

A Modified Sampson–NIH Equation with Improved Accuracy for Estimating Low Levels of Low-Density Lipoprotein-Cholesterol

Maureen Sampson ^{a,†} Rafael Zubiran ^{b,*†} Anna Wolska ^b Jeffrey W. Meeusen,^c Leslie J. Donato,^c Allan S. Jaffe ^{c,d} Giorgio E.M. Melloni,^e Robert P. Giugliano,^e Marc S. Sabatine,^e Nicholas A. Marston ^e and Alan T. Remaley ^b

BACKGROUND: Cardiovascular guidelines have long recommended low-density lipoprotein-cholesterol (LDL-C) as the primary target for lipid-lowering therapy. Recent guidelines have emphasized the importance of achieving low LDL-C levels; hence, the accurate measurement of low LDL-C is increasingly clinically relevant.

METHODS: Using lipid panel test results from the Mayo Clinic (n = 24 590) and the FOURIER clinical trial of evolocumab (n = 9605), the following modified Sampson equation was developed by least-squares regression to match LDL-C (mg/dL) by the β -quantification reference method, by combining terms into non High Density Lipoprotein Cholesterol (nonHDL-C = Total Cholesterol – High Density Lipoprotein Cholesterol) and forcing the coefficient to be one:

$$\text{modified Sampson LDL-C} = \text{nonHDL-C} - \frac{TG}{8.37} - \frac{(TG \times \text{nonHDL-C})}{2640} + \frac{TG^2}{17\,400}$$

RESULTS: The modified Sampson equation demonstrated significant improvement in its concordance to the reference method compared to other equations (the Lin Concordance Correlation Coefficient 0.992, $P < 0.001$). By overall kappa analysis, it showed the best agreement to the reference method at the 55 mg/dL cutpoint (1.4 mmol/L, 0.98 [$P < 0.001$], Sampson–NIH: 0.96, Martin–Hopkins: 0.96, Friedewald: 0.94) and the 70 mg/dL cutpoint (1.8 mmol/L, 0.97 [$P < 0.001$], Sampson–NIH: 0.94, Martin–Hopkins: 0.95, Friedewald: 0.92). The false classification rate of the modified Sampson equation was also significantly lower compared to the other equations at 55 mg/dL (15%, [$P < 0.001$], Sampson–NIH: 29%, Martin–Hopkins: 28%, Friedewald: 37%) and 70 mg/dL (18%, [$P < 0.001$]; Sampson–NIH: 30%, Martin–Hopkins: 28%, Friedewald: 34%). The new equation increases the percentage of correctly classified patients with low LDL-C by approximately 10% to 20% over the other equations based on its net reclassification index.

CONCLUSIONS: The modified Sampson equation shows improved accuracy compared to other equations for low LDL-C. It more accurately identifies high-risk patients, who are not at their LDL-C goals and could benefit from more intensive lipid-lowering therapy.

^aDepartment of Laboratory Medicine, Clinical Center, National Institutes of Health, Bethesda, MD, United States; ^bLipoprotein Metabolism Laboratory, Translational Vascular Medicine Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, United States; ^cDepartment of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, United States; ^dDepartment of Cardiovascular Medicine, Mayo Clinic, Rochester, MN, United States; ^eTIMI Study Group, Division of Cardiovascular Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States.

*Address correspondence to this author at: Lipoprotein Metabolism Laboratory, Translational Vascular Medicine Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, 10 Center Drive, Bldg. 10/Rm. 8N220, Bethesda, MD 20892, United States. E-mail rafael.zubiran@nih.gov.

[†]Authors contributed equally and share first authorship.

Disclaimer: This publication is solely the work of the authors and should not be interpreted to represent the views, perspectives, or policies of the U.S. Department of Health and Human Services, the National Institutes of Health, or the National Heart, Lung, and Blood Institute. Received March 11, 2025; accepted May 27, 2025. <https://doi.org/10.1093/clinchem/hvaf073>

Introduction

Low-density lipoprotein cholesterol (LDL-C) has long been endorsed by cardiovascular guidelines as a key biomarker for atherosclerotic cardiovascular disease (ASCVD) risk assessment. The 2018 US multi-society guidelines (1) also recommended LDL-C as an aid in clinical management. In high-risk individuals, the strategy is to reduce LDL-C by $\geq 50\%$. For very high-risk patients (secondary prevention and patients with diabetes), a threshold of 70 mg/dL (1.8 mmol/L) is used to consider adding other lipid-lowering therapies on top of statins. In case of severe primary hypercholesterolemia (LDL-C ≥ 190 mg/dL, [> 4.9 mmol/L]), if LDL-C levels remain ≥ 100 mg/dL (2.6 mmol/L) on high-

intensity statins, adding ezetimibe should be considered. More recent recommendations from the United States (2) and Europe (3) have proposed even lower LDL-C thresholds, namely 70 mg/dL (1.8 mmol/L) for high-risk and 55 mg/dL (1.4 mmol/L) for very high-risk individuals, based on evidence that more aggressive LDL-C lowering is both safe and effective (4, 5). For ASCVD patients who experience a second vascular event within 2 years, an even lower LDL-C goal of <40 mg/dL (1 mmol/L) should be considered, according to European guidelines (3). The development of new lipid-lowering drugs like inhibitors for proprotein convertase subtilisin-kexin type 9 (PCSK9i) has provided new treatment options that now make it possible to achieve these very low concentrations of LDL-C (6). This has made the accurate measurement of low LDL-C critical in the care of patients and for determining eligibility for insurance coverage of the new lipid-lowering medications (7).

The two most widely used methods for LDL-C (8) are direct (homogenous) assays or equations for calculating LDL-C based on the results of the standard lipid panel (total cholesterol [TC], triglycerides [TG], and high-density lipoprotein cholesterol [HDL-C]). Although direct HDL-C assays are routinely used by most clinical laboratories, direct LDL-C assays are not as widely used, particularly by smaller laboratories and by point-of-care devices. The analytical limitations of some direct LDL-C assays (9, 10) and their extra expense compared to a free calculation have limited their utilization.

Until recently, LDL-C was almost exclusively calculated by the Friedewald (F) equation ($\text{LDL-C} = \text{TC} - \text{HDL-C} - \text{TG}/5$), which subtracts HDL-C and estimated cholesterol on very low-density lipoproteins (VLDL-C = $\text{TG}/5$) from TC (11). In 2013, the Martin–Hopkins (M) equation was developed and shown to be more accurate than the F-equation (12). The two equations are nearly identical except for the use of a variable factor by the M-equation for dividing TG. This factor varies, depending on the TG and nonHDL-C level, and was empirically determined by the vertical auto profile (VAP) ultracentrifugation method (13). In 2020, the Sampson–NIH (S) equation (14) was developed by least-squares regression analysis to match the β -quantification (BQ) reference method for LDL-C (15), which is used by regulatory agencies throughout the world to standardize routine assays for LDL-C. The BQ method, like the VAP method, utilizes ultracentrifugation to separate lipoproteins, but also an LDL precipitation step (16). The VAP method, however, under-recovers VLDL-C on hypertriglyceridemic samples compared to the BQ method (13, 17). The S-equation is in the form of a bivariate (nonHDL-C and TG) quadratic equation and was shown to be more accurate than other LDL-C equations, especially on hypertriglyceridemic samples (14, 18–21).

Recently, the enhanced Sampson (eS) equation (19) was reported to be the most accurate estimation method. It incorporates apolipoprotein B (apoB) as an independent variable along with the standard lipid panel. This equation was specifically developed to improve LDL-C estimation in cases with atypical lipoprotein profiles, such as dysbetalipoproteinemia, where standard equations are unreliable and additional laboratory testing with apoB is required (19).

Given the growing importance of accurately estimating low LDL-C in the clinical management of patients (5), we investigated if we could improve upon the accuracy of the S-equation by including more test results from patients with low LDL-C. Using a combined database of BQ LDL-C results from the Mayo Clinic and the FOURIER clinical trial of the PCSK9i evolocumab (6, 22), we describe here a modified S-equation (mS) based on the lipid panel. It has improved accuracy for low LDL-C, while still maintaining good performance throughout the rest of the LDL-C range.

Materials and Methods

Deidentified lipid panel tests, BQ-derived LDL-C, and apoB from patients, for whom these tests were ordered for routine care at the Mayo Clinic, were used for analysis ($n = 24,590$, median LDL-C 116 mg/dL [3 mmol/L], inter quartile range [IQR] 90–146 mg/dL, [2.5–3.77 mmol/L]) (23). The treatment status of these patients was not available. The same lipid tests ($n = 9605$, median LDL-C 24 mg/dL [0.62 mmol/L], IQR 16–32 mg/dL, [0.41–0.82 mmol/L]) from the treatment arm of the FOURIER clinical trial were also utilized (6). Samples with TG above 3000 mg/dL (33.9 mmol/L) ($n = 39$), BQ LDL-C <1 mg/dL (0.02 mmol/L) ($n = 17$), or dysbetalipoproteinemia (TG 150–1000 mg/dL [1.7–11.3 mmol/L] and BQ VLDL-C/TG >0.3, $n = 116$) were excluded to prevent skewing model performance, as these lipoprotein profiles require specialized approaches not generalizable to broader populations. The combined data set was randomly divided into half to create training and validation data sets. To convert to SI units from mg/dL to mmol/L, divide the results for TC, LDL-C, and nonHDL-C by 38.67; for triglyceride, divide by 88.57.

The following mS-equation was established with the training data set by least-squares regression analysis, using BQ LDL-C as the dependent variable to calculate in mg/dL:

$$\text{modified Sampson LDLC} = \text{nonHDL-C} - \frac{\text{TG}}{8.37} - \frac{(\text{TG} \times \text{nonHDL-C})}{2640} + \frac{\text{TG}^2}{17400}$$

Regression error characteristic (REC) analysis (24) was performed as previously described. Agreement between the various LDL-C equations with BQ LDL-C for classifying patients at different LDL-C cutpoints (± 10 mg/dL [0.25 mmol/L] around cutpoint) was determined in the validation data set by kappa analysis (25). LDL-C was calculated as per the original equation descriptions, and for the M-equation the extended panel of factors for hypertriglyceridemic samples was used (11, 12, 14, 19, 26). All analysis was done with JMP software, R software, or with Excel. An Excel spreadsheet for calculating LDL-C by the various equations in either mg/dL or mmol/L can be downloaded at the following website: <https://doi.org/10.6084/m9.figshare.28557038>. The analysis of deidentified laboratory test results in this study was considered nonhuman subject research. A more detailed description of the methods can be found in the supplement.

Results

The median BQ LDL-C value for the combined database used in this study was 96 mg/dL (2.5 mmol/L), with 35% of LDL-C results below 70 mg/dL (1.8 mmol/L) (online Supplemental Table 1). In Fig. 1A, we examined how well the original S-equation compared against BQ LDL-C in the validation data set. A close linear relationship was observed, but as can be seen by the TG percentile color code for each data point, the greatest deviations (both high and low) from BQ LDL-C occur when TG is elevated.

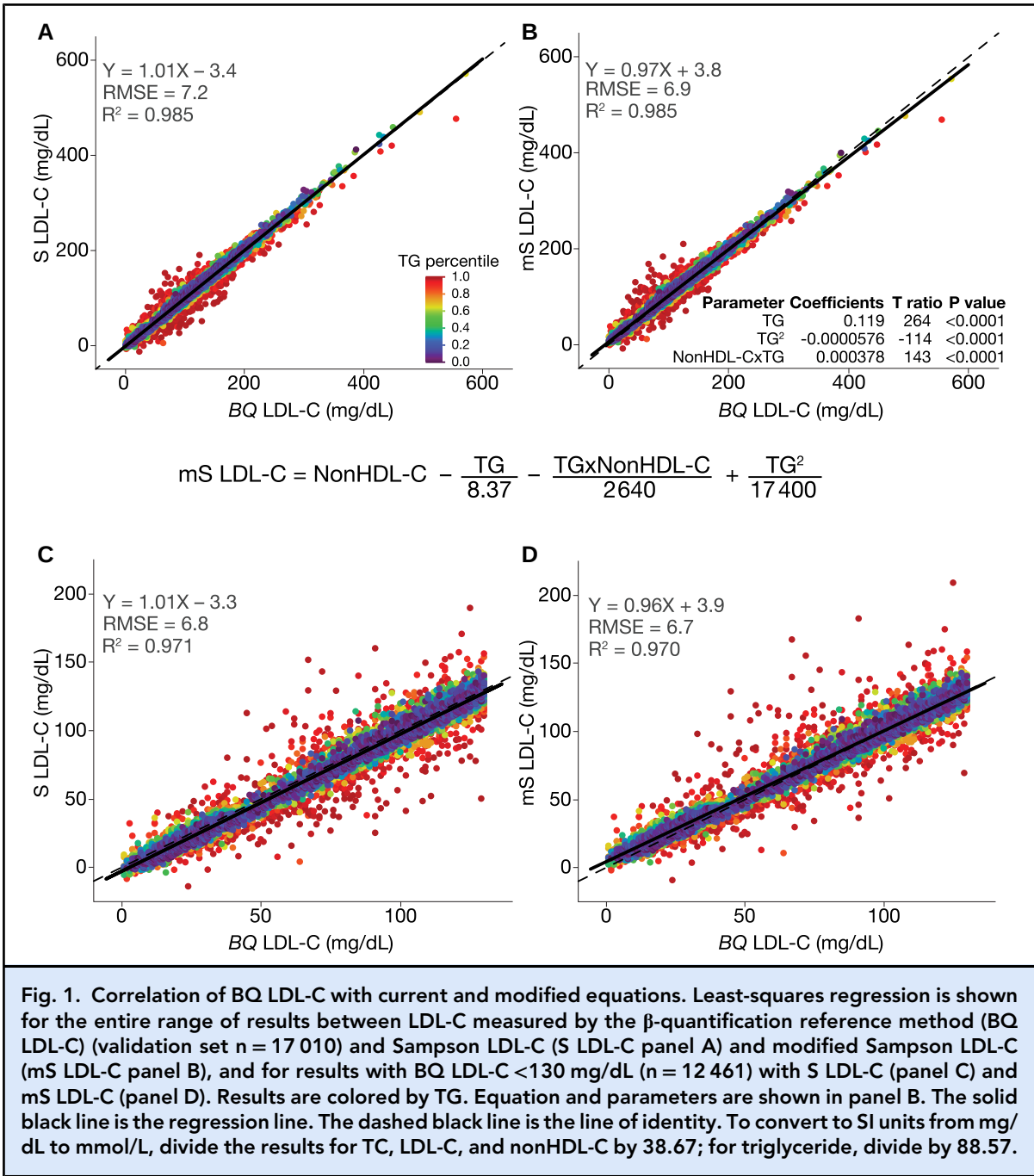
In the development of the modified equation, we first simplified the independent variables in the S-equation by eliminating the individual terms for TC and HDL-C, by combining them as nonHDL-C ($\text{nonHDL-C} = \text{TC} - \text{HDL-C}$) and forcing its coefficient to be one. We forced the intercept to zero. These changes were made to better align with the known distribution of cholesterol in the different major lipoprotein fractions. The other terms in the mS-equation (TG, $\text{TG} \times \text{nonHDL-C}$, and TG^2), as in the S-equation, estimate the amount of cholesterol in lipoprotein fractions besides LDL and lipoprotein (a) [Lp(a)], which includes VLDL, chylomicrons and their remnants, but will simply be referred to here as VLDL-C. The coefficients for these 3 terms in the mS-equation were fitted by least-squares regression analysis (Fig. 1B), using the training data set, and tested in the validation data set. Despite having the same R^2 as the S-equation, there was a small improvement in the root means square error score for the mS-equation (7.2 vs 6.9). It also had a statistically stronger fit to BQ LDL-C as determined by the F -test ($P < 0.001$). When only those samples with BQ LDL-C ≤ 130 mg/dL (3.36 mmol/L) ($n = 12\,461$) were analyzed (Fig. 1C and D), the mS-equation showed a slightly

stronger fit to the BQ method than did the S-equation. The inclusion of higher-order mathematical terms into the mS-equation, containing various combinations of the independent variables, did not further improve its fit.

We next examined the cumulative error distribution for the equations by REC plot analysis. A log scale for the x -axis was used to more closely examine small differences between the original and modified equations (Fig. 2A). At error limits of 15% or less, the mS-equation produced LDL-C results that more closely matched BQ LDL-C than the original S-equation. An even greater improvement for the mS-equation was observed in the REC plots when we restricted the analysis to $\text{LDL} \leq 130$ mg/dL (3.36 mmol/L) (Fig. 2B) or to $\text{TG} \geq 400$ mg/dL (4.5 mmol/L) (Fig. 2C). A direct comparison of the two equations (Fig. 2D) showed a very close correlation, but the mS-equation tended to result in higher LDL-C values for hypertriglyceridemic samples. The mS-equation also showed a statistically higher Lin's concordance correlation coefficient (CCC) to BQ LDL-C than the S-equation or the other commonly used LDL-C equations (online Supplemental Table 2).

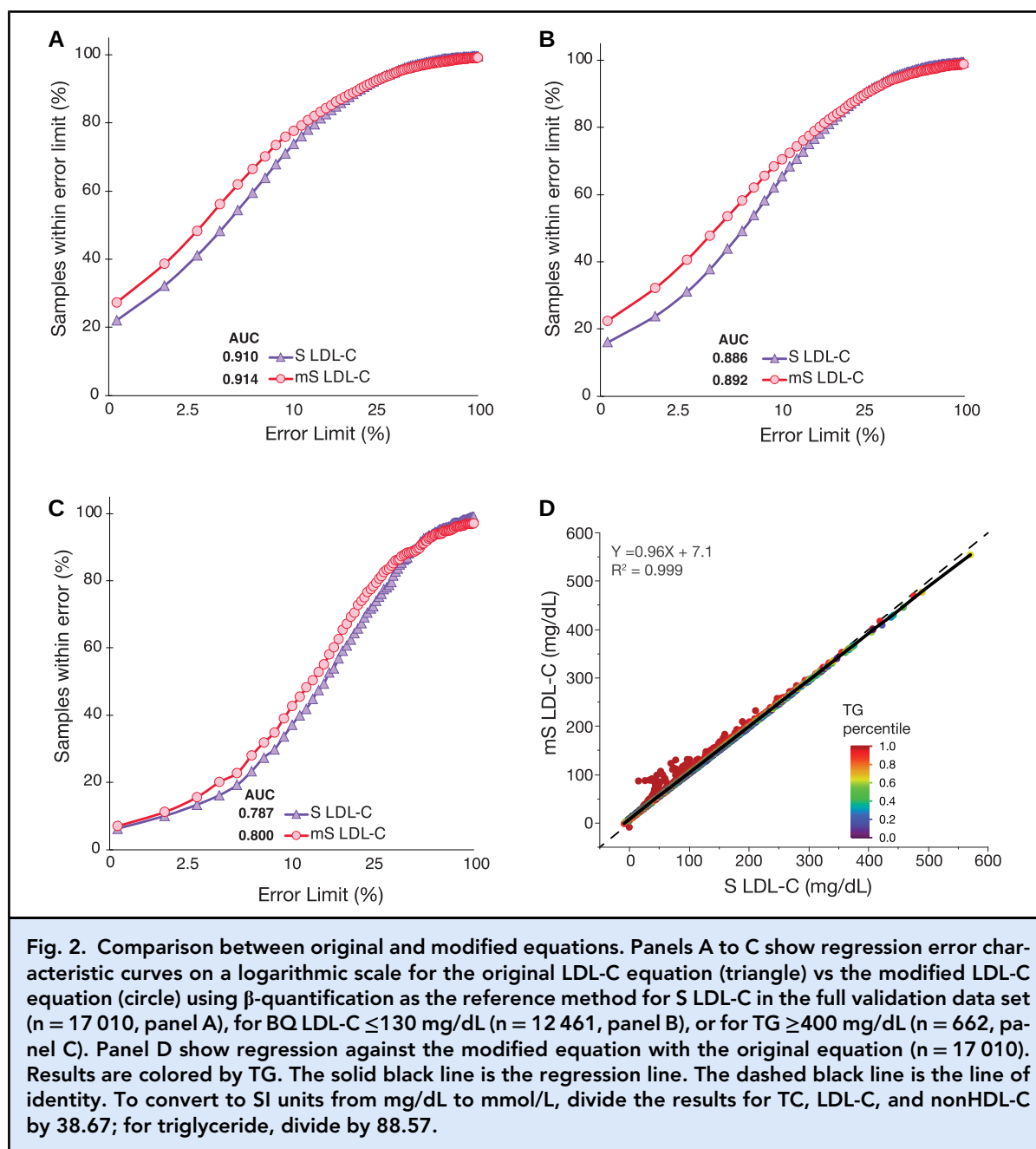
We next analyzed the mean bias of the various equations for increasing intervals of TG (Table 1). At the 45 to 65 mg/dL (1.16 to 1.68 mmol/L) interval, which encompasses the 55 mg/dL (1.4 mmol/L) cutpoint, the mS-equation showed a nonsignificant difference from the BQ method (0.01 mg/dL [0.0002 mmol/L]; 95% CI, -0.50 – 0.53 mg/dL, $P = 0.9623$), even after adjusting across TG levels (bias from -0.36 to 1.2 mg/dL, $P = 0.3507$). At this same interval, the M-equation also had a nonsignificant bias from the BQ method (-0.62 mg/dL [0.016 mmol/L]; 95% CI, -1.29 – 0.04 mg/dL, $P = 0.0668$) (Table 1). However, this was a result of a negative bias at low TG and positive bias at high TG (-3.8 to 17.3 mg/dL), which cancelled each other out. After adjustment by TG, the M-equation did show a significant difference from the BQ method ($P < 0.0001$). At the 180 to 200 mg/dL (4.65 to 5.17 mmol/L) interval, all the equations showed a similar bias (-1.27 to -3.7 mg/dL [-0.03 to 0.1 mmol/L]) (Table 1).

In Fig. 3A, we plotted the mean LDL-C bias for the mS-equation. It was approximately zero across almost the entire TG range. The mean bias line for the original S-equation was also relatively flat and slightly below zero, but it did start to show a small negative bias for TG intervals greater than 400 mg/dL (4.5 mmol/L). In contrast, the F-equation showed a strong mean negative bias with increasing TG intervals. The M-equation started out with a small negative bias for low TG intervals but a strong positive bias for high TG intervals. Because mean bias is not affected by the variance of the estimate, we also calculated the mean absolute difference (MAD) from the BQ method (Fig. 3B). All the equations showed an increase in their MAD scores



with increasing TG intervals, but the mS-equation was the least adversely affected. This occurs because hypertriglyceridemia increases the variance of the predicted LDL-C results, which can be seen in the residual error plots (online Supplemental Fig. 1). For nonHDL-C intervals, the mS-equation also had a mean bias close to zero, except for the highest and lowest nonHDL-C intervals (Fig. 3C). All other equations started out with

a negative mean LDL-C bias (3 to 7 mg/dL [0.077 to 0.18 mmol/L]) at low nonHDL-C intervals, which then decreased as nonHDL-C increased. In the case of the M-equation, it turned into a positive mean bias. Based on MAD scores, the mS-equation showed the smallest decrease in its accuracy as nonHDL-C increased, but the MAD scores for all the equations started to converge at higher nonHDL-C intervals (Fig. 3D).



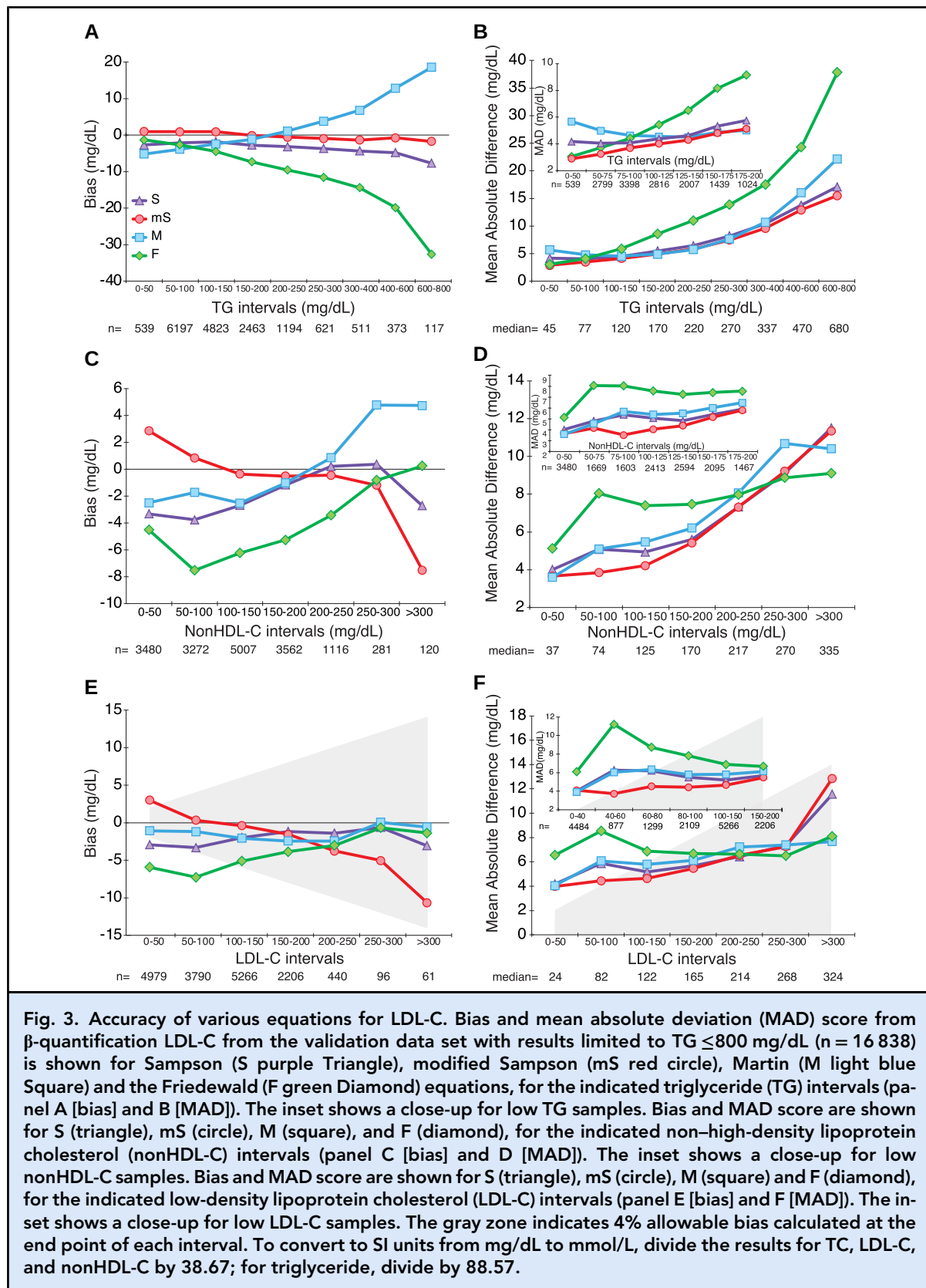
We also calculated the mean LDL-C bias and MAD scores by LDL-C intervals and included the recommended bias limit of 4% (27) for LDL-C (gray zone) in these graphs. For LDL-C intervals between 50 and 150 mg/dL (1.29 and 3.87 mmol/L), the mS-equation was the most accurate (Fig. 3E). It had a mean bias of only -0.13 mg/dL (0.003 mmol/L) (95% CI, -0.26 – 0.01 mg/dL, P value = 0.0713), which was not statistically different from BQ LDL-C (Table 1). Based on MAD scores (Fig. 3F), the mS-equation also showed

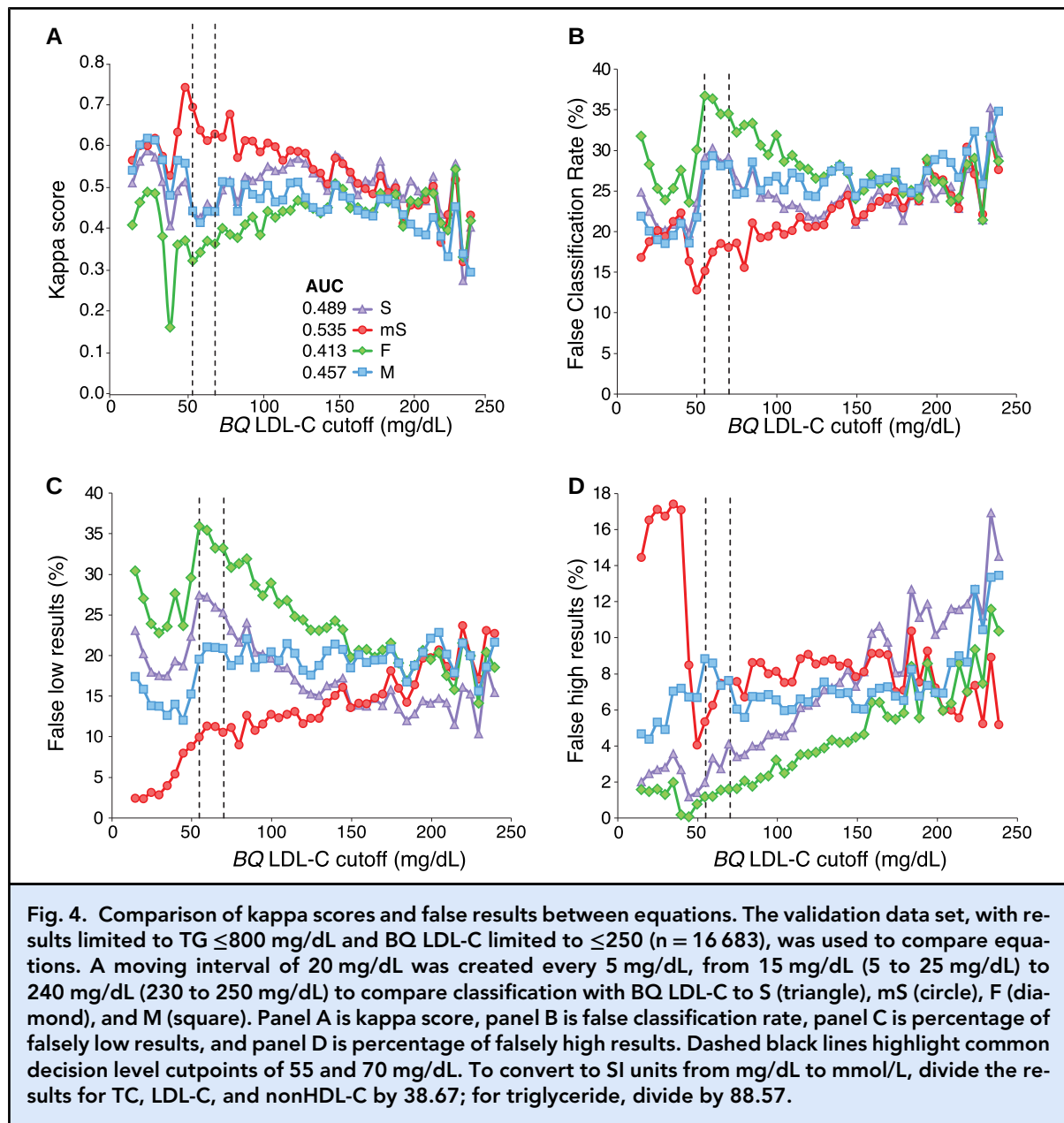
the best accuracy of 4.83 mg/dL (95% CI 4.66–4.99 mg/dL, P value <0.0001), but for LDL-C values below 40 mg/dL (1.03 mmol/L) it exceeded the recommended 4% (1.6 mg/dL [0.04 mmol/L]) bias limit (gray zone). The 4% bias limit, however, becomes an unrealistic target and less clinically relevant at low LDL-C values. The other LDL-C equations showed an even larger mean bias and MAD scores at these low LDL-C values, particularly the F-equation. At the highest LDL-C interval (>300 mg/dL [7.75 mmol/L]), the mS-equation did

Table 1. Comparative analysis of equation bias across LDL-C and triglyceride intervals.^a

Intervals BQ LDL-C, mg/dL	Equation	Mean bias, mg/ dL	Full TG interval (0–800 mg/dL)				TG intervals, mg/dL				Mean bias by pairs		Mean bias across TG groups	
			Upper 95% CI	Lower 95% CI	t value	P value	0–200 Mean bias	200–400 Mean bias	400–800 Mean bias	F value	P value	F value	P value	
1–572 (full range)	S	–2.47	–2.37	–2.57	–48.6	<0.0001	–2.17	–3.61	–5.50	102	<0.0001	60.4	<0.0001	
	mS	0.43	0.53	0.33	8.55	<0.0001	0.70	–0.86	–1.03	69.2	<0.0001	63.5	<0.0001	
	M	–1.65	–1.54	–1.76	–29.2	<0.0001	–2.99	3.05	14.2	2361	<0.0001	113	<0.0001	
	F	–5.60	–5.48	–5.73	–85.8	<0.0001	–4.07	–11.2	–22.9	2217	<0.0001	35.9	<0.0001	
45–65	S	–4.80	–4.27	–5.32	–17.9	<0.0001	–4.83	–5.01	–4.04	0.39	0.6797	1.55	0.2139	
	mS	0.01	0.53	–0.50	0.05	0.9623	–0.03	–0.36	1.21	1.05	0.3507	1.82	0.1626	
	M	–0.62	0.04	–1.29	–1.84	0.0668	–3.87	4.32	17.3	278	<0.0001	61.2	<0.0001	
	F	–10.5	–9.68	–11.4	–24.4	<0.0001	–6.36	–16.8	–33.8	295	<0.0001	121	<0.0001	
60–80	S	–3.72	–3.32	–4.13	–18.1	<0.0001	–4.05	–2.80	–1.08	6.70	0.0013	2.29	0.1018	
	mS	0.378	0.770	–0.0133	1.90	0.0583	0.00	1.30	3.94	11.2	<0.0001	2.29	0.1017	
	M	–1.22	–0.72	–1.72	–4.83	<0.0001	–3.59	5.22	18.8	380	<0.0001	94.7	<0.0001	
	F	–7.55	–7.01	–8.09	–27.5	<0.0001	–5.48	–13.0	–25.8	217	<0.0001	97.5	<0.0001	
50–150	S	–2.58	–2.43	–2.72	–35.4	<0.0001	–2.38	–3.60	–2.95	18.3	<0.0001	27.0	<0.0001	
	mS	–0.126	0.0110	–0.263	–1.80	0.0713	–0.0950	–0.740	1.66	17.8	<0.0001	22.3	<0.0001	
	M	–1.74	–1.58	–1.90	–21.0	<0.0001	–3.41	3.28	15.2	1623	<0.0001	8.39	0.0002	
	F	–6.03	–5.85	–6.21	–65.0	<0.0001	–4.37	–11.4	–21.5	1099	<0.0001	90.8	<0.0001	
180–200	S	–1.27	–0.544	–2.00	–3.43	0.0006	0.0572	–3.93	–15.7	40.2	<0.0001	20.0	<0.0001	
	mS	–2.48	–1.81	–3.16	–7.25	<0.0001	–1.64	–4.04	–12.5	19.6	<0.0001	11.6	<0.0001	
	M	–2.63	–1.92	–3.33	–7.27	<0.0001	–3.66	0.22	3.57	15.1	<0.0001	1.93	0.1464	
	F	–3.70	–2.95	–4.45	–9.73	<0.0001	–2.60	–5.86	–16.0	25.6	<0.0001	14.6	<0.0001	

Abbreviations: BQ LDL-C, β -quantification low density lipoprotein cholesterol; LDL-C, low density lipoprotein cholesterol; TG, triglycerides; S, Sampson LDL-C; mS, modified Sampson LDL-C; M, Martin LDL-C; F, Friedewald LDL-C.
*Bolded P values indicate no significant difference from BQ LDL-C. F and t values show magnitude of effect. To convert to SI units from mg/dL to mmol/L, divide the results for TC, LDL-C, nonHDL-C by 38.67; for triglyceride divide by 88.57.





show a higher mean (negative) bias and higher MAD score than the other equations.

Because LDL-C is used in a dichotomous manner in making clinical decisions, we compared the kappa scores for the various equations for their concordance with BQ LDL-C at multiple cutpoints, each with a ± 10 mg/dL (0.25 mmol/L) interval around it (Fig. 4A). Using area under the curve (AUC) as a metric, the best equation across all LDL-C intervals was the mS-equation (AUC: 0.535). As an alternative to the kappa score, we also examined the false classification rate (FCR) when compared to BQ LDL-C (Fig. 4B),

which was also categorized as either being a falsely low (Fig. 4C) or falsely high (Fig. 4D). The relative accuracy of the various equations by this approach is the same as for kappa analysis, but it provides additional information, such as the exact frequency of errors and whether they are due to a positive or a negative bias. The FCR of mS-equation at the 55 mg/dL (1.4 mmol/L) cutpoint was significantly lower (15%, $P < 0.0001$) compared to the other equations (S: 29%, M: 28%, F: 37%). Based on the net reclassification index (NRI), the mS-equation increased the correct classification of patients at this cutpoint compared to all other equations

(F: 21.5%, S: 14.0%, M: 12.9%). A similar significant reduction in FCR for the mS-equation at the 70 mg/dL (1.8 mmol/L) cutpoint (18%, $P < 0.0001$) was also observed compared to other equations (S: 30%, M: 28%, F: 34%). Again, based on NRI, the mS-equation increased the correct classification of patients at this cutpoint compared to other equations (F: 16.3%, S: 11.4%, M: 10.1%). Furthermore, the mS-equation showed, by the McNemar test, the closest agreement to the BQ method at the 55 mg/dL (1.4 mmol/L) cutpoint ($\chi^2 = \text{mS: } 12.1, \text{ M: } 35.3, \text{ S: } 190, \text{ F: } 281$) and the 70 mg/dL (1.8 mmol/L) cutpoint ($\chi^2 = \text{mS: } 7.7, \text{ M: } 86.1, \text{ S: } 400, \text{ F: } 221$).

Next, we divided the validation dataset into 4 quadrants based on the median value of TG and nonHDL-C, the two variables that have the largest impact in estimating LDL-C. The FCR by the various equations was then determined. In Fig. 5A (high nonHDL-C and low TG), the M-equation had the greatest FCR followed by the F-equation and then the other equations. In contrast, in Fig. 5B (high nonHDL-C and high TG), the F-equation clearly stood out in terms of its high level of inaccuracy for low LDL-C samples. In Fig. 5C (low nonHDL-C and low TG), all the equations showed a similar FCR except for the mS-equation, which was more accurate than the other equations. As shown in Fig. 5D (low nonHDL-C and high TG), the F-equation was once again the least accurate and the mS-equation was overall the most accurate.

Finally, we also compared the mS-equation to the eS-equation, which requires apoB. The results are shown in online Supplemental Figs. 2–4. The mS-equation was nearly as accurate as the eS-equation and showed a similar mean bias and MAD score as the eS-equation across TG, LDL-C, and non-HDL-C intervals. The two equations were also similar in their accuracy by kappa analysis in the classification of patients into different LDL-C intervals. In Supplemental Fig. 2, we show TG values up to 3000 mg/dL (33.9 mmol/L). At TG levels up to 1000 mg/dL (11.3 mmol/L), the mS-equation results in errors below the maximum recommended limit and thus has a slightly extended TG range compared to the 800 mg/dL (9 mmol/L) limit for the original S-equation.

Discussion

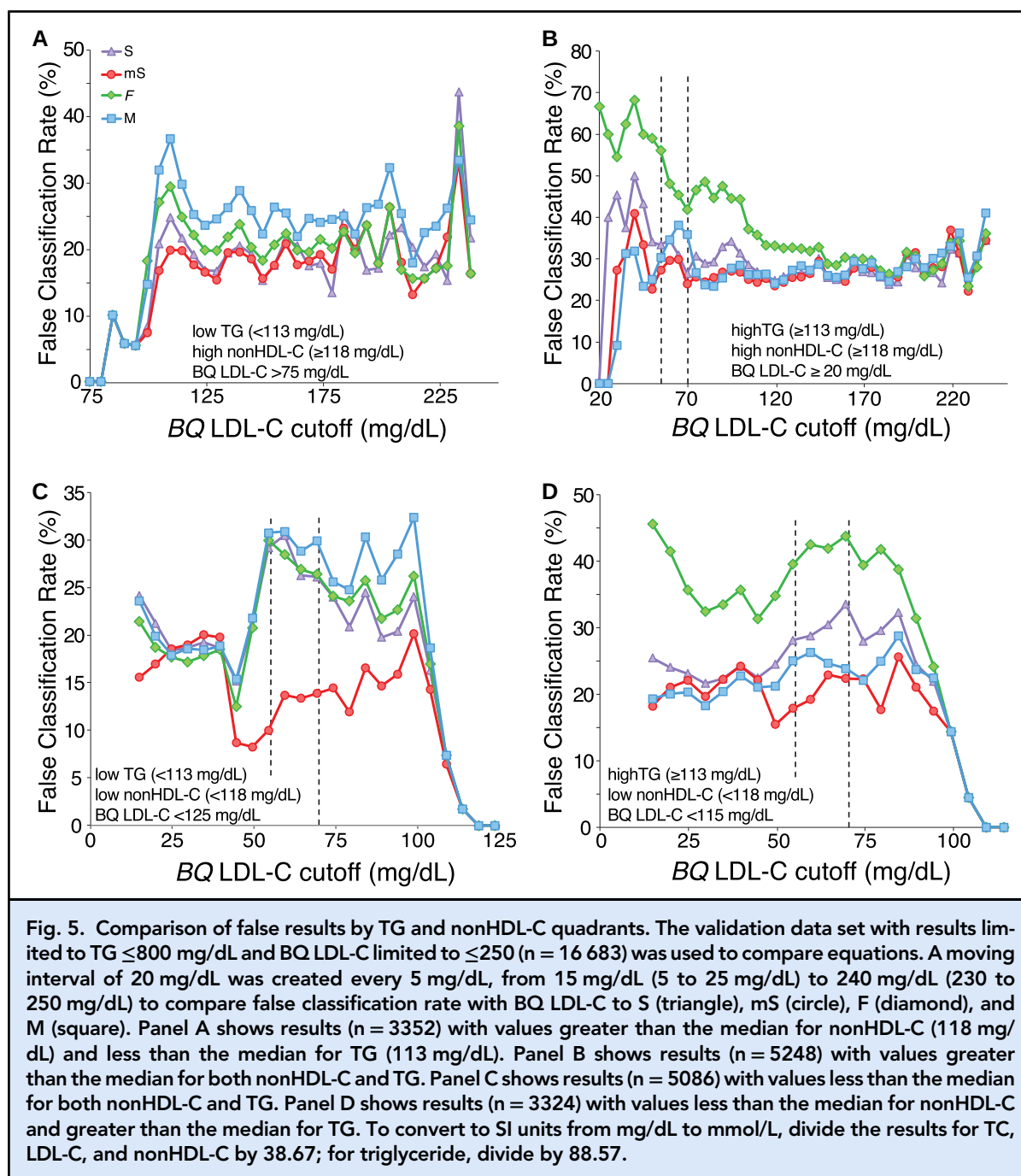
In this study, we describe the development of a more accurate LDL-C equation that could improve clinical decision-making at low LDL-C thresholds. By using an additional cohort of very low LDL-C values from patients treated with a PCSK9i drug, we were able to significantly improve the accuracy of estimated LDL-C by least-squares regression analysis. Use of the new mS-equation should better identify patients with low

LDL-C, who are not at their LDL-C target goals and could benefit from intensification of lipid-lowering therapy.

Based on recent College of American Pathology Proficiency test surveys, the great majority of clinical laboratories still use the F-equation despite its well-known limitations (8). When it was first developed over 50 years ago (11), the accurate measurement of low LDL-C was not clinically relevant given the lack of effective lipid-lowering drugs at the time. In 2018, when the latest comprehensive US cardiovascular guidelines were completed, it was recommended that either a direct assay or the M-equation be used instead of the F-equation for low LDL-C (<70 mg/dL, [1.8 mmol/L]) and high TG (>150 mg/dL, [1.7 mmol/L]) (1). We previously found that as many as a third of all individuals with an LDL-C <70 mg/dL by the F-equation are, in fact, falsely low by as much as 30 mg/dL (0.77 mmol/L) (20). This may result in patients not receiving PCSK9i therapy or may make it more difficult to get reimbursement for the new more effective but also more expensive lipid-lowering treatments (7). In our view, future cardiovascular guidelines should actively discourage the use of the F-equation for all purposes. It does not appear to offer any accuracy advantage over the newer alternative equations and may result in the undertreatment of patients, because of its negative bias.

The original S-equation, which was developed in 2020 (14), offers the following advantages over the M-equation: (a) it is based on the BQ reference method unlike the M-equation, (b) it is more accurate for high TG samples (20), and (c) it can be easily implemented into almost all laboratory information systems without the need of additional IT support unlike the M-equation (8). As can be seen in this and other studies, the two equations do yield similar results on most samples. Some reports have favored the M-equation, because it more closely compares with some direct LDL-C assays (28, 29). However, direct LDL-C assays are known to sometimes have a positive bias on high TG samples (9, 30), which is what has been observed for the M-equation when compared against the BQ method (14, 31).

The new mS-equation has the same advantages over the M-equation as the original S-equation, but it is even more accurate, especially for low LDL-C values. Its accuracy almost approaches the eS-equation (19), which includes apoB in the prediction. By retraining the new equation with a data set containing a greater number of low LDL-C values than was used for the original S-equation (14), the coefficients for the 3 terms related to VLDL-C differ from the original equation, which accounts for its improved accuracy. Based on its improved accuracy, especially at key clinical decision thresholds of 55 and 70 mg/dL (1.4 and 1.8 mmol/L), we recommend that the mS-equation be used to replace the



S-equation, as well as the other equations. The mS-equation did, however, show a larger mean bias and higher MAD scores at high LDL-C values greater than 300 mg/dL (7.75 mmol/L), but LDL-C values this high are quite rare (>99.9th percentile in the US National Health and Nutrition Survey general population survey) and do not contain any specific clinical thresholds. Importantly, at the 190 mg/dL (4.9 mmol/L) cutpoint, a critical decision threshold for initiating

statin therapy for primary prevention (1), the accuracy of the mS-equation was similar to the other equations.

In 2014, the Milan consensus conference on diagnostic testing recommended that direct outcome-based models, such as clinical outcomes, or indirect outcome models (impact on clinical decisions) as the preferred approaches for assessing diagnostic test performance (32). Although regression-related parameters are useful for assessing the predictive value of a linear regression equation,

they are only global measures of the overall fit and do not directly relate to clinical decision-making at a decision threshold. In addition, many of these types of metrics are known to be unduly affected by outliers (33). We found in this study that REC curve analysis (24) can provide a more detailed view of the performance of an LDL-C equation. A reduction in relatively small errors of less than 15%, which can be readily observed by REC curves when a log scale is used for the x -axis, can still have a large impact on the correct binary classification around a single decision threshold. The advantage of using kappa scores and FCR as performance metrics is that they both provide a measure for how inaccurate LDL-C measurements can affect clinical decisions. FCR, which can be evaluated as to whether it is due to falsely high or low results, is perhaps even more intuitive than the kappa score, which does not have units. The FCR of the mS-equation was reduced by about half compared to the other equations at both the 55 and 70 mg/dL (1.4 and 1.8 mmol/L) cutpoints. Based on NRI, the use of the mS-equation would increase the correct classification of patients at these low LDL-C cutpoints by approximately 10% to 20% compared to the other equations.

This study has several limitations. First, methodological differences in the BQ LDL-C assay across data sets may introduce variability, and the BQ reference method sometimes lacks specificity for LDL-C, potentially also capturing cholesterol from small dense remnants and Lp(a). Second, treatment status was unavailable for the Mayo Clinic cohort; however, the new equation matches well higher LDL-C values by the original S-LDL-C equation, which was largely based on untreated individuals (14). Lastly, individuals with dysbetalipoproteinemia were excluded; however, this condition is rare (<1%), and the eS-equation (19) was specifically designed for this subgroup.

In summary, we describe a simple modification of the Sampson–NIH equation with improved test accuracy, particularly for low LDL-C values. Use of this new equation should improve the identification of high-risk and very high-risk patients, who are not at their LDL-C goals and require additional lipid-lowering therapy. LDL-C, however, has inherent limitations as an ASCVD biomarker (34, 35) and other measures like apoB are now being recommended (36, 37) to guide lipid-lowering therapy. Even cholesterol in subfractions of LDL, which can also be calculated from the lipid panel (38, 39), is being investigated as an alternative marker, but until a transition is made in the preferred ASCVD biomarker, LDL-C should be calculated by the most accurate method available. It is also important to note the need to assess the new equation in additional cohorts even though we used several different large data sets in our study. Because the new equation only utilizes results from the standard lipid panel, it will make it relatively

easy to perform additional studies to determine the generalizability of the current findings. Importantly, this feature of the mS-equation would also facilitate its possible future implementation into routine clinical practice.

Supplemental Material

Supplemental material is available at *Clinical Chemistry* online.

Nonstandard Abbreviations: LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein-cholesterol; ASCVD, atherosclerotic cardiovascular disease; PCSK9i, proprotein convertase subtilisin-kexin type 9; TC, total cholesterol; TG, triglycerides; F, Friedewald; VLDL-C, very low-density lipoprotein cholesterol; M, Martin–Hopkins; S, Sampson–NIH; BQ, β -quantification; eS, enhanced Sampson; apoB, apolipoprotein B; mS, modified Sampson equation; REC, regression error characteristic; MAD, mean absolute difference; FCR, false classification rate.

Author Contributions: *The corresponding author takes full responsibility that all authors on this publication have met the following required criteria of eligibility for authorship: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; (c) final approval of the published article; and (d) agreement to be accountable for all aspects of the article thus ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved. Nobody who qualifies for authorship has been omitted from the list.*

Maureen Sampson (Conceptualization–Equal, Data curation–Equal, Formal analysis–Equal, Methodology–Equal, Validation–Equal, Visualization–Equal, Writing—original draft–Equal, Writing—review & editing–Equal), Rafael Zubiran (Conceptualization–Equal, Data curation–Equal, Formal analysis–Equal, Methodology–Equal, Project administration–Equal, Supervision–Equal, Validation–Equal, Visualization–Equal, Writing—original draft–Equal, Writing—review & editing–Equal), Anna Wolska (Formal analysis–Supporting, Methodology–Equal, Project administration–Equal, Writing—original draft–Supporting, Writing—review & editing–Supporting), Jeffrey Meeusen (Methodology–Equal, Writing—review & editing–Supporting), Leslie Donato (Methodology–Equal, Writing—review & editing–Supporting), Allan Jaffe (Methodology–Equal, Writing—review & editing–Supporting), Giorgio E.M. Melloni (Methodology–Equal, Writing—review & editing–Supporting), Robert P. Giugliano (Methodology–Equal, Writing—review & editing–Supporting), Marc S. Sabatine (Methodology–Equal, Writing—review & editing–Supporting), Nicholas A. Marston (Methodology–Equal, Writing—review & editing–Supporting), and Alan T. Remaley (Conceptualization–Lead, Data curation–Equal, Formal analysis–Equal, Funding acquisition–Equal, Investigation–Equal, Methodology–Equal, Project administration–Lead, Resources–Equal, Supervision–Lead, Visualization–Equal, Writing—original draft–Equal, Writing—review & editing–Equal)

Authors’ Disclosures or Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the author disclosure form.*

Research Funding: This work was supported by the Intramural Research Program of the National Heart, Lung, and Blood Institute (NHLBI) at the National Institutes of Health (to R. Zubiran,

A. Wolska, and A.T. Remaley) and a grant (23CVD02) from the Leducq Foundation for Cardiovascular Research (to A.T. Remaley).

Disclosures: A.S. Jaffe currently consults, or has consulted in the past, for Abbott, Siemens, Roche, Radiometer, ET Healthcare, 4Tean4, SpinChip, LuminaDx, Hytest, and Mindray, and holds stock options in RCE Technologies. M.S. Sabatine received research grant support through Brigham and Women's Hospital from Abbott, Anthos Therapeutics, Inc., AstraZeneca, Boehringer Ingelheim, Daiichi-Sankyo, Ionis, Marea, Merck, Novartis, Pfizer, Sagham Therapeutics, and Verve Therapeutics; consulted for Amgen, AMPEL BioSolutions, Anthos Therapeutics, Inc., AstraZeneca, Beren Therapeutics, Boehringer Ingelheim, Dr. Reddy's Laboratories, Fibrogen, Merck, Moderna, Novo Nordisk, Precision BioSciences, and Silence Therapeutics. R.P. Gugliano received research grants from: Amgen, Anthos Therapeutics, Daiichi Sankyo and Ionis CV; honoraria for lectures from: Amgen, CADECI,

Centrix, Daiichi Sankyo, Dr. Reddy's Laboratories, Korean Heart Rhythm Society, Medical Education Resources (MER), Menarini, Pfizer, SHAKEHEART, and SUMMEET; and acted as consultant for Amgen, Astra Zeneca, Beckman Coulter, Daiichi Sankyo, Gilead, Inventiva, Novartis, Persosphere, Samsung, and Syneos Health. N.A. Marston received research grants from National Institutes of Health and Amgen, Ionis, Pfizer, Novartis, and AstraZeneca; honoraria for lectures from: Amgen, Illumina, and Medical Education Speakers Network (MESN); and acted as consultant for Beckman Coulter, Cleerly, Inc. (CPC Clinical Research), Janssen, and Viz.ai. L.J. Donato received speaker support from Helena and Novartis and acted as a consultant for Bosch Medical.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, preparation of manuscript, or final approval of manuscript.

References

1. Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APHA/ASPC/NLA/PCNA guideline on the management of blood cholesterol. *Circulation* 2019;139:e1082–143.
2. Rao SV, O'Donoghue ML, Ruel M, Rab T, Tamis-Holland JE, Alexander JH, et al. 2025 ACC/AHA/ACEP/NAEMSP/SCAI guideline for the management of patients with acute coronary syndromes: a report of the American College of Cardiology/American Heart Association Joint Committee on clinical practice guidelines. *Circulation* 2025;151:e771–862.
3. Visseren FLJ, Mach F, Smulders YM, Carballo D, Koskinas KC, Bäck M, et al. 2021 ESC guidelines on cardiovascular disease prevention in clinical practice. *Eur Heart J* 2021;42:3227–337.
4. Schwartz GG, Steg PG, Szarek M, Bhatt DL, Bittner VA, Diaz R, et al. Alirocumab and cardiovascular outcomes after acute coronary syndrome. *N Engl J Med* 2018;379:2097–107.
5. Rodriguez F, Khera A. How low can you go? New evidence supports no lower bound to low-density lipoprotein cholesterol level in secondary prevention. *Circulation* 2023;147:1204–7.
6. Sabatine MS, Giugliano RP, Keech AC, Honarpour N, Wiviott SD, Murphy SA, et al. Evolocumab and clinical outcomes in patients with cardiovascular disease. *N Engl J Med* 2017;376:1713–22.
7. Cohen JD, Cziraky MJ, Jacobson TA, Maki KC, Karalis DG. Barriers to PCSK9 inhibitor prescriptions for patients with high cardiovascular risk: results of a healthcare provider survey conducted by the National Lipid Association. *J Clin Lipidol* 2017;11:891–900.
8. Wolska A, Remaley AT. Measuring LDL-cholesterol: what is the best way to do it? *Curr Opin Cardiol* 2020;35:405–11.
9. Miller WG, Myers GL, Sakurabayashi I, Bachmann LM, Caudill SP, Dziekonski A, et al. Seven direct methods for measuring HDL and LDL cholesterol compared with ultracentrifugation reference measurement procedures. *Clin Chem* 2010;56:977–86.
10. Miida T, Nishimura K, Okamura T, Hirayama S, Ohmura H, Yoshida H, et al. A multicenter study on the precision and accuracy of homogeneous assays for LDL-cholesterol: comparison with a beta-quantification method using fresh serum obtained from non-diseased and diseased subjects. *Atherosclerosis* 2012;225:208–15.
11. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* 1972;18:499–502.
12. Martin SS, Blaha MJ, Elshazly MB, Toth PP, Kwiterovich PO, Blumenthal RS, Jones SR. Comparison of a novel method vs the friedewald equation for estimating low-density lipoprotein cholesterol levels from the standard lipid profile. *JAMA* 2013;310:2061–8.
13. Chung BH, Segrest JP, Ray MJ, Brunzell JD, Hokanson JE, Krauss RM, et al. Single vertical spin density gradient ultracentrifugation. *Methods Enzymol* 1986;128:181–209.
14. Sampson M, Ling C, Sun Q, Harb R, Ashmaig M, Warnick R, et al. A new equation for calculation of low-density lipoprotein cholesterol in patients with normolipidemia and/or hypertriglyceridemia. *JAMA Cardiol* 2020;5:540–8.
15. McNamara JR, Leary ET, Ceriotti F, Boersma-Cobbaert CM, Cole TG, Hassemer DJ, et al. Point: status of lipid and lipoprotein standardization. *Clin Chem* 1997;43:1306–10.
16. Hainline A Jr. Manual of laboratory operations: lipid and lipoprotein analysis. DHEW Pub. No.(NIH) 75-628 (rev.). Publication No. 1982-361-132:678. Washington (DC): US Government Printing Office; 1982.
17. Hopkins PN, Brinton EA, Nanjee MN. Hyperlipoproteinemia type 3: the forgotten phenotype. *Curr Atheroscler Rep* 2014;16:440.
18. Zubirán R, Vargas-Vazquez A, Olvera FDR, Cruz-Bautista I, Martagón-Rosado A, Sampson M, et al. Performance of the enhanced Sampson-NIH equation for VLDL-C and LDL-C in a population with familial combined hyperlipidemia. *Atherosclerosis* 2023;386:117364.
19. Coverdell TC, Sampson M, Zubirán R, Wolska A, Donato LJ, Meeusen JW, et al. An improved method for estimating low LDL-C based on the enhanced Sampson-NIH equation. *Lipids Health Dis* 2024;23:43.
20. Sampson M, Wolska A, Cole J, Zubirán R, Otvos JD, Meeusen JW, et al. Accuracy and clinical impact of estimating low-density lipoprotein-cholesterol at high and low levels by different equations. *Biomedicines* 2022;10:3156.
21. Sampson M, Wolska A, Meeusen JW, Otvos J, Remaley AT. The Sampson-NIH equation is the preferred calculation method for LDL-C. *Clin Chem* 2024;70:399–402.
22. Gaba P, O'Donoghue ML, Park JG, Wiviott SD, Atar D, Kuder JF, et al. Association between achieved low-density lipoprotein cholesterol levels and long-term cardiovascular and safety outcomes: an analysis of FOURIER-OLE. *Circulation* 2023;147:1192–203.
23. Meeusen JW, Lueke AJ, Jaffe AS, Saenger AK. Validation of a proposed novel equation for estimating LDL cholesterol. *Clin Chem* 2014;60:1519–23.
24. Aşıkil B, Erar A. Regression error characteristic curves based on the choice of best estimation method. *Selçuk J Appl Math* 2013;14:1–9.
25. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37–46.
26. Sajja A, Park J, Sathiyakumar V, Varghese B, Pallazola VA, Marvel FA, et al. Comparison of methods to estimate low-density lipoprotein cholesterol in patients with high triglyceride levels. *JAMA Netw Open* 2021;4:e2128817.

27. Bachorik PS, Ross JW. National cholesterol education program recommendations for measurement of low-density lipoprotein cholesterol: executive summary. The national cholesterol education program working group on lipoprotein measurement. *Clin Chem* 1995;41:1414–20.
28. Zararsız G E, Bolat S, Cephe A, Kochan N, Yerlitaş Sİ, Doğan HO, Zararsız G. Validation of Friedewald, Martin-Hopkins and Sampson low-density lipoprotein cholesterol equations. *PLoS One* 2022;17:e0263860.
29. Grant JK, Kaufman HW, Martin SS. Extensive evidence supports the Martin-Hopkins equation as the LDL-C calculation of choice. *Clin Chem* 2024;70:392–8.
30. Langlois MR, Descamps OS, van der Laarse A, Weykamp C, Baum H, Pulkki K, et al. Clinical impact of direct HDLc and LDLc method bias in hypertriglyceridemia. A simulation study of the EAS-EFLM collaborative project group. *Atherosclerosis* 2014;233:83–90.
31. Ginsberg HN, Rosenson RS, Hovingh GK, Letierce A, Samuel R, Poulouin Y, Cannon CP. LDL-C calculated by Friedewald, Martin-Hopkins, or NIH equation 2 versus beta-quantification: pooled alirocumab trials. *J Lipid Res* 2022;63:100148.
32. Sandberg S, Fraser CG, Horvath AR, Jansen R, Jones G, Oosterhuis W, et al. Defining analytical performance specifications: consensus statement from the 1st strategic conference of the European Federation of Clinical Chemistry and Laboratory Medicine. *Clin Chem Lab Med* 2015;53:833–5.
33. Shipe ME, Deppen SA, Farjah F, Grogan EL. Developing prediction models for clinical use using logistic regression: an overview. *J Thorac Dis* 2019;11(Suppl 4):S574–84.
34. Sniderman AD, Navar AM, Thanassoulis G. Apolipoprotein B vs low-density lipoprotein cholesterol and non-high-density lipoprotein cholesterol as the primary measure of apolipoprotein B lipoprotein-related risk: the debate is over. *JAMA Cardiol* 2022;7:257–8.
35. Marston NA, Giugliano RP, Melloni GEM, Park JG, Morrill V, Blazing MA, et al. Association of apolipoprotein B-containing lipoproteins and risk of myocardial infarction in individuals with and without atherosclerosis: distinguishing between particle concentration, type, and content. *JAMA Cardiol* 2022;7:250–6.
36. Soffer DE, Marston NA, Maki KC, Jacobson TA, Bittner VA, Peña JM, et al. Role of apolipoprotein B in the clinical management of cardiovascular risk in adults: an expert clinical consensus from the National Lipid Association. *J Clin Lipidol* 2024;18:e647–63.
37. Wolska A, Lloyd-Jones DM, Remaley AT. Measure apolipoprotein B if we believe what we say about precision medicine. *Circulation* 2025;151:257–9.
38. Sampson M, Wolska A, Warnick R, LuceroD, Remaley AT. A new equation based on the standard lipid panel for calculating small dense low-density lipoprotein-cholesterol and its use as a risk-enhancer test. *Clin Chem* 2021;67:987–97.
39. Wolska A, Sampson M, Zubirán R, Meeusen JW, Donato LJ, Jaffe AS, Remaley AT. An equation for estimating low-density lipoprotein-triglyceride content and its use for cardiovascular disease risk stratification. *Front Cardiovasc Med* 2024;11:1452869.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/393520796>

A Modified Sampson–NIH Equation with Improved Accuracy for Estimating Low Levels of Low-Density Lipoprotein–Cholesterol

Article in *Clinical Chemistry* · July 2025

DOI: 10.1093/clinchem/hvaf073

CITATIONS

2

READS

137

11 authors, including:



Maureen L. Sampson

National Institutes of Health

165 PUBLICATIONS 3,620 CITATIONS

[SEE PROFILE](#)



Rafael Zubiran

National Institutes of Health

36 PUBLICATIONS 196 CITATIONS

[SEE PROFILE](#)



Anna Wolska

National Heart Lung and Blood Institute, National Institutes of Health

87 PUBLICATIONS 1,269 CITATIONS

[SEE PROFILE](#)



Jeffrey W. Meeusen

Mayo Foundation for Medical Education and Research

86 PUBLICATIONS 1,472 CITATIONS

[SEE PROFILE](#)