Clinica Chimica Acta 496 (2019) 68–75

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/cca

Validation of the novel Martin method for LDL cholesterol estimation

Efi Petridou, Kostas Anagnostopoulos*

Laboratory of Biochemistry, Department of Medicine, Democritus University of Thrace, 68100 Alexandroupolis, Greece

ARTICLE INFO

Keywords:
LDL
Cholesterol
Triglycerides
HDL
Friedewald
Validation

ABSTRACT

Introduction: Accurate determination of low-density lipoprotein cholesterol (LDL) is important for coronary heart disease (CHD) risk assessment among others. A low-cost method used is the Friedewald equation, calculating LDL from total cholesterol after subtracting high-density lipoprotein cholesterol (HDL) and triglycerides divided by 5. A new calculation method has been proposed where the value of 5 is not fixed but depends on the values of the other parameters.

Results: We validated this method in a Greek population sample, by comparing direct LDL, the Friedewald equation and the novel method. Some clinical laboratories use the direct determination when TG > 200 mg/dl (2.26 mmol/lt). We performed segmented linear regression to check if this value makes sense. Bayesian linear regression was performed to compare the direct determination to the Friedewald and novel one.

Conclusions: We found that TG > 200 mg/dl is a sensible threshold value since it is a saddle point for the standard error of the regression. For Bayesian linear regression, the results were inconclusive. When the LDL values were used for classification of CHD, it turned out that the novel method was better than the Friedewald equation at correctly classifying LDL levels for CHD risk assessment.

1. Introduction

Low density lipoprotein cholesterol (LDL) is a significant risk factor for atherosclerotic vascular disease [1,2] that can be modified either by pharmaceutical intervention or lifestyle changes. In the Adult Treatment Panel III (ATP-III) of the third report of the National Cholesterol Education Panel (NCEP), it is classified in 5 categories (from optimal to very high) according to its measured value for each patient and is a risk factor for coronary heart disease (CHD). Many laboratories measure LDL not directly but by calculation from total cholesterol (CHOL), high-density lipoprotein cholesterol (HDL) and the triglycerides (TG) using the Friedewald Eq. [3], as LDL = CHOL-HDL-(TG/5), where TG/5 approximates the very low density lipoprotein cholesterol (VLDL), which is a good approximation when LDL < 400 mg/dl (4.52 mmol/lt). The reference method for LDL determination (as well as for the other fractions of cholesterol) involves ultracentrifugation and is therefore not well suited for routine application [4]. There are also direct enzymatic assays for LDL that have been developed [5,6], can be automated and are used in the routine of the clinical laboratory.

Martin et at in a study [7] proposed a novel method involving the differentiation of the denominator of the (TG/5) fraction in the Friedewald equation, whose value is not to be fixed at 5, but to be dependent on the values of TG and non-HDL cholesterol (nonHDL = CHOL-HDL). In the Martin paper this value can range from 3.2 to 24.0, being a function of the values of TG and nonHDL (classification tables with 10 to 2000 cells were studied), and found that the 180-cell table was an optimum for number of cells and method performance.

However, there is a lot of inertia when changes in clinical determination methods are proposed, especially when the method under scrutiny has been used for a long time with satisfactory results. To advocate and promote new methods or changes to existing ones used globally, independent verification using multiple data sets is required. Two studies have independently validated the novel Martin method in general populations, an American [8] and a Korean one [9]. Other studies have also compared the formulas in population subgroups, such as diabetics [10], subjects eligible for statin treatment [11], statin-treated patients [12,13] and patients with familial combined hyperlipidemia [14].

Most laboratories in Greece use the Friedewald equation, mostly when TG < 400 mg/dl (4.52 mmol/lt). Some laboratories set this threshold at 200 mg/dl (2.26 mmol/lt), above which they use a direct enzymatic method and others use the direct method exclusively. We used data from the «Sismanoglio" General Hospital of Komotini in Greece that uses a direct enzymatic method in order to compare the novel method to the Friedewald one. We also performed segmented

* Corresponding author.
E-mail address: kanagnos@med.duth.gr (K. Anagnostopoulos).

https://doi.org/10.1016/j.cca.2019.06.023
Received 3 June 2019; Received in revised form 25 June 2019; Accepted 27 June 2019
Available online 29 June 2019
0009-8981/

linear regression to see if the threshold values of 200 mg/dl (2.26 mmol/lt) or 400 mg/dl (4.52 mmol/lt) above which to use direct determination of LDL has merit and if when the novel method is as good as the direct one.

2. Materials and methods

2.1. Measurements

Data from the biochemical laboratory of the "Sismanoglio" General Hospital of Komotini, Greece were used after approval from its Board. The data were anonymous, containing only the values of CHOL, HDL, TG and enzymatically directly determined LDL (henceforth referred to as LDL_d), as well as the age and the gender of the subjects. All samples were from overnight-fasting individuals. All parameters were determined using an ADVIA 1800 Chemistry Systems by Siemens. The LDL values calculated according to Friedewald are referred to as LDL_f. Also, the LDL values using the novel method were calculated using the 180-, 360- and 2000- cell tables (referred to as LDL_{180}, LDL_{360} and LDL_{2000}, respectively) [Suppl. Section 3.1]. The enzymatically directly determined LDL is referred to as LDL_d. Non-HDL cholesterol (nonHDL) was calculated by subtracting HDL from total cholesterol. $VLDL_f$ was calculated by dividing TG by 5.

2.2. Statistical analysis

Analysis of the data were performed using the R programming language version 3.5.3 [15]. Bayesian analysis was performed using Gibbs sampling with the RJAGS interface to JAGS [16]. Hypothesis testing for the difference between group means was done by determining the Bayes factor [17]. The Bayes Factor (BF) is a numerical value that quantifies how well a hypothesis predicts the empirical data (alternative hypothesis H_1) relative to a competing hypothesis (null hypothesis H_0). E.g., if the $BF_{H1/H0}$ is 3, this means that the observed data is 3 times more probable if H_1 were true than if H_0 were true. Therefore, the experimental evidence points towards H_1. Although it is not straightforward and it is arbitrary to put an interpretation for the strength of the evidence based on the value of the BF, according to [18] one possible interpretation is as follows: When BF is between 1 and 3 the strength of the evidence in favor of H_1 is weak, not worth more than a bare mention. If BF is between 3 and 20 the evidence for H_1 is positive. When BF is between 20 and 150 there is strong evidence and when > 150 there is very strong evidence for H_1. Of course when $BF < 1$, the evidence supports H_0. All analyses are also shown in the supplementary material sections, denoted as [Suppl. Section Number].

Segmented linear regression was performed in order to find out if there are critical values of TG upon which the quality of the regression (prediction of the value of LDL_d from LDL_f and TG) suffers. In segmented linear regression there exist knot points on which the slope and intercept of the regression line change. The model used in the regression was:

$$LDL_f = intercept + slope \cdot LDL_d \cdot I_{TG < knot}$$

where $I_{TG < knot}$ is the indicator function (1 when TG $<$ knot value and 0 otherwise). The above model describes a linear regression only when the TG value is below a certain threshold.

Bayesian linear regression analysis was performed to test the response of the calculated LDL methods to the direct one, using LDL_d as the independent variable and the calculated LDL methods as the dependent ones (with each regression performed separately). In linear regression, usually the independent variable (LDL_d in this case) values are considered to be exactly known. However, this is clearly not the case for LDL_d and TG, which (as clinical laboratory measurements) have their own random error. We therefore performed the Bayesian linear regression by treating not only the calculated LDL variables as random, but the LDL_d as well. We also factored in the model the stratification of TG levels that resulted from the segmented linear regression. The model used was:

$$\begin{aligned} LDL_{calculated} &= \beta_0 + \beta_1 \cdot LDLd + \beta_2 \cdot I_{TG < 195} + \beta_3 \cdot I_{TG < 195} \cdot TG + \beta_4 \cdot I_{195 \le TG < 315} + \beta_5 \\ &\quad \cdot I_{195 \le TG < 315} \cdot TG + \beta_6 \cdot I_{315 \le TG < 452} + \beta_7 \cdot I_{315 \le TG < 452} \cdot TG + \beta_8 \\ &\quad \cdot I_{TG \ge 452} + \beta_9 \cdot I_{TG \ge 452} \cdot TG \end{aligned}$$

where $LDL_{calculated}$ is either LDL_f, LDL_{180}, LDL_{360} or $LDL2000$. The slope and intercept of the model depend not only on the dependent ($LDL_{calculated}$) and the two random independent (LDL_d and TG) variables but also on the TG values. When e.g. TG $<$ 195 all but the first 3 terms of eq. (2) become zero. In Table 4, the values of the intercept and slopes of the model by TG level are shown.

Table 1
Descriptive statistics of the study cohort.

	Male (%)	Female (%)	
Gender	1945 (42.3%)	2653 (57.7%)	
	Median	2.5 Percentile	97.5 Percentile
Total Cholesterol	167	84	277
HDL Cholesterol	43	15	82
Triglycerides	118	51	349
LDL_f	94	33	189
LDL_d	100	31	198
LDL_{180}	98	35	190
LDL_{360}	98	35	190
LDL_{2000}	98	36	190
$VLDL_f$	24	10	70

3. Results

3.1. Study cohort and comparison between various LDL values

The study cohort consisted of 4598 measurements. Of these, 1945 (42.3%) were female and 2653 (57.7%) were male (Table 1) [Suppl Section 4.1]. They were mainly of middle age (age median 67 years, 2.5–97.5 percentile = 1–91 years) [Suppl Section 4.2]. In supplementary fig. S1 the distribution of the ages is shown. 568 (12.4%) of the patients were 1–9 years old, 70 (1.5%) were 11–17 years, 1435 (31.2%) were between 18 and 64 years and the majority (2525, 54.9% were of age 65 years or older). Notable is also that 142 patients (3.1%) and 316 (6.9%) were 1 and 2 years old, respectively [Suppl Section 4.3].

The Friedewald LDL (LDL_f) as well as the LDL values based on the novel method were calculated. In Table 1 the median and the 2.5 and 97.5 percentiles for the various LDL parameters as well as the rest of the parameters (CHOL, HDL and TG) are shown. [Suppl Section 4.2]. In Fig. 1 and [Suppl Section 4.4] the densities of all parameters except gender and age are depicted. For the LDL parameters the median and the 2.5 and 97.5 percentiles are also shown in the figure.

We can see that the density distribution of LDL_d is skewed towards higher values compared to LDL_f. The same is the case for LDL_{180}, LDL_{360} and LDL_{2000}, which are also higher than LDL_f but not as high as LDL_d.

We performed Bayesian testing to check if the differences between the various LDL parameters are statistically significant, by determining the Bayes Factor. Results are shown in Table 2 and [Suppl Section 4.5]. We can see that LDL_d is statistically significantly different from LDL_f, LDL_{180}, LDL_{360} and LDL_{2000}. The same is the case for the difference of LDL_f from the rest of the LDL parameters. On the other hand the statistical differences between the novel methods are from very weak to non-existent.

3.2. Determination of critical TG values

For economic reasons many clinical laboratories in Greece (as well

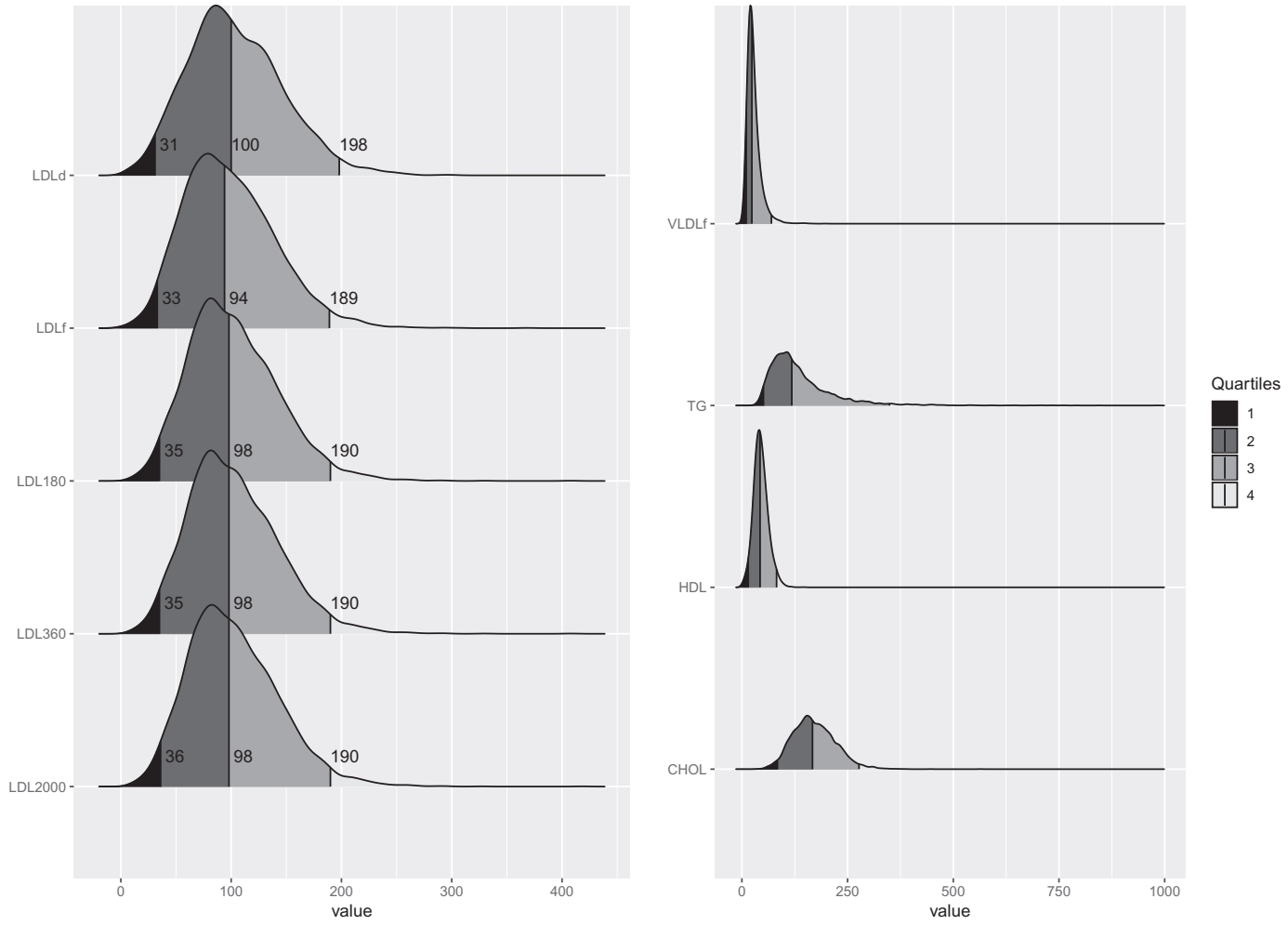

Fig. 1. Distribution densitites of LDL_d, LDL_f, LDL_{180}, LDL_{360}, LDL_{2000}, $VLDL_f$, triglycerides, HDL and total cholesterol. The vertical lines and adjacent numbers are (from left to right) the 2.5, 50 (median) and 97.5 percentiles.

Table 2
Bayes factors for the statistical significance for the difference between the various LDL parameters (the table is symmetric).

	LDL_d	LDL_f	LDL_{180}	LDL_{360}	LDL_{2000}
LDL_d	–	$1.32 \cdot 10^{115}$	$4.42 \cdot 10^{10}$	$3.90 \cdot 10^{10}$	$5.52 \cdot 10^{10}$
LDL_f	$1.32 \cdot 10^{115}$	–	$1.63 \cdot 10^{198}$	$7.76 \cdot 10^{197}$	$3.36 \cdot 10^{195}$
LDL_{180}	$4.42 \cdot 10^{10}$	$1.63 \cdot 10^{198}$	–	0.0213	3.51
LDL_{360}	$3.90 \cdot 10^{10}$	$7.76 \cdot 10^{197}$	0.0213	–	1.94
LDL_{2000}	$5.52 \cdot 10^{9}$	$3.36 \cdot 10^{195}$	3.15	1.94	–

as the rest of the world), even if they have the necessary reagents for the direct determination of LDL, they do not perform it by default but only if the TG levels exceed a threshold, usually 200 mg/dl (2.26 mmol/lt) or in rarer cases 400 mg/dl (4.52 mmol/lt). To find out if there are critical values of TG upon which the quality of the regression (prediction of the value of LDL_d from LDL_f and TG) suffers, we performed segmented linear regression with the threshold value not fixed, but set to move from the minimum value of TG in our dataset to the maximum one (therefore, many linear regressions were performed, each one with a different knot point). For each regression, the R^2 and the standard error of the regression (SER) were calculated [Suppl Section 5.1]. The plot of the SER versus the TG knot point is shown in Fig. 2 and [Suppl Section 5.2]. We can see that there are 3 saddle points (local minima, corresponding to the dashed vertical lines) at TG values of 195, 315 and 452 mg/dl (2.54, 4.10 and 5.88 mmol/lt, respectively) [Suppl Section 5.3].

3.3. Comparison of LDL by TG level

We subsequently categorized the LDL values in 4 levels (TG < 195 (2.20), 195 ≤ TG < 315 (3.56), 315 ≤ TG < 452 (5.11) and TG ≥ 452 (5.11) mg/dl (mmol/lt)). In supplementary Fig. S2 [Suppl Section 5.3], the scatter plot of LDL_f versus LDL_d by these TG levels is shown. The majority of points (3835, 83.4%) are for TG < 195 mg/dl (2.20 mmol/lt). In supplementary fig. S3 [Suppl Section 6.1] the densities of the various LDL measurements by TG classification are shown and the medians and 2.5 and 97.5 percentiles of these parameters are in the [Suppl Section 6.2] table. We subsequently tested for statistical significance of the differences between the LDL parameters by TG level. Results are in Table 3 and [Suppl Section 6.3]. For TG < 195 mg/dl and 195 ≤ TG < 315 mg/dl, things are the same as for the whole dataset (significant differences for LDL_d and/or LDL_f from the rest, no significant differences between the novel calculations), although for 195 ≤ TG < 315 mg/dl the Bayes factors are not as high. However, for 315 ≤ TG < 452 mg/dl only a weak statistical significance was found for the difference LDL_d - LDL_f and very strong difference between LDL_{180} and LDL_{360}. For TG ≥ 452 mg/dl no significant difference was evident for LDL_d – LDL_f. However, the number of measurements available for these TG levels is relatively low to reach any definite conclusions. For LDL_{180} – LDL_{360} a Bayes factor could not be calculated because the values were exactly the same and therefore the difference is

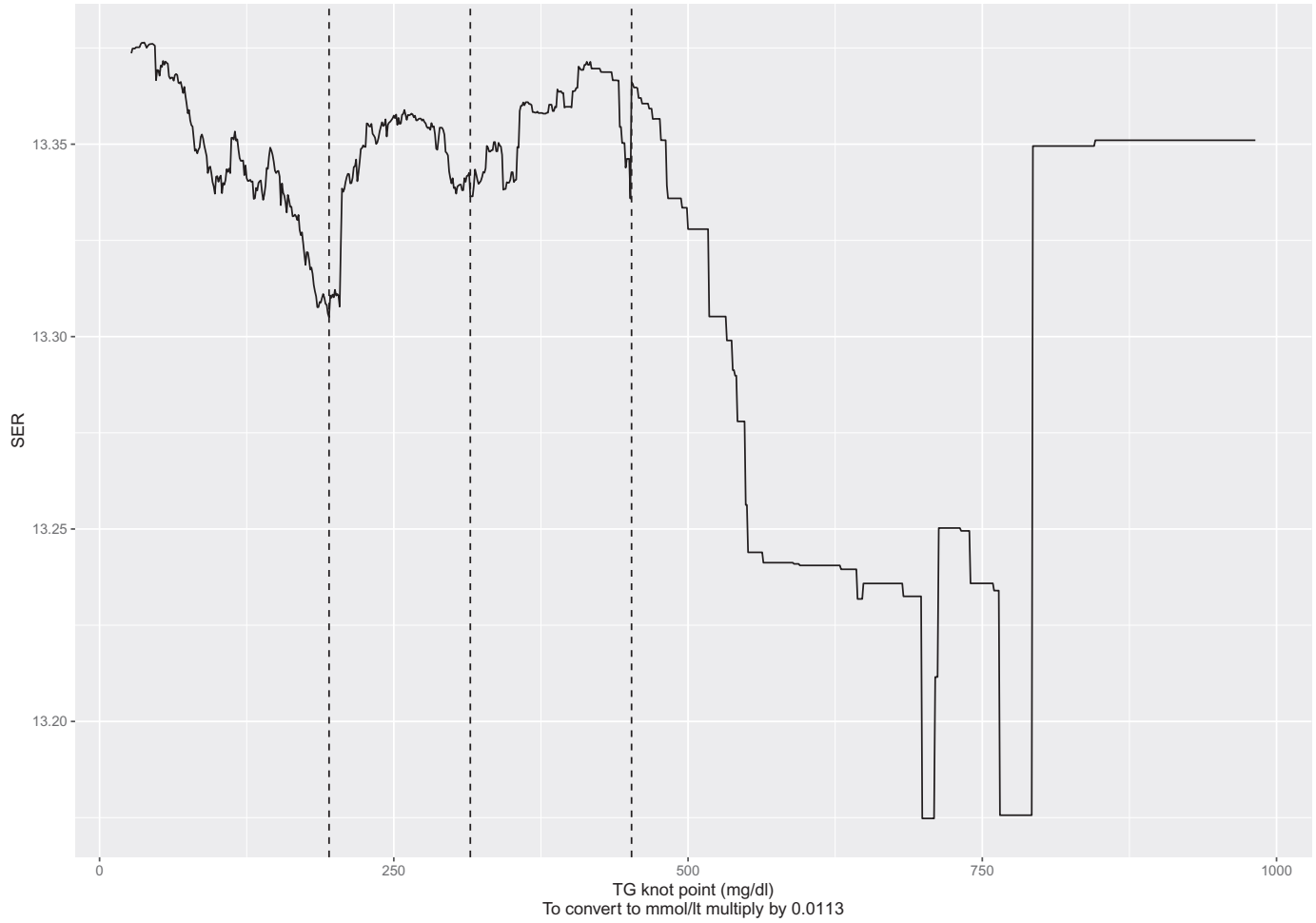

Fig. 2. Standard error of the regression (SER) for the segmented linear regression versus changing TG knot point value.

Table 3
Bayes factors for the statistical significance for the difference between the various LDL parameters by TG level (the tables are symmetric).

	LDL_d	LDL_f	LDL_{180}	LDL_{360}	LDL_{2000}
TG < 195 mg/dl (N = 3835)					
LDL_d	–	$8.87 \cdot 10^{129}$	$1.45 \cdot 10^{75}$	$4.05 \cdot 10^{75}$	$4.97 \cdot 10^{73}$
LDL_f	$8.87 \cdot 10^{129}$	–	$1.45 \cdot 10^{75}$	$5.12 \cdot 10^{84}$	$9.59 \cdot 10^{87}$
LDL_{180}	$1.45 \cdot 10^{75}$	$1.45 \cdot 10^{75}$	–	0.0316	0.369
LDL_{360}	$4.05 \cdot 10^{75}$	$5.12 \cdot 10^{84}$	0.0316	–	4.92
LDL_{2000}	$4.97 \cdot 10^{73}$	$9.59 \cdot 10^{87}$	0.369	4.92	–
195 ≤ TG < 315 (N = 603)					
LDL_d	–	$1.64 \cdot 10^{14}$	965	918	1590
LDL_f	$1.64 \cdot 10^{14}$	–	$1.31 \cdot 10^{232}$	$1.34 \cdot 10^{240}$	$4.00 \cdot 10^{239}$
LDL_{180}	965	$1.31 \cdot 10^{232}$	–	0.0603	0.210
LDL_{360}	918	$1.34 \cdot 10^{240}$	0.0603	–	0.969
LDL_{2000}	1590	$4.00 \cdot 10^{239}$	0.210	0.969	–
315 ≤ TG < 452 (N = 121)					
LDL_d	–	4.56	60,500	177,000	69,200
LDL_f	4.56	–	$1.97 \cdot 10^{58}$	$4.78 \cdot 10^{58}$	$1.16 \cdot 10^{61}$
LDL_{180}	60,500	$1.97 \cdot 10^{58}$	–	14,300	0.204
LDL_{360}	177,000	$4.78 \cdot 10^{58}$	14,300	–	0.204
LDL_{2000}	69,200	$1.16 \cdot 10^{61}$	0.204	2.99	–
TG ≥ 452 (N = 39)					
LDL_d	–	0.173	2730	2730	4590
LDL_f	0.173	–	$3.21 \cdot 10^{22}$	$3.21 \cdot 10^{22}$	$1.50 \cdot 10^{19}$
LDL_{180}	2730	$3.21 \cdot 10^{22}$	–	–	35
LDL_{360}	2730	$3.21 \cdot 10^{22}$	–	–	35
LDL_{2000}	4590	$1.50 \cdot 10^{19}$	35	35	–

0.

3.4. Correlation of LDL_d with the other LDL variables

To test how well the calculated (both Friewewald and novel formulae) LDL values correlate with LDL_d, we performed Bayesian linear regression analysis using LDL_d as the independent variable and the calculated LDL methods as the dependent ones (each regression was performed separately), treating all LDL variables as random. In the model, the TG stratification imposed above was also factored in, as described in the Statistical analysis subsection of Material and Methods. In Table 4, the values of the intercept and slopes of the model by TG level are shown.

The minimum effective sample size was calculated at 8831 [Suppl Section 7.1]. Since in some parameters there was a high degree of autocorrelation, all models were initialized for 10,000 iterations and run for 150,000 iterations so that the effective sample size could be at least reached and possibly exceeded [Suppl Section 7.2]. The model files used in JAGS are also in the supplementary material. The models

Table 4
Intercepts and slopes of linear model defined by eq. (2).

TG level	Intercept	LDL_d slope	TG slope
TG < 195	$\beta_0 + \beta_2$	β_1	β_3
195 ≤ TG < 315	$\beta_0 + \beta_4$		β_5
352 ≤ TG < 452	$\beta_0 + \beta_6$		β_7
TG ≥ 452	$\beta_0 + \beta_8$		β_9

Table 5
R^2 and standard error of the regression for the 4 Bayesian linear regression models with LDL_d and TG by levels as the independent variables and the calculated LDL parameters as the dependent variable.

	LDL_f	LDL_{180}	LDL_{360}	LDL_{2000}
R^2	0.895	0.897	0.898	0.897
SER	13.34	12.75	12.73	12.71

converged properly after the run. The standardized residuals plot is in supplementary fig. S4 [Suppl Section 7.4], where we can see that most points are within ± 2SD, although a few points have significant deviations from the mean and could be considered as outliers. However, since in order to remove an outlier there must be some justification (which is not forthcoming here), these points can be considered as natural patient variation and were not excluded from the analysis. The R^2 and SER for the 4 models are shown in Table 5 and [Suppl Section 7.5.1].

We can see mainly from the SER values that the regression improves when using the novel calculation compared to the Friedewald one (SER from 13.34 becomes 12.75 for LDL_{180}). On the other hand, using higher cell-count tables (LDL_{360} and LDL_{2000}) improves the SER very slightly (12.73 and 12.71 respectively). The same pattern is evident when we calculated the SER by LDL method and TG levels [Suppl Section 7.5.2], although for 195 ≤ TG < 315 there was a notable improvement of the SER for LDL_{2000}.

The posterior distributions of the intercept and slopes are shown in Figs. 3 and 4 and in [Suppl Section 8.3.1 and 8.3.2], respectively. The vertical lines and the numbers in the figures represent the 2.5, 50 (median) and 97.5 percentiles. For the intercept a pattern is evident across the various TG levels: the intercepts for the novel calculations are lower than that of the Friedewald method, although even in this case they are not high (95% credible interval-CrI = 1.88–4.68). For TG < 195 the credible intervals between LDL_f and novel methods do not overlap, but for the rest of the TG levels they do overlap, although the distributions are still shifted towards lower values. Between the novel methods (180-, 360- and 2000- cell tables) there is almost complete overlap and there seems to be no statistically significant difference. The same pattern appears for the LDL_d slope (Fig. 4), which is nearer to 1 for LDL_f (95% CrI = 0.89–0.91) than that of the novel methods. For the TG slopes the opposite pattern is evident: the TG slopes for the novel methods are higher than those for LDL_f.

3.5. NCEP classification

We then classified the LDL levels by NCEP III, which classifies it in 5 risk levels according to its value [2]. In Table 6 and [Suppl Section 9.1] the NCEP III LDL percentages by class of the 4598 total samples with LDL belonging to each one by LDL method are shown (the subscripts in the NCEP abbreviation denote the method used corresponding to LDL method determination, e.g. $NCEP_d$: NCEP classification by LDL_d).

In Table 7 and [Suppl Section 9.2], the Bayes factor proportional ratio tests of each LDL calculation method are shown compared to the $NCEP_d$ classification which is used as the reference method. The higher the BF, the more likely it is that the differences in the proportions are statistically significant.

First, it is evident from both the percentages and the comparison BFs that there are no significant differences between the various novel methods (180-, 360 and 2000 -cell tables). Between the direct LDL determination and the Friedewald calculation there are significant differences for all NCEP classifications except the “Near optimal/above optimal” level. The evidence for the existence of these differences, after inspection of the BF values, is very strong (for “Optimal”, “Borderline high” and “High”) or positive (for “Very high”). Between LDL_d and the novel methods there is evidence of very strong difference for “High” and somewhat positive for “Very high”. On the other hand, when comparing the Friedewald to the novel methods, very strong evidence of difference appears only for the “Optimal” classification.

We can see in Table 6 that the Friedewald method classifies a larger

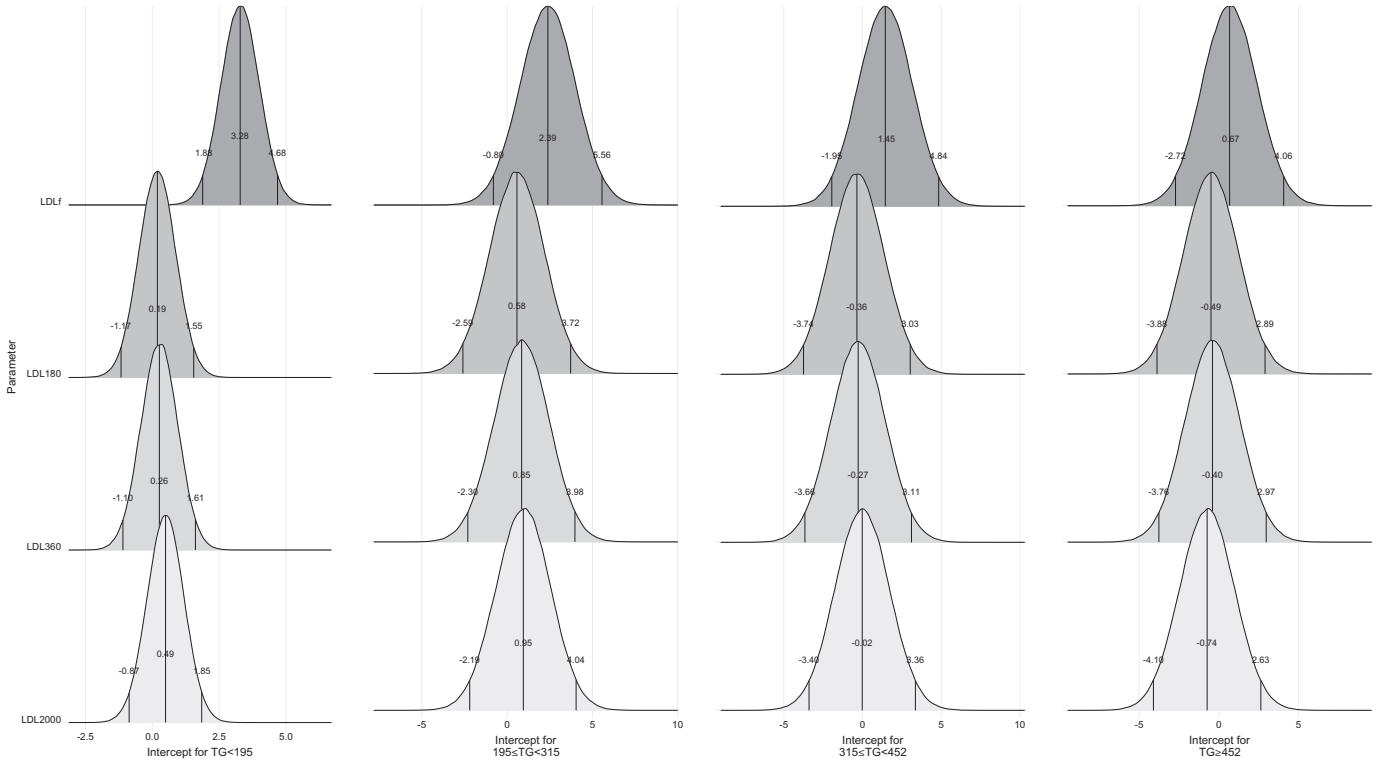

Fig. 3. Posterior distributions for the intercepts of the Bayesian linear regression for the various TG classification levels. The vertical lines and adjacent numbers are (from left to right) the 2.5, 50 (median) and 97.5 percentiles.

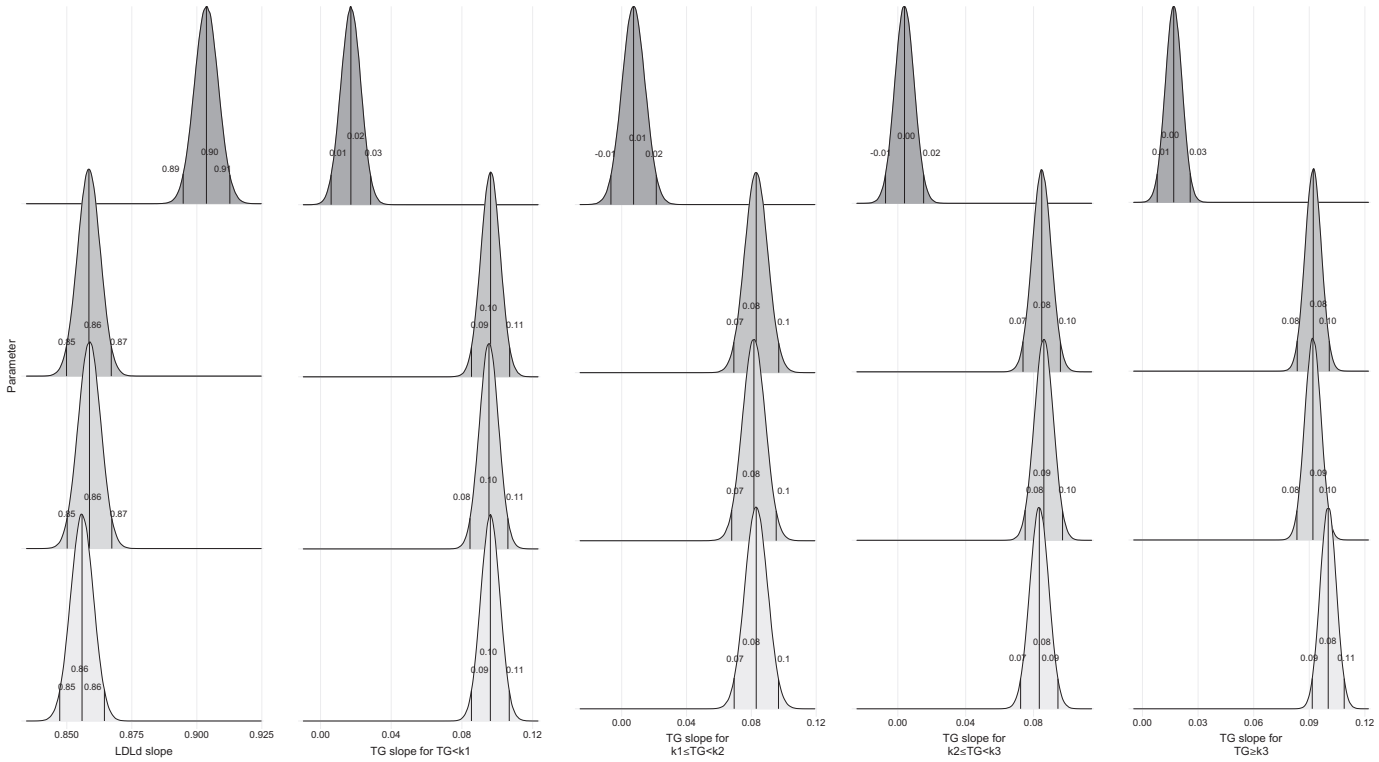

Fig. 4. Posterior distributions for the slopes of the Bayesian linear regression for the various TG classification levels. The vertical lines and adjacent numbers are (from left to right) the 2.5, 50 (median) and 97.5 percentiles.

Table 6
Percentages of LDL NCEP III classification by LDL determination method.

LDL levels and classes	$NCEP_d$	$NCEP_f$	$NCEP_{180}$	$NCEP_{360}$	$NCEP_{2000}$
Optimal	49.91%	54.72%	51.50%	51.52%	51.52%
Near optimal/Above optimal	23.12%	23.38%	24.51%	24.51%	24.73%
Borderline high	16.18%	13.85%	15.31%	15.33%	15.20%
High	7.48%	5.55%	6.07%	6.02%	5.92%
Very high	3.31%	2.50%	2.61%	2.61%	2.63%

percentage of LDL levels as “Optimal” (significant difference), whereas the novel methods are closer to LDL_d. This leads to a lower percentage of the clinically worse classifications (“Borderline high” 13.85% versus 16.18%, “High” 5.55% versus 7.47%, “Very high” 2.50% versus 3.31% for the Friedewald and LDL_d, respectively). The Bayes factors for these differences are very high. For the novel methods there is also underestimation of the percentages for those levels, but the differences are not as pronounced and the Bayes factors for them (although still high) are much lower than the ones for the Friedewald method.

4. Discussion

We used a sample of 4598 measurements from a Greek population to validate the novel LDL determination method. Both the age and the gender composition were similar to those in two other studies comparing LDL_f to the novel methods in a general population [8,9] as well as to the original paper by Martin et al. [7]. LDL was determined directly using an enzymatic method and also calculated using both the Friedewald and the novel formulae.

First, we saw that there were statistically significant differences for LDL_d and LDL_f, both between them and from the novel LDLs (Fig. 1 and Table 2), with the LDL_d shifted towards higher values compared to LDL_f and the novel method being in between these two. This is similar to the finding in [8] where LDL_f tended to underestimate the true LDL values and the novel method to overestimate it and also in [9] where the novel method LDL estimates were closer to LDL_d than to LDL_f. However, it should be noted that although some statistical test gives significant results, it does not necessarily mean that this fact has practical implications. In the case of LDL determination, the practical implication is the categorization according to NCEP, discussed below. It is however a

Table 7
Bayes factors comparing the differences in classification percentages between the various LDL determination methods.

	Optimal	Near optimal/above optimal	Bordeline high	High	Very high
$NCEP_d$-$NCEP_f$	62084582****	0.0	735.8****	57662****	17.3**
$NCEP_d$-$NCEP_{180}$	0.4	0.5	0.2	85.7***	4.6**
$NCEP_d$-$NCEP_{360}$	0.4	0.5	0.2	135.3***	4.6**
$NCEP_d$-$NCEP_{2000}$	0.4	1.1	0.3	453.3****	3.6**
$NCEP_f$-$NCEP_{180}$	526.4****	0.2	2.7*	0.2	0.1
$NCEP_f$- $NCEP_{360}$	462.9****	0.2	3.0*	0.2	0.1
$NCEP_f$- $NCEP_{2000}$	462.9****	0.4	1.5*	0.1	0.1
$NCEP_{180}$- $NCEP_{360}$	0.0	0.0	0.1	0.1	0.1
$NCEP_{180}$- $NCEP_{2000}$	0.0	0.0	0.1	0.1	0.1
$NCEP_{360}$-$NCEP_{2000}$	0.0	0.0	0.1	0.1	0.1

Evidence for the difference *: not worth more than a bare mention, **: Positive, ***: Strong, ****: Very strong.

strong indicator that there may exist practical implications when using one or the other method.

When multiple segmented linear regressions correlating LDL_f to LDL_d by simultaneously factoring in increasing TG values were performed and the SER was plotted versus those TG values, three saddle points at TG values of 195, 315 and 452 mg/dl (2.54, 4.10 and 5.88 mmol/lt, respectively) were observed. Therefore, the use of TG = 200 mg/dl as a point above which to perform direct LDL determination, if available, instead of using the Friedewald equation, seems reasonable (as is common practice in many clinical chemistry laboratories). Furthermore, according to the American Heart Association, one of the criteria for dyslipidemia is TG levels > 150 mg/dl (1.70 mmol/lt) [19]. In potentially dyslipidemic patients it makes sense to go to greater lengths to ascertain more accurate levels of TG and lipids in general.

We then classified the various LDL measurements according to their respective TG value and tested for statistical significance of the difference of the stratified LDL values (Table 3). Up to TG values of 315 mg/dl (3.56 mmol/lt), − that is the two first strata, TG < 195 and 195 ≤ TG < 315 mg/dl - the differences were the same as those for the whole population. For 315 ≤ TG < 452 mg/dl the difference between LDL_d and LDL_f was statistically weak and for TG ≥ 452 mg/dl (5.11 mmol/lt) no significant difference was found. On the other hand a very strong difference between LDL_{180} and LDL_{360} was observed for the 315 ≤ TG < 452 stratum. Although, the Friedewald equation does not hold for TG values > 400 mg/dl [3], the values of LDL resulting from it may also be not very trustworthy at TG values less than but near to 400 mg/dl. As TG increases, the fraction [TG/denominator] that is calculated VLDL constitutes a larger proportion of the equation and errors in this fraction will result in larger relative errors in the resulting LDL_f [20]. Also, since the number of observations in these last two groups were relatively small compared to those with lower TG values, no definite conclusions can be drawn.

When Bayesian linear regression was performed to see how well LDL_d correlated with the calculated LDL variables, the SER improved when using the novel calculation with the 180-cell table (Table 5). Use of higher cell-count tables had no impactful improvement of the overall regression, which is also the case when the TG strata were taken under consideration [Suppl Section 7.5.2]. In Figs. 3 and 4 the posterior distributions of the intercept and slopes respectively, along with the median, 2.5 and 97.5 percentiles are shown. Although it is more convenient to use one or a few numbers for the result of a regression (e.g. the mean and standard error of the slope and intercept), in Bayesian analysis the best way to describe the result is to show the posterior distribution, since Bayesian analysis has the uncertainty in the results incorporated [21]. In Fig. 3 it is evident that the intercepts for the novel calculations are lower than that of the Friedewald method, although the 95% credible intervals do not overlap only for TG < 195. The LDL_d slope (Fig. 4) is higher for LDL_f that that of the novel methods, with no overlap. Oppositely, the TG slopes are lower for LDL_f than those for the novel methods, again with no overlap.

In a linear regression it is generally better for an intercept to be zero and the slope to be 1, because this means that the dependent variable is directly predicted from the independent one(s). In our case, the intercept for LDL_f is higher compared to that of the novel methods (which indicates that the novel methods would serve as a better predictor of LDL_d). On the other hand, the LDL_d slope for LDL_f as independent variable is nearer to 1 than that of the novel methods (which indicates the opposite, that LDL_f would be a better predictor for LDL_d). The TG slopes are lower for LDL_f (near zero) and higher (near 0.10) for the novel methods. A TG slope near zero for LDL_f means that the TG levels would not affect the regression of LDL_f as much when compared to that of the novel methods.

The previous results are somewhat contradictory. On the one hand we have the regression goodness-of-fit measure of SER and the intercept indicating that the novel methods are better than LDL_f at predicting the LDL_d values and on the other hand the regression slopes indicate the opposite. To resolve this conundrum and since what ultimately matters is not the LDL value itself but its risk classification (and therefore the effect it will have on a clinical intervention decision), we examined the percentages (Table 6) and the statistical significance of their differences (Table 7) in the classification of LDL levels according to NCEP III [2] which includes 5 levels (“Optimal”, “Near optimal/Above optimal”, “Borderline high”, “High”and “Very high”).

In Table 7, no significant differences between the novel methods were evident (Table 7). Oppositely, between LDL_d and LDL_f there were significant differences for all NCEP III levels except the “Near optimal/above optimal” one. Between LDL_d and the novel methods significant differences appeared for the “High” and “Very high” levels. Between LDL_d and the novel methods there significant differences appeared for the “High” and “Very high” levels. For the comparison of LDL_f to the novel methods, a very strong difference appears only for the “Optimal” classification. In Table 6, where the percentages of each NCEP III class are given by LDL determination method, the novel methods percentages are closer to LDL_d. This leads to a statistically significant lower percentage of the clinically worse classifications (“Borderline high”, “High”, “Very high”) when using LDL_f compared to that when using the novel methods. The same pattern, although formulated somewhat differently appears also in the concordance in Guideline Classification of the original Martin et al. paper [7], where the concordance of the novel LDL compared to the one determined by the reference ultracentrifugation method is higher than that of LDL_f. This is the case also in the validation of the novel method in an American polulation [8] and in a Korean one [9]. It seems therefore that the novel methods improve the agreement of LDL classification, bringing it closer to that of LDL_d when the LDL values used for this classification are calculated using the novel method compared to the Friedewald equation. It should be noted that this result holds only for the direct LDL determination method used here (Siemens ADVIA 1800 chemistry system). Since there is a significant variability between the various direct methods for LDL [22,23], we can only be certain that this holds for the specific assay, although it is highly likely that the same is the case for other direct enzymatic assays.

5. Conclusions

There are two main conclusions we can draw from the above results. First, the practice in many clinical laboratories to determine LDL from total cholesterol, HDL and triglycerides using a calculation method and not direct determination when TG is less that 200 mg/dl (2.26 mmol/lt) seems warranted as evidenced by the existence of a saddle point at 195 mg/dl in segmented linear regression for the prediction of LDL_d from LDL_f and TG. Of course, if possible it would be preferable to use direct LDL determination in all measurements.

The second conclusion is that the direct determination of LDL used in this study is a better method for NCEP III classification compared to calculated methods (unfortunately, the reference method -involving ultracentrifugation- is too labor intensive and time consuming to be used for routine lipid screening). If however the direct LDL determination is not available (for financial or other reasons), the novel method should be employed instead of the traditional Friedewald calculation, since it classifies results more in accordance with LDL_d. Therefore the novel equation improves concordance in risk classification of calculated LDL with the direct LDL assay employed in this paper. Which cell-count table is used (180-, 360- or 2000- cell) does not make a difference; the 180-cell table is as good as the higher cell count tables. On the other hand, since the novel method will be implemented by software code and not calculated by hand (which is almost always the case even for the Friedewald calculation which is computationally much less demanding than the novel one), the higher cell-count tables could be used as well.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cca.2019.06.023.

References

[1] D. Steinberg, Thematic review series: the pathogenesis of atherosclerosis: an interpretive history of the cholesterol controversy, part III: mechanistically defining the role of hyperlipidemia, J. Lipid Res. (2005), https://doi.org/10.1194/jlr.R500010-JLR200.

[2] National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III), Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) Final Report., Circulation, (2002).

[3] W.T. Friedewald, R.I. Levy, D.S. Fredrickson, Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge, Clin. Chem. 18 (6) (1972) 499–502.

[4] K.R. Kulkarni, Cholesterol profile measurement by vertical auto profile method, Clin. Lab. Med. (2006), https://doi.org/10.1016/j.cll.2006.07.004.

[5] M. Rasouli, H. Mokhtari, Calculation of LDL-cholesterol vs. direct homogenous assay, J. Clin. Lab. Anal. (2017), https://doi.org/10.1002/jcla.22057.

[6] S. Vanavanan, P. Srisawasdi, M. Rochanawutanon, J. Kerdmongkol, M.H. Kroll, Comparison of a direct enzymatic assay and polyacrylamide tube gel electrophoresis for measurement of small, dense low-density lipoprotein cholesterol, Clin. Chem. Lab. Med. (2015), https://doi.org/10.1515/cclm-2014-0396.

[7] S.S. Martin, M.J. Blaha, M.B. Elshazly, P.P. Toth, P.O. Kwiterovich, R.S. Blumenthal, S.R. Jones, Comparison of a novel method vs the Friedewald equation for estimating low-density lipoprotein cholesterol levels from the standard lipid profile, JAMA (2013), https://doi.org/10.1001/jama.2013.280532.

[8] J.W. Meeusen, A.J. Lueke, A.S. Jaffe, A.K. Saenger, Validation of a proposed novel equation for estimating LDL cholesterol, Clin. Chem. (2014), https://doi.org/10.1373/clinchem.2014.227710.

[9] J. Lee, S. Jang, H. Son, Validation of the martin method for estimating low-density lipoprotein cholesterol levels in Korean adults: findings from the Korea national health and nutrition examination survey, 2009-2011, PLoS One (2016), https://doi.org/10.1371/journal.pone.0148147.

[10] H. Chaen, S. Kinchiku, M. Miyata, S. Kajiya, H. Uenomachi, T. Yuasa, K. Takasaki, M. Ohishi, Validity of a novel method for estimation of low-density lipoprotein cholesterol levels in diabetic patients, J. Atheroscler. Thromb. (2016), https://doi.org/10.5551/jat.35972.

[11] D. Shin, C. Bohra, K. Kongpakpaisarn, Novel method versus the Friedewald method for estimating low-density lipoprotein cholesterol in determination of the eligibility for statin treatment for primary prevention in the United States, Med. (United States). (2018), https://doi.org/10.1097/MD.0000000000010612.

[12] T. Sonoda, T. Takumi, M. Miyata, D. Kanda, I. Kosedo, S. Yoshino, M. Ohishi, Validity of a novel method for estimating low-density lipoprotein cholesterol levels in cardiovascular disease patients treated with statins, J. Atheroscler. Thromb. (2018), https://doi.org/10.5551/jat.44396.

[13] M.K. Palmer, P.J. Barter, P. Lundman, S.J. Nicholls, P.P. Toth, B.W. Karlson, Comparing a novel equation for calculating low-density lipoprotein cholesterol with the Friedewald equation: A VOYAGER analysis, Clin. Biochem. (2019), https://doi.org/10.1016/j.clinbiochem.2018.10.011.

[14] R. Mehta, E. Reyes-Rodríguez, O. Yaxmehen Bello-Chavolla, A.C. Guerrero-Díaz, A. Vargas-Vázquez, I. Cruz-Bautista, C. A. Aguilar-Salinas, Performance of LDL-C calculated with Martin's formula compared to the Friedewald equation in familial combined hyperlipidemia, Atherosclerosis. (2018), https://doi.org/10.1016/j.atherosclerosis.2018.06.868.

[15] R. Core Team, R: A Language and Environment for Statistical Computing, (2019).

[16] M. Plummer, JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling, Int. Work. Distrib. Stat. Comput. (2003) (doi:ISSN 1609-395X).

[17] S.N. Goodman, Toward evidence-based medical statistics. 2: the Bayes factor, Ann. Intern. Med. (1999), https://doi.org/10.7326/0003-4819-130-12-199906150-00019.

[18] R.E. Kass, A.E. Raftery, Bayes factors, J. Am. Stat. Assoc. (1995), https://doi.org/10.1080/01621459.1995.10476572.

[19] R.E.W. Kavey, S.R. Daniels, R.M. Lauer, D.L. Atkins, L.L. Hayman, K. Taubert, American Heart Association guidelines for primary prevention of atherosclerotic cardiovascular disease beginning in childhood, J. Pediatr. (2003), https://doi.org/10.1067/mpd.2003.205.

[20] L. Thomas, E.D. Peterson, The value of statistical analysis plans in observational research: defining high-quality research from the start, JAMA (2012), https://doi.org/10.1001/jama.2012.9502.

[21] R. Willink, R. White, Disentangling Classical and Bayesian Approaches to Uncertainty Analysis, (2011), pp. 1–19 http://www.bipm.org/cc/CCT/Allowed/26/Disentangling_uncertainty_v14.pdf.

[22] W.G. Miller, G.L. Myers, I. Sakurabayashi, L.M. Bachmann, S.P. Caudill, A. Dziekonski, S. Edwards, M.M. Kimberly, W.J. Korzun, E.T. Leary, K. Nakajima, M. Nakamura, G. Nilsson, R.D. Shamburek, G.W. Vetrovec, G.R. Warnick, A.T. Remaley, Seven direct methods for measuring HDL and LDL cholesterol compared with ultracentrifugation reference measurement procedures, Clin. Chem. (2010), https://doi.org/10.1373/clinchem.2009.142810.

[23] M.R. Langlois, M.J. Chapman, C. Cobbaert, S. Mora, A.T. Remaley, E. Ros, G.F. Watts, J. Borén, H. Baum, E. Bruckert, A. Catapano, O.S. Descamps, A. Von Eckardstein, P.R. Kamstrup, G. Kolovou, F. Kronenberg, A. Langsted, K. Pulkki, N. Rifai, G. Sypniewska, O. Wiklund, B.G. Nordestgaard, Quantifying atherogenic lipoproteins: current and future challenges in the era of personalized medicine and very low concentrations of ldl cholesterol. A consensus statement from EAS and EFLM, Clin. Chem. (2018), https://doi.org/10.1373/clinchem.2018.287037.