medians and ranges. Inter-laboratory imprecision was assessed by calculating the coefficient of variation (CV) of reported values both overall and per manufacturer/method group. In our EQA scheme we gathered test results from single measurements under routine circumstances; hence, we do not have intra-laboratory imprecision data of the individual measurements. The mean bias compared to the “true” target value of measurement was calculated as reported concentration minus the target concentration, and bias differences in HTG vs. NTG samples were evaluated by the Wilcoxon test. Differences in SCORE classification and treatment goal classification of the fictitious patients were evaluated with the χ² test. Statistical significance was considered at the level P < 0.05. MedCalc software was used for the statistical analysis.

3. Results

3.1. HDL-adjusted SCORES

Median HDLc concentrations measured by the various methods did not show any major differences in the NTG serum (Table 2A). However, inter-laboratory imprecision caused a variability of HDLc concentrations reported among the participating laboratories. The application of HDL-multipliers to a fictitious woman (SCORE 3%) or man (SCORE 4%) yielded high-risk categories (SCORE >5%) in a minority of all HDLc measurements (3% and 1%, respectively). According to the HDLc target value measured in the reference laboratory, it is expected that 100% of these men and women would maintain a moderate risk level below 5%.

The overall HDLc concentration measured in the HTG serum (mean 0.92 mmol/l, SD 0.11 mmol/l) showed a negative bias of −0.16 mmol/l (−14.8%) compared to the target value 1.08 mmol/l (Table 2B). This resulted in higher HDL-adjusted SCOREs compared to the NTG sample (P < 0.0001). Striking inter-method differences in mean HDLc bias were observed in the HTG sample, −0.03 mmol/l (−2.7%) (Abbott), −0.07 mmol/l (−6.5%) (Beckman), −0.10 mmol/l (−9.3%) (Olympus), −0.21 mmol/l (−19.4%) (Roche), and −0.24 mmol/l (−22.2%) (Siemens). As a consequence, HDL-adjusted SCOREs yielded considerable differences in numbers of high-risk categories depending on the method (Table 2B).

3.2. Treatment goals

Calculated LDLc obtained with the various direct HDLc data (and cholesterol and TG measured on the same instruments) showed a
wide range due to inter-laboratory imprecision within each method group (Table 3A). With this median value of 2.40 mmol/l, 33% of overall calculated values were above the high-risk goal of therapeutic intervention (2.50 mmol/l). With direct LDLc measurements in the NTG sample, 29% of reported values were above high-risk goal (Table 3B). In the HTG sample, direct LDLc measurements were overall 57% above high-risk goal (Table 3C). Abbott and Beckman method means showed considerably positive biases of +0.39 mmol/l (+15.9%) and +0.34 mmol/l (+13.8%), respectively, compared to the target value 2.46 mmol/l in the HTG sample, resulting in 100% of values above treatment goal; Roche direct LDLc showed a negative mean bias of −0.16 mmol/l (−6.5%) and only 39% above treatment goal in HTG sera. It should be noted that these biases were observed in a fresh HTG serum pool 2012.2E, prepared to exclude freeze-thawing effect on the measurements.

### 3.3. Secondary treatment goals

Overall mean bias of non-HDLc calculations in the HTG serum pool 2012.2E (n = 194) was +7.1% (95% CI 5.4–8.4%) compared to the expected value based on total cholesterol and HDLc reference measurement procedures. Largest biases were observed with Roche (mean + 9.1%, 95% CI 7.9–10.3%) and Siemens (mean + 10.8%, 95% CI 6.4–12.3%). Overall inter-laboratory imprecision of non-HDLc calculations in the 2012.2E survey was 8.5%; within-method inter-laboratory CVs were 7.1% (Abbott), 7.6% (Beckman), 6.6% (Olympus), 7.3% (Roche), and 9.7% (Siemens). We were unable to simulate misclassifications in non-HDLc-based treatment groups, because the range of calculated non-HDLc values in HTG sera from the EQA survey did not encompass the high-risk goal (3.3 mmol/l).

Overall apoB measurements showed comparable imprecision (8.8%) but less bias (+1.0%) than non-HDLc calculations in the same HTG serum pool 2012.2E (Table 4). Overall biases of apoB measurements did not differ between NTG and HTG samples (P = 0.382). In hypertriglyceridemia, misclassification of high risk by apoB measurement (29%) was less frequent than by direct LDLc measurement (57%; χ² = 5.80, P = 0.016) (Table 4). However, apoB data are not representative of the total survey due to the lower number of reporting laboratories.

### 4. Discussion

In this study on a Dutch EQA database, representing all LDLc and HDLc methods used in the clinical laboratories in The Netherlands and most other countries, we observed marked deviations from the analytical target value (bias) with measurements in HTG sera. These

---

### Table 1

Analytical target values of SKML EQA-specimens as assigned with CDC Reference methods for total cholesterol, HDLc, and LDLc. ApoB values are traceable to WHO/IFCC standards.

<table>
<thead>
<tr>
<th>EQA-specimen</th>
<th>Consensus mean</th>
<th>CDC RMP</th>
<th>CDC BQ</th>
<th>Consensus mean</th>
<th>CDC RMP</th>
<th>LUMC</th>
<th>Non-HDLc (mmol/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012.1D</td>
<td>6.78</td>
<td>5.16</td>
<td>1.94</td>
<td>1.97</td>
<td>1.08</td>
<td>0.86</td>
<td>4.08</td>
</tr>
<tr>
<td>2012.1F</td>
<td>1.10</td>
<td>4.09</td>
<td></td>
<td>2.62</td>
<td>1.09</td>
<td>0.77</td>
<td>3.00</td>
</tr>
<tr>
<td>2012.2D</td>
<td>1.20</td>
<td>3.82</td>
<td>2.43</td>
<td>0.89</td>
<td>0.95</td>
<td>0.75</td>
<td>2.93</td>
</tr>
<tr>
<td>2012.2E</td>
<td>7.11</td>
<td>5.37</td>
<td>2.45</td>
<td>2.40</td>
<td>0.98</td>
<td>0.75</td>
<td>4.52</td>
</tr>
<tr>
<td>2012.2F</td>
<td>1.13</td>
<td>5.13</td>
<td>2.46</td>
<td>2.40</td>
<td>3.19</td>
<td>1.46</td>
<td>3.65</td>
</tr>
</tbody>
</table>

---

### Table 2

HDLc-adjusted risk scores simulated from direct HDLc methods in patients with initial SCORE 3% (women) and 4% (men).

<table>
<thead>
<tr>
<th>Method</th>
<th>Laboratories (N)</th>
<th>HDLc (mmol/l) median (range)</th>
<th>Imprecision (%CV)^a</th>
<th>Bias (mmol/l)^b mean (95% CI)</th>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SCORE (%)</td>
<td>SCORE &gt;5% N (%)</td>
</tr>
<tr>
<td>A. NTG sample 2012.1F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target value</td>
<td>1.09</td>
<td>1.10 (0.79–1.40)</td>
<td>6.1%</td>
<td>+0.01 (0.00; 0.02)</td>
<td>4.2 (3.0–5.4)</td>
<td>6 (3%)</td>
</tr>
<tr>
<td>Overall</td>
<td>197</td>
<td>1.10 (0.79–1.40)</td>
<td>8.3%</td>
<td>+0.01 (0.00; 0.04)</td>
<td>4.2 (3.6–5.4)</td>
<td>2 (11%)</td>
</tr>
<tr>
<td>Abbott</td>
<td>18</td>
<td>1.10 (0.79–1.40)</td>
<td>8.3%</td>
<td>+0.01 (0.00; 0.04)</td>
<td>4.2 (3.6–5.4)</td>
<td>2 (11%)</td>
</tr>
<tr>
<td>Beckman</td>
<td>39</td>
<td>1.04 (0.92–1.20)</td>
<td>4.4%</td>
<td>+0.05 (0.00; 0.03)</td>
<td>4.2 (3.5–5.1)</td>
<td>5 (5%)</td>
</tr>
<tr>
<td>Olympus</td>
<td>8</td>
<td>1.05 (1.01–1.10)</td>
<td>3.6%</td>
<td>+0.04 (0.01; 0.06)</td>
<td>4.2 (4.2–4.5)</td>
<td>0</td>
</tr>
<tr>
<td>Roche</td>
<td>114</td>
<td>1.11 (0.99–1.30)</td>
<td>4.2%</td>
<td>+0.03 (0.02; 0.04)</td>
<td>3.9 (3.3–4.5)</td>
<td>0</td>
</tr>
<tr>
<td>Siemens</td>
<td>14</td>
<td>1.10 (0.95–1.40)</td>
<td>9.6%</td>
<td>+0.02 (0.00; 0.08)</td>
<td>4.2 (3.0–5.1)</td>
<td>2 (14%)</td>
</tr>
<tr>
<td>B. HTG sample 2012.1D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target value</td>
<td>1.08</td>
<td>1.09 (0.62–1.25)</td>
<td>10.6%</td>
<td>+0.01 (0.05; 0.05)</td>
<td>4.2 (4.2–4.5)</td>
<td>0</td>
</tr>
<tr>
<td>Overall</td>
<td>196</td>
<td>0.99 (0.62–1.25)</td>
<td>10.6%</td>
<td>+0.01 (0.05; 0.05)</td>
<td>4.2 (4.2–4.5)</td>
<td>0</td>
</tr>
<tr>
<td>Abbott</td>
<td>18</td>
<td>1.06 (0.98–1.10)</td>
<td>4.3%</td>
<td>+0.03 (0.03; 0.05)</td>
<td>4.2 (4.2–4.5)</td>
<td>0</td>
</tr>
<tr>
<td>Beckman</td>
<td>39</td>
<td>1.00 (0.79–1.17)</td>
<td>6.5%</td>
<td>+0.07 (0.01; 0.10)</td>
<td>4.3 (4.2–5.4)</td>
<td>9 (23%)</td>
</tr>
<tr>
<td>Olympus</td>
<td>8</td>
<td>0.98 (0.92–1.03)</td>
<td>2.8%</td>
<td>+0.07 (0.04; 0.10)</td>
<td>4.5 (4.5–5.1)</td>
<td>2 (25%)</td>
</tr>
<tr>
<td>Roche</td>
<td>113</td>
<td>0.98 (0.68–1.25)</td>
<td>6.6%</td>
<td>+0.07 (0.03; 0.10)</td>
<td>5.1 (4.2–6.0)</td>
<td>11 (98%)</td>
</tr>
<tr>
<td>Siemens</td>
<td>14</td>
<td>0.79 (0.62–1.20)</td>
<td>15.3%</td>
<td>+0.07 (0.03; 0.10)</td>
<td>5.4 (4.2–6.6)</td>
<td>13 (93%)</td>
</tr>
</tbody>
</table>

---

NTG, normotriglyceridemic; HTG, hypertriglyceridemic.

^P < 0.0001 compared to NTG sample (Wilcoxon test).

**P < 0.0001, χ² = 186.6 (women), 98.2 (men) compared to NTG sample (χ² test).

^ Inter-laboratory coefficient of variation of reported values.

^a HDLc—target HDLc.

^b Also includes Ortho Clinical Diagnostics HDLc methods.
tematic difference in results between a method and the reproducibility of a particular method while bias refers to a systematic imprecision and systematic bias. Precision refers to the discordances.

Direct LDLc measurements did not improve the particularly in HTG sera. The errors in HDL-c measurements also depending on the laboratory where HDLc has been measured, reference value. The discordance between measurements and ApoB values above high-risk treatment goal obtained with ApoB methods in a patient with SCORE

\[ \text{Table 4} \]

<table>
<thead>
<tr>
<th>Method</th>
<th>Laboratories (N)</th>
<th>ApoB (g/l) median (range)</th>
<th>Imprecision (%CV)</th>
<th>Bias (g/l)</th>
<th>ApoB &gt;1.00 g/l N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. NTG sample 2012.2F</td>
<td>Overall</td>
<td>0.92</td>
<td>0.94 (0.82–1.09)</td>
<td>8.5%</td>
<td>+0.01 (-0.01; +0.04)</td>
</tr>
<tr>
<td></td>
<td>Abbott</td>
<td>0.85 (0.82–0.91)</td>
<td>5.0%</td>
<td>-0.07 (-0.11; -0.02)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Beckman</td>
<td>0.94 (0.88–1.00)</td>
<td>4.3%</td>
<td>-0.02 (-0.01; +0.05)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Roche</td>
<td>0.97 (0.83–1.08)</td>
<td>6.8%</td>
<td>+0.04 (0.00; +0.08)</td>
<td>2 (14%)</td>
</tr>
<tr>
<td></td>
<td>Siemens</td>
<td>0.90 (0.83–1.09)</td>
<td>9.9%</td>
<td>+0.02 (-0.06; +0.10)</td>
<td>1</td>
</tr>
<tr>
<td>B. HTG sample 2012.2E</td>
<td>Overall</td>
<td>0.98</td>
<td>0.99 (0.84–1.17)</td>
<td>8.8%</td>
<td>+0.01 (-0.02; +0.04)</td>
</tr>
<tr>
<td></td>
<td>Abbott</td>
<td>0.92 (0.84–1.17)</td>
<td>11.2%</td>
<td>-0.01 (-0.15; +0.16)</td>
<td>4 (7%)</td>
</tr>
<tr>
<td></td>
<td>Beckman</td>
<td>0.95 (0.93–1.03)</td>
<td>3.9%</td>
<td>-0.01 (-0.05; +0.02)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Roche</td>
<td>1.00 (0.85–1.09)</td>
<td>7.3%</td>
<td>+0.03 (-0.03; +0.09)</td>
<td>5 (16%)</td>
</tr>
<tr>
<td></td>
<td>Siemens</td>
<td>1.00 (0.90–1.10)</td>
<td>8.1%</td>
<td>+0.04 (-0.05; +0.10)</td>
<td>3 (38%)</td>
</tr>
</tbody>
</table>

NTG, normotriglyceridemic; HTG, hypertriglyceridemic.

\( ^a \) P = 0.007 bias compared to NTG sample (Wilcoxon test).

\( ^b \) Also includes Ortho Clinical Diagnostics and Olympus HDLc methods.

\( ^c \) Inter-laboratory coefficient of variation of reported values.

\( ^d \) Also includes Ortho Clinical Diagnostics, Olympus, and Siemens LDLc methods.

Table 3

<table>
<thead>
<tr>
<th>Method</th>
<th>Laboratories (N)</th>
<th>LDLc (mmol/l) median (range)</th>
<th>Imprecision (%CV)</th>
<th>Bias (mmol/l)</th>
<th>LDLc &gt;2.50 mmol/l N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Calculated LDLc, NTG sample 2012.2D</td>
<td>Target value</td>
<td>2.43</td>
<td>2.40 (2.10–3.00)</td>
<td>6.2%</td>
<td>-0.01 (-0.04; +0.03)</td>
</tr>
<tr>
<td></td>
<td>Overall ( ^a )</td>
<td>123</td>
<td>2.40 (2.28–2.56)</td>
<td>4.3%</td>
<td>-0.02 (-0.09; +0.06)</td>
</tr>
<tr>
<td></td>
<td>Abbott</td>
<td>7</td>
<td>2.37 (2.20–2.60)</td>
<td>4.2%</td>
<td>-0.04 (-0.10; +0.02)</td>
</tr>
<tr>
<td></td>
<td>Beckman</td>
<td>24</td>
<td>2.40 (2.26–3.00)</td>
<td>6.7%</td>
<td>+0.03 (-0.02; +0.09)</td>
</tr>
<tr>
<td></td>
<td>Siemens</td>
<td>8</td>
<td>2.25 (2.10–2.50)</td>
<td>7.5%</td>
<td>-0.16 (-0.42; +0.12)</td>
</tr>
<tr>
<td>B. Direct LDLc, NTG sample 2012.2D</td>
<td>Target value</td>
<td>2.43</td>
<td>2.44 (2.28–2.70)</td>
<td>3.4%</td>
<td>0.00 (-0.02; +0.02)</td>
</tr>
<tr>
<td></td>
<td>Overall ( ^a )</td>
<td>49</td>
<td>2.43 (2.40–2.46)</td>
<td>1.1%</td>
<td>+0.01 (+0.03; +0.03)</td>
</tr>
<tr>
<td></td>
<td>Abbott</td>
<td>5</td>
<td>2.43 (2.40–2.70)</td>
<td>5.0%</td>
<td>+0.01 (+0.05; +0.10)</td>
</tr>
<tr>
<td></td>
<td>Beckman</td>
<td>5</td>
<td>2.44 (2.28–2.55)</td>
<td>3.1%</td>
<td>0.00 (-0.04; +0.02)</td>
</tr>
<tr>
<td>C. Direct LDLc, HTG sample 2012.2E</td>
<td>Target value</td>
<td>2.46</td>
<td>2.40 (2.20–3.19)</td>
<td>6.3%</td>
<td>-0.06 (-0.10; -0.01) ( ^a )</td>
</tr>
<tr>
<td></td>
<td>Overall ( ^a )</td>
<td>47</td>
<td>2.88 (2.70–2.90)</td>
<td>3.0%</td>
<td>+0.39 (+0.28; +0.49)</td>
</tr>
<tr>
<td></td>
<td>Abbott</td>
<td>5</td>
<td>2.80 (2.75–3.19)</td>
<td>6.8%</td>
<td>+0.34 (+0.29; +0.67)</td>
</tr>
<tr>
<td></td>
<td>Beckman</td>
<td>5</td>
<td>3.33 (2.35–2.70)</td>
<td>3.7%</td>
<td>-0.16 (-0.20; -0.13)</td>
</tr>
</tbody>
</table>

NTG, normotriglyceridemic; HTG, hypertriglyceridemic.

\( ^a \) P = 0.007 bias compared to NTG sample (Wilcoxon test).

\( ^b \) Also includes Ortho Clinical Diagnostics and Olympus HDLc methods.

\( ^c \) Inter-laboratory coefficient of variation of reported values.

\( ^d \) Also includes Ortho Clinical Diagnostics, Olympus, and Siemens LDLc methods.

biases exceeded far beyond the U.S. National Cholesterol Education Program (NCEP) recommendations for LDLc (<4%) and HDLc (<5%) [21,22]. We simulated that application of HDL-multipliers of SCORE may yield different risk classifications (moderate vs. high risk) depending on the laboratory where HDLc has been measured, particularly in HTG sera. The errors in HDL-c measurements also affect calculated LDLc, leading to discordant treatment goals even in NTG sera. Direct LDLc measurements did not improve the discordances.

Analytical laboratory error can be divided into two components: random imprecision and systematic bias. Precision refers to the reproducibility of a particular method while bias refers to a systematic difference in results between a method and the “true” or reference value. The discordance between measurements and target values, observed by us, is caused by both bias (which should be avoided) and imprecision (which can only be minimized but not avoided) of the tests. Since the target values are very close to the decision point, one would expect that also an ideal test without any bias will inevitably cause some misclassifications due to analytical imprecision even in NTG sera.

Manufacturers of lipid assays certify and standardize their assays by comparison with a CRMNL laboratory. The CRMNL laboratories employ LDLc and HDLc reference measurement procedures that are traceable to the CDC reference methods, i.e., betatitation/heparin-Mn\(^{2+}\) precipitation/Abell-Kendall cholesterol analysis for HDLc [23,24]. This process ensures that the calibrators and reagents sold by manufacturers produce test results that are traceable to the CDC reference methods. Although total cholesterol standardization is generally viewed as a success, concerns remain about the effectiveness of standardization programs for LDLc and HDLc. An important limitation of the current CDC standardization protocol is the lack of testing with specimens from individuals with dyslipidemias to better evaluate “real world” assay performance.
Of particular concern for the isolation and quantification of LDL and HDL is the heterogeneity of the lipoprotein fractions. Both LDL and HDL comprise different subclasses of particles that vary in size, density, shape, lipid and apolipoprotein composition [5], making development of specific assays difficult. Direct assays based on different principles may measure different subclasses of LDL or HDL that may or may not be equally quantified, depending on the assay procedure and reagents (non-specificity bias) [5–8]. In contrast to calibration bias, non-specificity bias cannot be overcome with better calibration; it is inevitable and varies per sample. Most discrepancies are observed in samples from patients with hypertriglyceridemia, mixed dyslipidemia, or other conditions with altered lipoprotein composition and remodeling such as diabetes and kidney disease [5–8]. Miller et al. [9,25] and Miida et al. [26] observed that direct LDLc and HDLc measurements in dyslipidemic samples frequently failed to meet the NCEP analytical performance criteria [21,22]. In other studies, direct LDLc methods did not offer advantage over calculated LDLc in classifying patients into NCEP risk categories in a dyslipidemic population, when compared to the reference method [27–29], while non-HDLc calculation and apoB measurement showed better concordance in risk classification [29]. In our EQA database, we were unable to simulate non-HDLc-based classification in treatment groups because none of the calculated values in HTG sera encompassed the high-risk goal.

It is important to realize that current guidelines and risk estimation models rely on epidemiological observations using early precipitation methods for HDLc measurement and LDLc calculation [5,8]. Therefore, the clinician should be aware that the values currently reported by clinical laboratories with direct methods are not measured by the same methods as those of risk cutpoints recommended in guideline documents. However, considering the heterogeneity of the lipoproteins, there is no evidence that the particular LDL and HDL fractions obtained by earlier methods are better indicators of CVD risk than the fractions obtained by current methods [5,30–32].

In addition, intraindividual biological variability should be considered when using lipid methods for risk score classification [15,33]. The NCEP guidelines recommend that at least 2 serial specimens, 1 week apart, are necessary to reduce biological variation. The relative change of the two results can be used to determine whether additional patient specimens are required because of unusual high variation. Studies demonstrate that on lipid retesting, subjects may change one risk category and even two risk categories, such as from low to high risk, or vice versa [15].

Even when the NCEP recommendations are followed, it remains probable that misclassifications occur, leading to unnecessary expensive treatment or lack of appropriate treatment. Difficulties arise when lipid values are near critical values that determine the partition between adjacent risk categories, particularly at levels of SCORE around 5% because the decision of therapeutic intervention depends on whether the patients are below or above the 5% threshold of “high risk” [1]. For example, in patients with initial SCORE 3–4%, an underestimation of HDLc may falsely bring the SCORE to a value above 5%. In these cases, the clinician may wrongly recommend drug treatment when lifestyle changes alone may be a more appropriate action. Conversely, for a true SCORE of 6%, possibilities of falsely reporting a HDL-adjusted SCORE <5% and thus undertreatment exist because of the wide risk distributions resulting from analytical error. In a patient at high risk, statin therapy is recommended if LDLc is ≥2.50 mmol/l [1]. Hence underestimation or overestimation of the treatment goal LDL lead to insufficient or exaggerated LDL reduction, respectively.

From the results of our study, it is clear that direct LDLc and HDLc measurements should be interpreted as potentially unreliable when making clinical decisions in patients with hypertriglyceridemia. However, it is important to emphasize that we tested a rather severely HTG serum (7 mmol/l). For example in the Copenhagen City Heart Study less than 2% of the population had nonfasting TG above 5 mmol/l [34]. Accordingly in the fasting state, the prevalence will be even lower. Therefore and because most people with increased risk because of low HDLc have moderate (2–5 mmol/l) rather than severe hypertriglyceridemia (>5 mmol/l), the problem of biased measurements and misclassification of CVD risk and target values is probably smaller than indicated in our simulation, although risk misclassifications have been reported to occur in ~20–30% of moderate HTG sera [29].

In clinical practice, LDLc is either estimated by the Friedewald formula or directly measured with a homogeneous assay. As the calculation is based on serum cholesterol, TG, and HDLc, it includes the sum of errors in all three measurements. The calculation is not valid for specimens with TG >4.5 mmol/l, type III hyperlipoproteinemia, or fasting chylomicronemia [30]. Furthermore, if there is free glycerol in the blood (i.e., in patients with diabetes) falsely elevated TG may be measured, so the result of calculated LDLc is probably too low [35]. Errors in HDLc measurement also affect the calculation of the secondary treatment goal non-HDLc. Therefore not only LDLc but also non-HDLc goals must be taken with caution in severe hypertriglyceridemia.

As the EAS/ESC guidelines allow the use of apoB as a secondary treatment goal, we recommend apoB measurements in case of severe hypertriglyceridemia and/or invalid Friedewald equation. In our study the effect of hypertriglyceridemia on apoB measurements showed less misclassifications than with direct LDLc, although the data cannot be compared to direct LDLc data because apoB reflects not only LDL but the total number of atherogenic particles including very low density lipoproteins (VLDL), intermediate density lipoproteins (IDL), and lipoprotein (a) [36]. ApoB is a clearly defined measurand, and its measurement can be standardized across laboratories worldwide because of the availability of WHO reference materials and methods [19]. The clinical utility of apoB testing is similar to LDLc in estimating CVD risk [37–39]. ApoB concentration has better concordance with LDLc than non-HDLc when determining the need for lipid-lowering therapy [40,41], but its (cost-) efficacy to improve the prediction of treatment success beyond LDLc or non-HDLc needs to be confirmed [42–45].

It is obvious, that even a perfect classification of a patient’s lipoprotein profile does not guarantee adequate clinical diagnosis and monitoring for every individual. In clinical practice, the ultimate risk judgment and management should remain the role of the clinician. Calculated SCOREs must always be interpreted in the context of the patient’s unique situation and underlying risk factors such as unhealthy lifestyle, obesity, family history, low socio-economic status, and psychosocial risk factors. Therefore, the decision to change a patient’s risk category on the basis of his or her HDLc must be based on clinical judgment. Similarly, when making decisions on initiation of drug therapy on the basis of LDLc, particularly at borderline LDLc value near the treatment goal, clinicians need to consider with the patient whether the magnitude of the expected benefit would outweigh the disadvantage of statin therapy, e.g., side-effects, the disutility of taking a pill every day, and financial burden.

The true evaluation of a risk biomarker, beyond its prognostic value, relates to its effectiveness in helping the clinician to improve patient outcome and its cost-effectiveness with social and economic healthcare implications [39]. Even modest changes in risk prediction with a disease as common as CVD translates into thousands of people in Europe that may or may not be treated adequately. The HDLc and LDLc inter-method differences we observed are clinically relevant considering the prevalence of risk scores in a representative Western-European population in which
the SCORE model can be applied for [46]. Among 6212 non-diabetic Belgian men and women aged 40–74 years, free of CVD, 14.2% of the population had a SCORE 3–4% and thus are candidates for HDL-adjusted risk reclassification [46]. In the same study, 17.8% of the population had a high-risk SCORE (5–9%) in whom choice of therapeutic intervention depends on accuracy of LDLc measurement [46]. Because no population cohort and clinical outcome data are available in our study, we cannot assess the clinical accuracy of the risk classification simulated with the various methods.

Other lipoprotein variables such as small dense LDL (not investigated in the present study) may potentially help clinicians to better manage the risk in the group of HTG subjects [47]. Lipoprotein-associated phospholipase A2 (Lp-PLA2), a pro-inflammatory enzyme primarily associated with small dense LDL, is another promising marker for identifying "hidden" high-risk patients; the measurement of Lp-PLA2 mass or activity is related specifically to vascular inflammation and rupture-prone plaque [45,48]. Precise and fully automated methods are now commercially available for rapid measurements of small dense LDL and Lp-PLA2 mass and activity on routine laboratory instruments, which allow the analysis of large numbers of samples and facilitate the evaluation of the clinical utility of these emerging risk markers in ongoing trials. However, epidemiological studies regarding the added value of small dense LDL, Lp-PLA2 mass, or Lp-PLA2 activity beyond traditional risk markers have been inconsistent, likely because analytical methods and lipoprotein particle definitions have not yet been standardized, thus complicating the interpretation of studies that have used different methods [49,50]. This lack of standardization is a major disadvantage of the new biomarkers compared to apolipoprotein measurements [19]. Although there are previous analytical studies that evaluated methodologies used to measure the emerging biomarkers, there is no detailed comparison of the different methodologies, as done in the present study. The present study may represent important ground work for future investigations in the field.

5. Conclusion

As pointed out in this simulation study and previously by Warnick, Nauck, and Rifai [6], “Laboratories supporting lipid clinics with a high proportion of specimens with atypical lipoproteins could observe discrepant results on certain specimens that might confound treatment decisions.” It is vitally important for clinical laboratories to consider assay reliability and specificity when choosing methods, particularly in dyslipidaemic samples. Additionally, more efforts are needed to address specificity issues in the HDLc and LDLc manufacturers’ certification programs offered by CDC and the CRMLN reference laboratories. As a patient’s HDLc and LDLc concentrations depend on the method chosen by the laboratory where it is measured, therapeutic decision cutpoints and multipliers of risk may not be considered universally applicable. This should be taken into account when deciding to treat a patient with lipid-lowering drug therapy for life.

Acknowledgments

We thank Alberico Catapano (EAS President), John Chapman (EAS Past-president), and Sverre Sandberg (EFLM Chair of Scientific Division) for establishing the EAS-EFLM Collaborative Project, intended to produce laboratory testing recommendations and guidelines for management of dyslipidaemia and prevention of CVD.

References


