Atherosclerosis 404 (2025) 119178

Contents lists available at ScienceDirect

Atherosclerosis

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

Reliability of different formulas for estimating plasma Apolipoprotein B levels in a large cohort of South European individuals

Federica Fogacci [a,1], Serra Ilayda Yerlitas Tastan [a,b,c,1], Gozde Erturk Zararsiz [b,c], Halef Okan Dogan [d,e], Andrea Balbinot [a], Marina Giovannini [a], Claudio Borghi [a,e], Gokmen Zararsiz [b,c,f,2], Arrigo Francesco Giuseppe Cicero [a,e,2,*]

[a] Hypertension and Cardiovascular Risk Research Unit, Alma Mater Studiorum University of Bologna, Bologna, Italy
[b] Department of Biostatistics, Erciyes University School of Medicine, Kayseri, Türkiye
[c] Drug Application and Research Center (ERFARMA), Erciyes University, Kayseri, Türkiye
[d] Department of Biochemistry, Sivas Cumhuriyet University, School of Medicine, Sivas, Türkiye
[e] Cardiovascular Medicine Unit, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy
[f] Erciyes Teknopark, Hematainer Biotechnology and Health Products Inc, Kayseri, Türkiye

ARTICLE INFO

Keywords:
Apolipoprotein B
ApoB
Non-HDL-C
Comparison
Estimation
Formulas

ABSTRACT

Background and aims: Direct measurement of apolipoprotein B (ApoB) is not always standardized and is relatively expensive, making it unavailable in several low-income settings. To address this issue, several formulas have been developed to estimate ApoB levels. Therefore, our study aims to compare the reliability of 23 formulas for estimating ApoB levels in a large cohort of South-European individuals.

Methods: We retrospectively assessed 4.577 clinical records in which ApoB measurements were obtained using the same standardized method. Overall concordance was defined as the proportion of cases where the directly measured ApoB level fell within the same category as the estimated ApoB level, based on ApoB quartiles (<80 mg/dL, 80–94 mg/dL, 95–114 mg/dL, and ≥115 mg/dL). In addition, overall concordance was assessed for different lipoprotein(a) (Lp(a)) and non-high density lipoprotein cholesterol (non-HDL-C) sub-levels. Ordinary least squares linear regression analyses were performed to compare estimated and measured ApoB values. Residual error plots were generated to visualize the difference between each estimation method and the actual ApoB measurements, stratified by Lp(a) and non-HDL-C levels.

Results: Plasma ApoB levels were best predicted by a non-HDL-C based formula and a formula using Friedewald's low-density lipoprotein cholesterol (LDL-C), regardless of ApoB plasma levels. Non-HDL-C levels did not significantly affect the concordance between measured and estimated ApoB across the different formulas, except at low non-HDL-C levels. Similarly, Lp(a) levels did not significantly impact concordance. However, the highest concordance level was 41 %.

Conclusion: Some simple formulas based on low-cost and widely available parameters can estimate ApoB levels independently of ApoB, non-HDL-C, and Lp(a) plasma levels. This approach may be particularly useful for estimating ApoB levels in low-resource settings.

1. Introduction

Apolipoprotein B (ApoB) is a hydrophobic protein present on low-density lipoproteins (LDL), lipoprotein(a) (Lp(a)), and triglyceride (TG)-rich lipoproteins like chylomicrons, very low-density lipoproteins (VLDL), and intermediate-density lipoproteins (IDL) (also known as remnant lipoproteins) [1], thus representing the entire spectrum of atherogenic lipoproteins in the bloodstream [1].

The INTERHEART study, a large case-control investigation, has recently concluded that ApoB is a more accurate marker of cardiovascular risk than non-high-density lipoprotein cholesterol (non-HDL-C) [1,2]. Other analyses, including those from the National Health and

* Corresponding author. Hypertension and Cardiovascular Risk Research Unit, Alma Mater Studiorum University of Bologna, Bologna, Italy.
E-mail address: arrigo.cicero@unibo.it (A.F.G. Cicero).
[1] These authors are co-first authors.
[2] These authors are co-last authors.

https://doi.org/10.1016/j.atherosclerosis.2025.119178
Received 27 November 2024; Received in revised form 4 March 2025; Accepted 19 March 2025
Available online 28 March 2025

Table 1
Formula, drawbacks and limitations of ApoB estimation equations.

Model	Formula	Drawbacks and Limitations
Kulkarni [14]	$Apo\ B = 20.67944 + (0.551614 \times nonHDLc)$	This model was developed using non-HDL-C measured by Vertical Auto Profile (VAP)
Hermans Model-1 [15]	$Apo\ B = (0.60 \times nonHDLc) + 12.0$	This equation was developed using uncorrected linear regression
Hwang Model-1 [16]	$Apo\ B = (0.68 \times TC) - (0.62 \times HDLc) - (0.02 \times TG)$	This model was developed using the Friedewald equation
Hwang Model-2 [16]	$TG \leq 270$: $Apo\ B = (0.65 \times TC) - (0.59 \times HDLc) + (0.01 \times TG)$ $TG > 270$: $Apo\ B = 25.6 + (0.58 \times TC) - (0.38 \times HDLc) + (0.06 \times TG)$	This model was evaluated using pre-defined TG cutoff points
Cho Model-1 [17]	$Apo\ B = 16.177816 + (0.735235 \times FLDLc)$	In this model LDL-C was estimated using the Friedewald formula
Cho Model-2 [17]	$Apo\ B = 16.177816 + (0.735235 \times SLDLc)$	In this model LDL-C was estimated using the Sampson formula
Cho Model-3 [17]	$Apo\ B = 16.177816 + (0.735235 \times MLDLc)$	In this model LDL-C was estimated using the Martin-Hopkins formula
Cho Model-4 [17]	$Apo\ B = -24.77 + (0.72 \times LDLc) + (11.43 \times ln(age))$	In this model LDL-C was estimated using the Friedewald formula
Cho Model-5 [17]	$Apo\ B = -49.13 + (0.70 \times LDLc) + (21.81 \times ln(BMI))$	In this model LDL-C was estimated using the Friedewald formula
Cho Model-6 [17]	$Apo\ B = -78.91 + (0.696 \times LDLc) + (9.81 \times ln(Age)) + (19.98 \times ln(BMI))$	In this model LDL-C was estimated using the Friedewald formula
Cho Model-7 [17]	$Apo\ B = -59.40 + (0.67 \times LDLc) + (11.63 \times ln(tg)) + (7.68 \times ln(age))$	In this model LDL-C was estimated using the Friedewald formula
Cho Model-8 [17]	$Apo\ B = -35.99 + (0.67 \times LDLc) + (11.84 \times ln(tg)) + (1.11 \times ln(BMI))$	In this model LDL-C was estimated using the Friedewald formula
Cho Model-9 [17]	$Apo\ B = -59.53 + (0.67 \times LDLc) + (11.62 \times ln(tg)) + (7.68 \times ln(age)) + (0.05 \times ln(BMI))$	In this model LDL-C was estimated using the Friedewald formula
Hermans Model-2 [18]	$Apo\ B = (0.65 \times nonHDLc) + 6.3$	The equation is proposed for use under both fasting and non-fasting conditions
Hermans Model-3 [18]	$Apo\ B = -33.12 + (0.675 \times FLDLc) + (11.95 \times \ln[TG])$	The equation is proposed for use only with fasting lipids. In this model, LDL-C was estimated using the Friedewald formula
Hermans Model-4 [18]	$Apo\ B = -33.12 + (0.675 \times SLDLc) + (11.95 \times \ln[TG])$	The equation is proposed for use only with fasting lipids. In this model, LDL-C was estimated using the Sampson formula
Hermans Model-5 [18]	$Apo\ B = -33.12 + (0.675 \times MLDLc) + (11.95 \times \ln[TG])$	The equation is proposed for use only with fasting lipids. In this model, LDL-C was estimated using the Martin-Hopkins formula
Dorairaj Model-1 [19]	$Apo\ B = 25.199 + (0.266 \times FLDLc) + (0.062 \times TG) + (0.248 \times nonHDLc)$	The original equation used direct measurements of LDL-C. In this model, LDL-C was estimated using the Friedewald formula
Dorairaj Model-2 [19]	$Apo\ B = 25.199 + (0.266 \times SLDLc) + (0.062 \times TG) + (0.248 \times nonHDLc)$	The original equation used direct measurements of LDL-C. In this model, LDL-C was estimated using the Sampson formula
Dorairaj Model-3 [19]	$Apo\ B = 25.199 + (0.266 \times MLDLc) + (0.062 \times TG) + (0.248 \times nonHDLc)$	The original equation used direct measurements of LDL-C. In this model, LDL-C was estimated using the Martin-Hopkins formula
Dorairaj Model-4 [19]	$Apo\ B = 25.077 + (0.528 \times FLDLc) + (0.138 \times TG)$	The 2nd equation developed by Dorairaj et al. used direct measurements of LDL-C. In this model, LDL-C was estimated using the Friedewald formula
Dorairaj Model-5 [19]	$Apo\ B = 25.077 + (0.528 \times SLDLc) + (0.138 \times TG)$	The 2nd equation developed by Dorairaj et al. used direct measurements of LDL-C. In this model, LDL-C was estimated using the Sampson formula
Dorairaj Model-6 [19]	$Apo\ B = 25.077 + (0.528 \times MLDLc) + (0.138 \times TG)$	The 2nd equation developed by Dorairaj et al. used direct measurements of LDL-C. In this model, LDL-C was estimated using the Martin-Hopkins formula

ApoB: Apolipoprotein B; TC: Total cholesterol; HDL-C: High-density lipoprotein cholesterol; TG: Triglycerides; LDL-C: Low-density lipoprotein cholesterol; FLDL-C: LDL-C calculated by Friedewald formula; SLDL-C: LDL-C calculated by Sampson formula; MLDL-C: LDL-C calculated by Martin-Hopkins formula; BMI: Body Mass Index.

Table 2
Patient characteristics.

Characteristics	
Age (years)	55.07 ± 13.6
Sex	
Female (n (%))	2341 (51.3)
Male (n (%))	2231 (48.7)
BMI (Kg^2/m)	25.91 ± 4.49
Lipid values	
TC (mg/dL)	211 (183.5–238)
HDL-C (mg/dL)	49 (41–58)
TG (mg/dL)	98 (69–141)
Non-HDL-C (mg/dL)	160 (134–187)
FLDLc (mg/dL)	137.2 (114.4–161.8)
SLDLc (mg/dL)	139.43 (116.15–164.03)
MLDLc (mg/dL)	137.42 (114.18–161.5)
ApoB (mg/dL)	95 (80–115)
Lp(a) (mg/dL)	9.2 (3.5–24.85)

Values are expressed as N (%), mean ± SD or median (1st–3rd quartiles). TC: total cholesterol; TG: triglycerides; HDL-C: high-density lipoprotein cholesterol, LDL-C: low-density lipoprotein cholesterol; FLDLc: LDLc calculated by Friedewald formula; SLDLc: LDL-C calculated by Sampson formula; MLDLc: LDL-C calculated by Martin-Hopkins formula; ApoB: ApoB measured by Immuno-Turbidimetric Assay; Lp(a): Lipoprotein(a).

Nutrition Examination Survey (NHANES) and the UK Biobank, have further confirmed that ApoB levels are strongly associated with cardiovascular risk and mortality [3,4]. Moreover, ApoB changes have proven to be reliable predictors of statin efficacy and the impact of other lipid-lowering therapies [5,6]. Unfortunately, direct measurement of ApoB is not always standardized and is relatively expensive, making it unavailable in several low-income settings. To address this issue, several formulas have been developed to estimate ApoB levels [7–12].

In 2008, Kulkarni et al. proposed a formula based on non-HDL-C, which eliminated the need for direct ApoB measurement by immunoturbidimetry or immunonephelometry, offering a more cost-effective and time-saving approach that became a patented method [7]. Other formulas, including those by Hermans et al. (2011), Hwang et al. (2012), Cho et al. (2012), and Dorairaj et al. (2023), also estimate ApoB from various lipid parameters [8–12]. Despite their utility, a common limitation of these formulas is the lack of validation across diverse populations, making it unclear which formuala is most suitable in different context, particularly for individuals with varying non-HDL-C and Lp(a) levels. Additionally, the performance of these formulas in specific subgroups, such as those with metabolic disorders or high cardiovascular

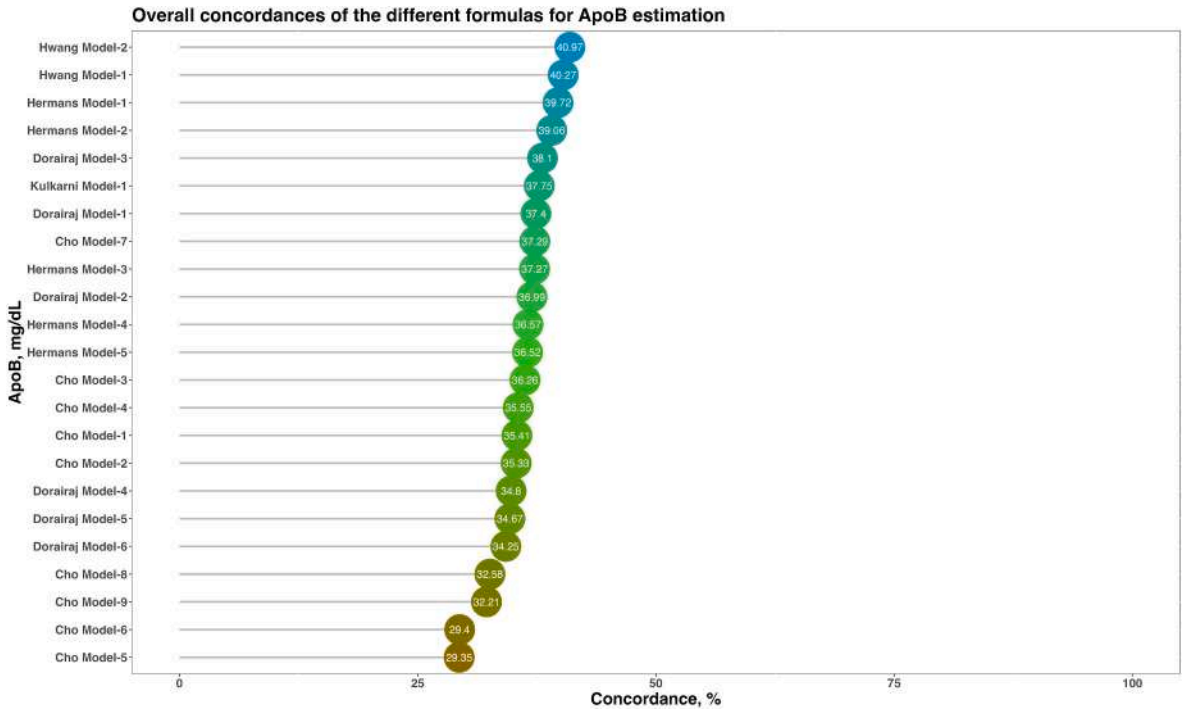

Fig. 1. Overall concordance between measured and estimated ApoB levels across the entire sample. The lollipop chart illustrates the concordance for various ApoB estimation formulas, comparing each formulas' estimates to directly measured ApoB levels.

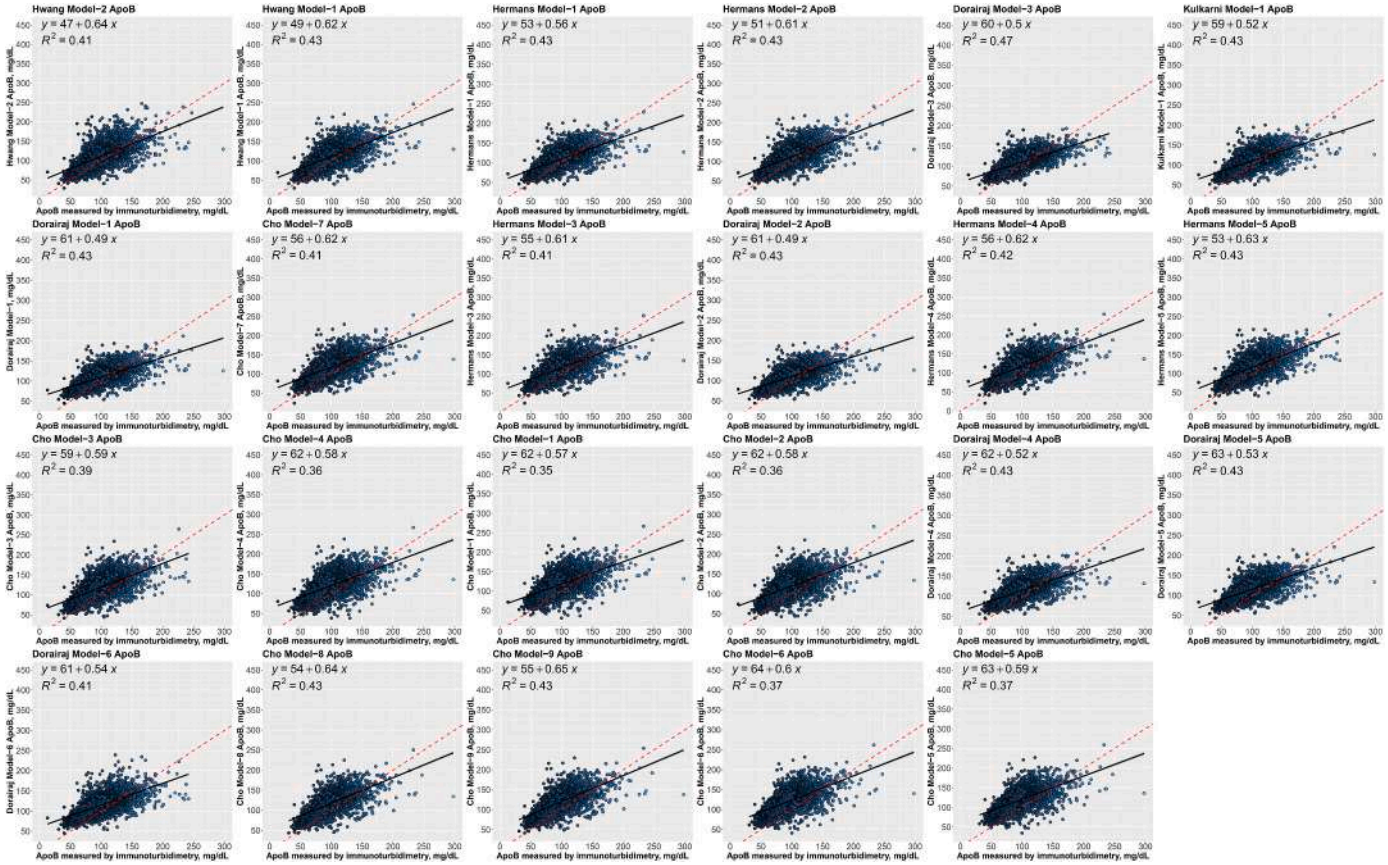

Fig. 2. Ordinary least squares linear regression analyses comparing the estimated and directly measured ApoB levels.

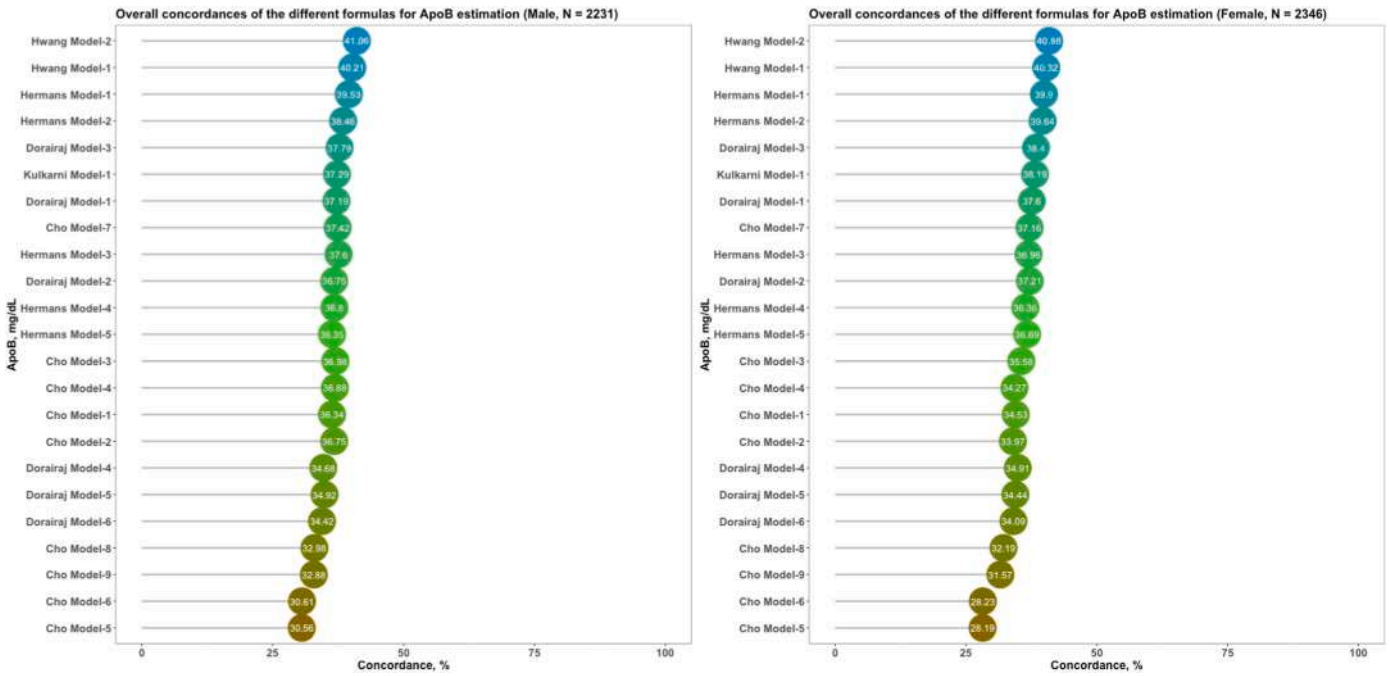

Fig. 3. Overall concordance of measured and estimated ApoB in men and women. The lollipop chart displays the concordance of various ApoB estimation formulas, comparing each formulas' estimates to directly measured ApoB levels.

Table 3
Comparisons between participants using lipid-lowering medications and those not using them.

Lipid Values	Entire Cohort	Statin Group		*P-value*
		Non-statin (n = 4437)	Statin (n = 140)	
TC	211 (183.5–238)	211 (185–239)	182 (164–210.75)	**<0.001**
HDL-C	49 (41–58)	49 (41–59)	42 (38–50)	**<0.001**
TG	98 (69–141)	97 (69–140)	109.5 (84–168)	**<0.001**
Non-HDL-C	160 (134–187)	160 (134–187)	138.5 (121–164.75)	**<0.001**
FLDLc	137.2 (114.4–161.8)	138 (115.2–162.4)	112 (95.85–136.15)	**<0.001**
SLDLc	139.42 (116.14–164.03)	140.19 (117.2–164.71)	114.65 (98.27–138.43)	**<0.001**
MLDLc	137.41 (114.18–161.5)	137.88 (115.11–162.09)	113.87 (100.08–138.65)	**<0.001**
Lp(a)	9.2 (3.5–24.85)	9.2 (3.5–24.85)	NA	-
Measured ApoB	95 (80–115)	96 (80–116)	80.5 (69–91)	**<0.001**
Kulkarni	108.93 (94.59–123.83)	108.93 (94.59–123.83)	97.07 (87.42–111.55)	**<0.001**
Hermans Model-1	108 (92.4–124.2)	108 (92.4–124.2)	95.10 (84.60–110.85)	**<0.001**
Hwang Model-1	95 (80–115)	109.5 (92.51–127.48)	93.23 (82.58–112.02)	**<0.001**
Hwang Model-2	109.08 (92.06–127.15)	108.71 (91.59–127.15)	94.67 (83.25–112.11)	**<0.001**
Cho Model-1	117.05 (100.28–135.13)	117.64 (100.87–135.58)	98.52 (86.65–116.28)	**<0.001**
Cho Model-3	117.21 (100.12–134.91)	117.55 (100.81–135.35)	99.9 (89.76–118.12)	**<0.001**
Cho Model-2	118.68 (101.57–136.77)	119.25 (102.35–137.28)	100.47 (88.43–117.95)	**<0.001**
Cho Model-4	119.16 (101.69–137.5)	119.5 (102.05–137.96)	103.76 (92.09–122.42)	**<0.001**
Cho Model-5	118.47 (101.66–135.27)	118.95 (102.28–135.85)	101.5 (90.09–119.23)	**<0.001**
Cho Model-6	121.24 (103.96–138.01)	121.78 (104.45–138.53)	105.73 (94.66–123.58)	**<0.001**
Cho Model-7	116.31 (98.60–134.19)	116.77 (98.97–134.49)	102.09 (91.61–119.17)	**<0.001**
Cho Model-8	113.99 (97.14–131.42)	114.58 (97.75–132.01)	98.67 (87.30–113.03)	**<0.001**
Cho Model-9	116.67 (99.20–134.16)	117.16 (99.67–134.46)	102.05 (91.49–117.1)	**<0.001**
Hermans Model-2	110.3 (93.4–127.85)	110.3 (93.4–127.85)	96.32 (84.95–113.38)	**<0.001**
Hermans Model-3	114.35 (97.03–132.02)	114.7 (97.51–132.6)	99.27 (87.59–116.05)	**<0.001**
Hermans Model-4	116.07 (98.22–133.88)	116.39 (98.72–134.45)	100.88 (89.92–118.51)	**<0.001**
Hermans Model-5	115.15 (96.79–132.53)	115.45 (97.25–132.93)	101.22 (90.64–119.52)	**<0.001**
Dorairaj Model-1	108.24 (91.02–126.75)	108.53 (95.08–122.71)	97.36 (88.94–111.3)	**<0.001**
Dorairaj Model-2	109 (95.11–122.59)	109.23 (95.58–123.52)	98.20 (90.15–112.05)	**<0.001**
Dorairaj Model-3	108.18 (94.58–122.31)	109.28 (95.43–122.89)	97.75 (90.55–111.55)	**<0.001**
Dorairaj Model-4	108.9 (95.03–123.13)	113.41 (98.81–128.42)	102.21 (94.01–116.45)	**<0.001**
Dorairaj Model-5	114.2 (99.43–129.61)	114.62 (99.66–129.96)	104.10 (95.01–117.56)	**<0.001**
Dorairaj Model-6	112.83 (98.11–127.95)	113.16 (98.35–128.35)	104.73 (93.85–116.63)	**<0.001**

Values are expressed as n (%), mean ± SD or median(1st—3rd quartiles). TC: total cholesterol; TG: triglycerides; HDL-C: high-density lipoprotein cholesterol; nonHDLc: non-high-density lipoprotein cholesterol; LDLc: low-density lipoprotein cholesterol; FLDLc: LDL-C calculated by Friedewald formula; SLDLc: LDL-C calculated by Sampson formula; MLDLc: LDL-C calculated by Martin-Hopkins formula; Lp(a): Lipoprotein(a); NA: not available.

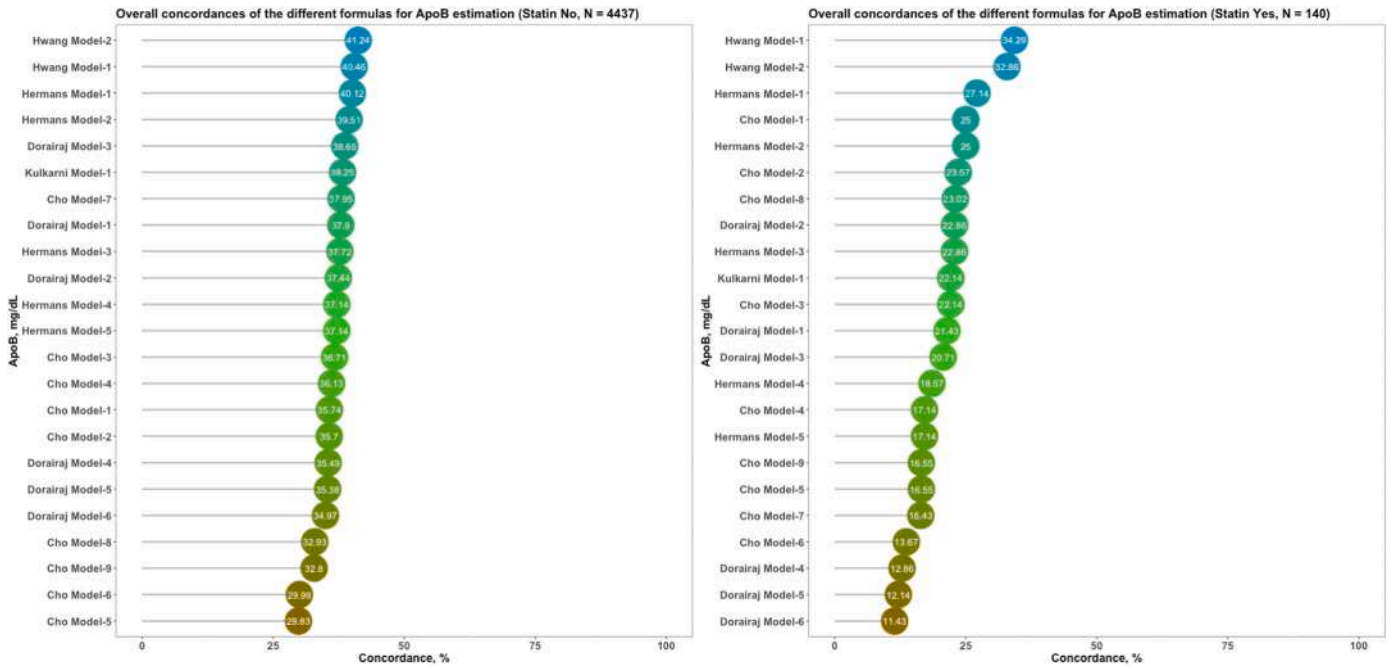

Fig. 4. Concordance between measured and estimated ApoB levels across different cohorts. This figure examines the impact of cohort structure on ApoB estimation from lipid variables. Concordance rates are generally lower in the statin group compared to the non-statin group. The top three models with the highest concordance remain consistent with the overall population, although the rankings of other models vary. Notably, the Cho models perform slightly better in the statin group than in the non-statin group.

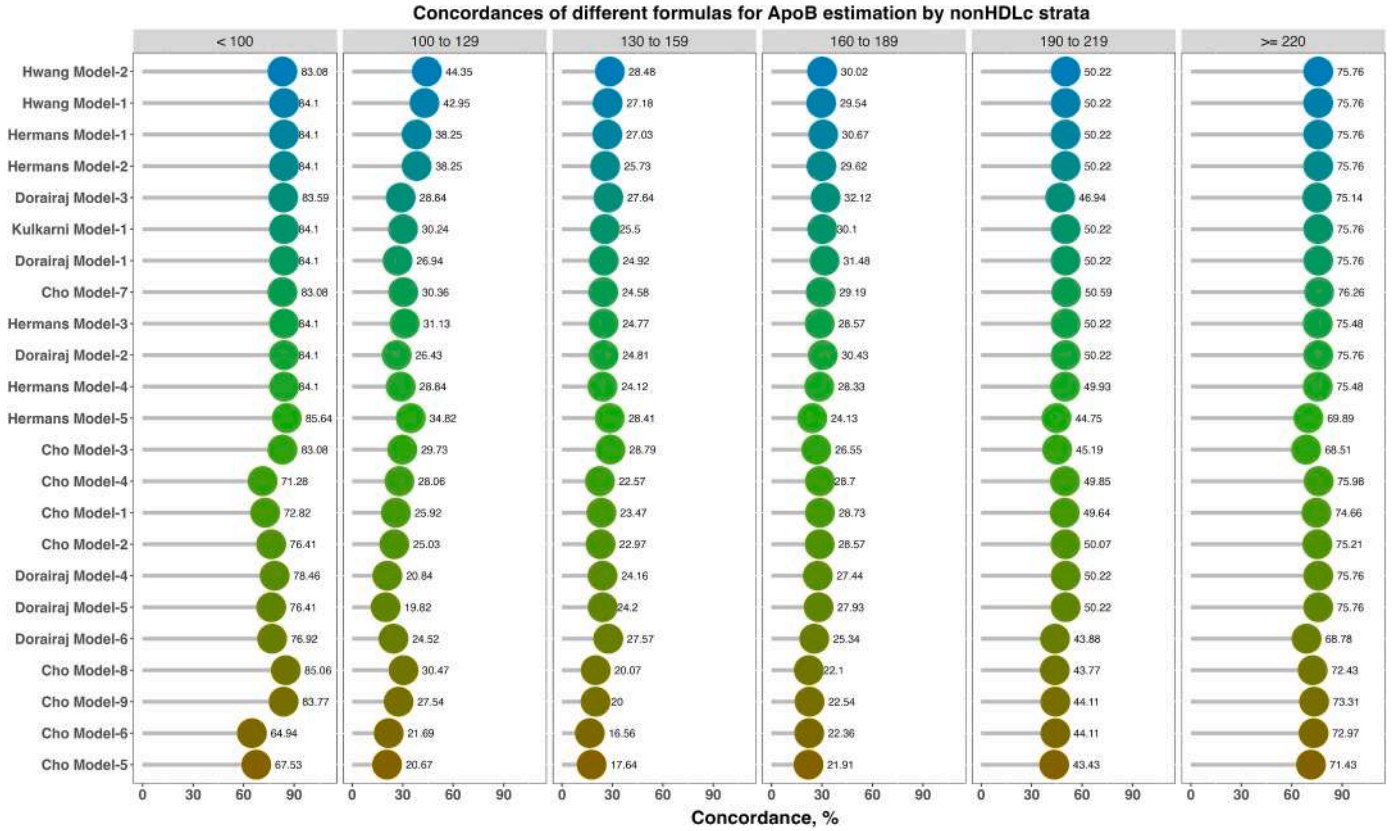

Fig. 5. Concordance between measured and estimated ApoB levels using different formulas across non-HDL-C levels. The clustered lollipop chart shows the concordance of various ApoB estimation formulas for each non-HDL-C group, based on directly measured ApoB levels.

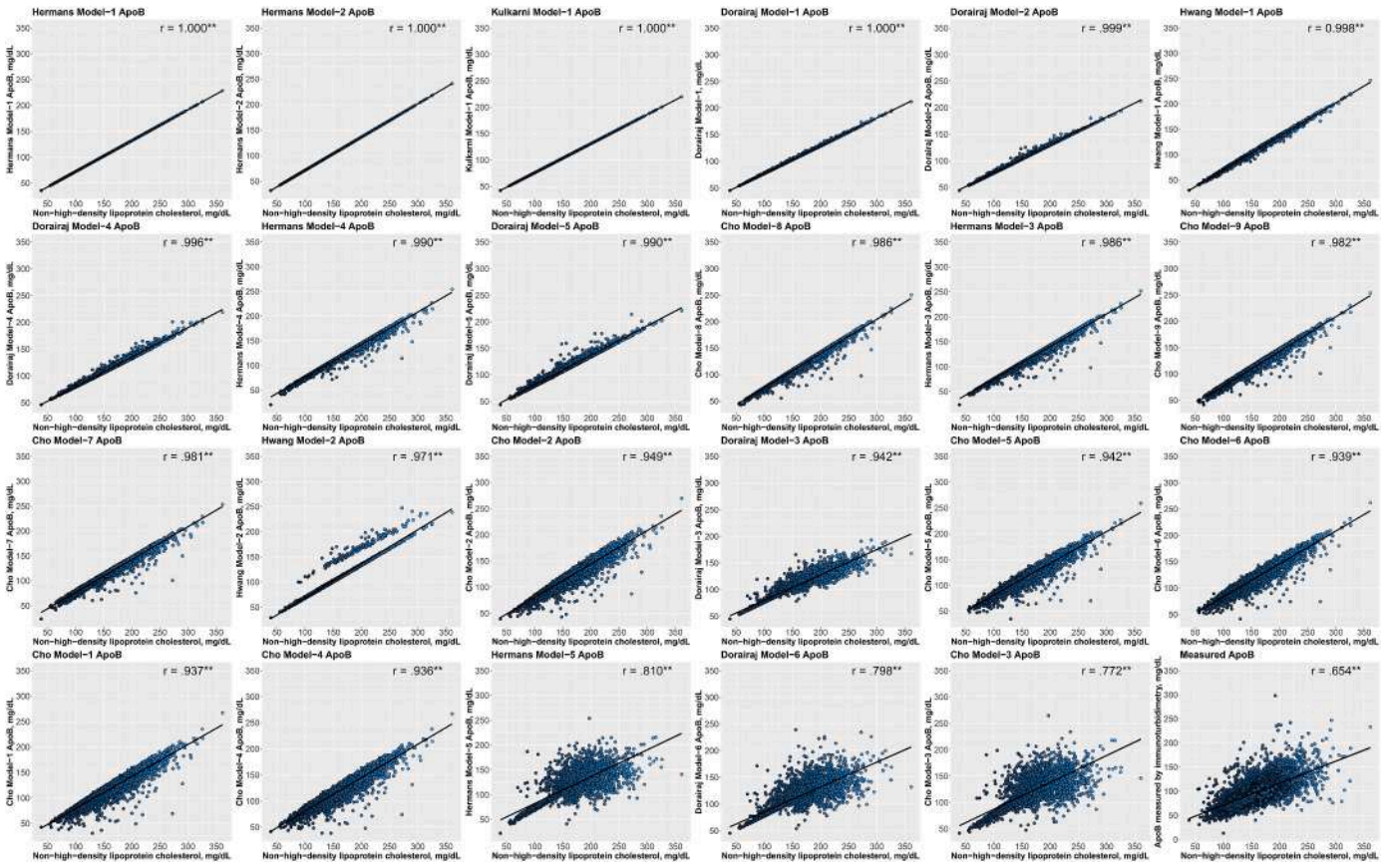

Fig. 6. Correlation between measured ApoB, non-HDL-C, and the 23 tested formulas. Formulas based on non-HDL-C (e.g., Hermans and Kulkarni models) show perfect correlation (r = 1) with non-HDL-C. In contrast, the correlation between measured ApoB and non-HDL-C is weaker (r = 0.65**). Models incorporating LDL-C and TG, such as Hwang Model-1 and Dorairaj Models 1, 2, and 4, show high correlations with non-HDL-C (r ≈ 1).

risk, has not been fully evaluated. Therefore, our study aims to compare the reliability of these formulas in estimating ApoB levels within a large cohort of South European individuals, providing a comprehensive assessment of their applicability across different populations and clinical contexts.

2. Methods

We retrospectively reviewed and assessed 11,341 clinical records of outpatients evaluated at the Lipid Clinic of the IRCCS University Hospital of Bologna from January 2020 to January 2024.

We included only clinical records in which ApoB measurement was obtained using the same standardized method. Records of patients under 18 years of age, those not of South European descent, and/or those without a complete lipid profile were excluded from the analysis.

Body mass index (BMI) was calculated as body weight in kilograms divided by height squared in meters (Kg/m^2).

Biochemical analyses were carried out on venous blood samples collected after overnight fasting (at least 12 h). Plasma was obtained by adding disodium ethylenediaminetetraacetate (Na_2EDTA) (1 mg/mL) to the blood, followed by centrifugation at 3000 RPM for 15 min at 25 °C. Immediately after centrifugation, trained personnel conducted standardized laboratory analyses [7]. The following parameters were directly measured: TC, TG, and HDL-C. Non-HDL-C was calculated by subtracting HDL-C from TC.

Friedewald's LDL-C estimation (FLDL-C) was calculated with the following formula [8]:

$$FLDLc = TC - HDLc - \left(\frac{TG}{5}\right)$$

Sampson's LDL-C estimation (SLDL-C) was calculated using the least squares formula described by Sampson [9]:

$$SLDLc = \frac{TC}{0.948} - \frac{HDLc}{0.971} - \left(\frac{TG}{8.56} + \frac{TG \times Non - HDLc}{2140} - \frac{TG^2}{16100}\right) - 9.44$$

Martin-Hopkins LDL-C estimation (MLDL-C) was calculated using the Martin-Hopkins formula [10]:

$$MLDLc = TC - HDLc - \left(\frac{TG}{\zeta}\right)$$

In this formula, ζ is an adjustable factor calculated using the median TG/VLDL-C ratio, taking into account the sublevels of TG and non-HDL-C. For MLDL-C estimation, ζ was derived from the strata-specific median ratio of TG/VLDL-C in the 180-cell table suggested by Martin et al. [10].

A standardized Immuno-Turbidimetric Assay on Roche Cobas® 6000 directly measured ApoB plasma levels [11,12], as used in NHANES [13].

We conducted an extensive search for available formulas estimating ApoB levels. The following databases were searched from their inception through May 1st, 2024: Medline, Embase, Cochrane CENTRAL and Web of Science. A total of 23 formulas were identified (Table 1), and ApoB levels were estimated using all the available formulas.

Overall concordance of ApoB estimates was calculated as the ratio of direct ApoB values that fell within the same quartile as the estimated ApoB values (i.e. <80 mg/dL, 80–94 mg/dL, 95–114 mg/dL, and ≥115 mg/dL). In addition, concordance for ApoB estimates was assessed for sublevels of Lp(a) and non-HDL-C.

Ordinary least squares linear regression analyses were performed to compare the estimated and measured ApoB values. Residual error plots were also generated to visualize the differences between each ApoB

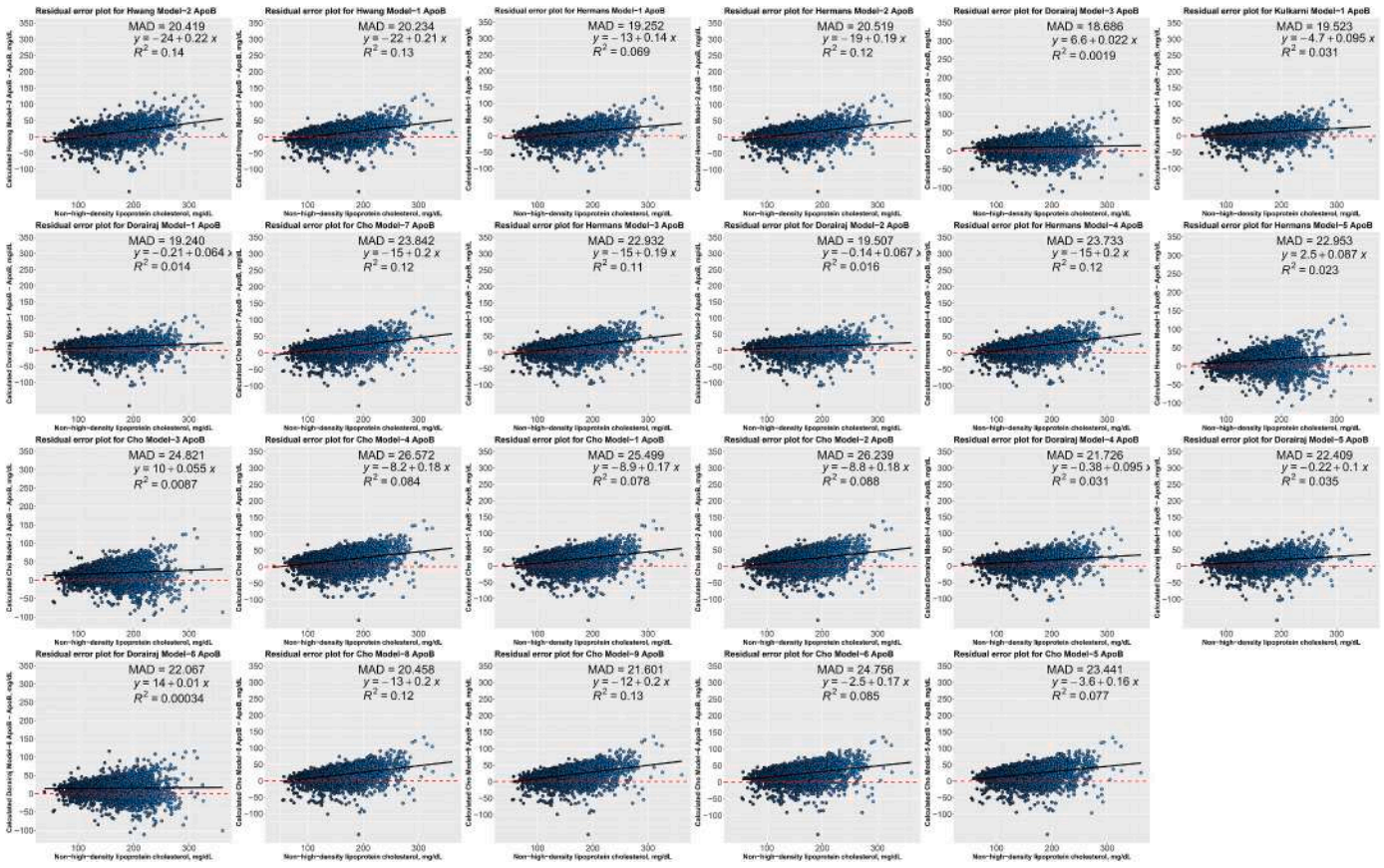

Fig. 7. Ordinary least squares linear regression analyses comparing estimated and measured ApoB values using the different tested formulas by non-HDL-C level. The x-axis represents non-HDL-C levels, and the y-axis shows the difference between estimated ApoB (calculated using 23 formulas) and directly measured ApoB levels (obtained through a standardized method). The mean absolute deviation (MAD) for each dataset is also provided in each panel for each dataset.

estimation method and ApoB measurements, categorized by Lp(a) and non-HDL-C levels.

All analyses were conducted using R 4.0.4 (www.r-project.org).

The study protocol was approved by the Institutional Ethical Board of the University Hospital of Bologna (code: LLD-RP2018) and conducted in accordance with the ethical standards outlined in the 1964 Declaration of Helsinki and its subsequent amendments. All patients provided voluntarily informed consent to participate in the study.

3. Results

The study included 4577 samples. Table 2 summarizes the main characteristics of the participants.

Hwang's models (Model-2: 40.97 %, Model-1: 40.27 %) performed the best, followed closely by Hermans (Model-1: 39.72 %) and Dorairaj Model-3 (38.1 %) (Fig. 1). In contrast, Cho's models exhibited the lowest concordance rates, with Cho Model-5 and Model-6 performing poorly (29.35 % and 29.4 %, respectively). Notably, the similar performance between Hermans and Kulkarni models reflects their reliance on non-HDL-C as a parameter, which was confirmed by the ordinary least squares linear regression analyses comparing the estimated and measured ApoB values (Fig. 2).

In the linear regression analyses, ApoB levels estimated using Dorairaj Model-3 demonstrated the highest correlation with measured ApoB levels, with an R^2 of 0.47. Hwang Model-2, which ranked first in concordance ratios, achieved an R^2 of 0.41, while the next three formulas (Hwang Model-1, Hermans Model-1, and Hermans Model-2) each had an R^2 of 0.43. In contrast, the Cho models consistently showed the lowest correlation.

Similar results to those observed in the overall cohort were found in both men and women (Fig. 3). It was noted that sex does had no significant impact on ApoB estimation. Comparisons between participants using lipid-lowering medications and those not using them are presented in Table 3.

Among the participants, 140 individuals (3.1 %) were on statin therapy, while 4437 individuals (96.9 %) were not. All lipid parameters, except for TG, were significantly higher in the non-statin group compared to the statin group ($p < 0.05$). In contrast, TG levels were significantly higher in the statin group ($p < 0.05$). Due to missing data, Lp(a) levels could not be calculated for the statin group.

Fig. 4 was generated to examine whether the structure of the cohort influences the ability to estimate ApoB from lipid variables. Concordance rates were generally lower in the statin group compared to the non-statin group. The top three models with the highest concordance rates were consistent with those observed in the overall population; however, the rankings of subsequent models varied. Notably, the Cho models performed slightly better in the statin group than the non-statin group.

Fig. 5 demonstrates the overall concordances of the various ApoB-estimating formulas based on the non-HDL-C levels. The performance of these formulas is influenced by non-HDL-C levels. As shown in the figure, the performance of each formula gradually decreases until the non-HDL-C level reaches 160 mg/dL, after which it begins to improve. The concordance ratios of the formulas exhibit a narrow range within each non-HDL-C category.

The correlation between measured ApoB, non-HDL-C, and the 23 tested formulas in shown in Fig. 6. As expected, formulas derived directly from non-HDL-C (e.g., the Hermans and Kulkarni models) had a perfect correlation ($r = 1$) with non-HDL-C. However, the correlation between measured ApoB and non-HDL-C was the weaker compared to

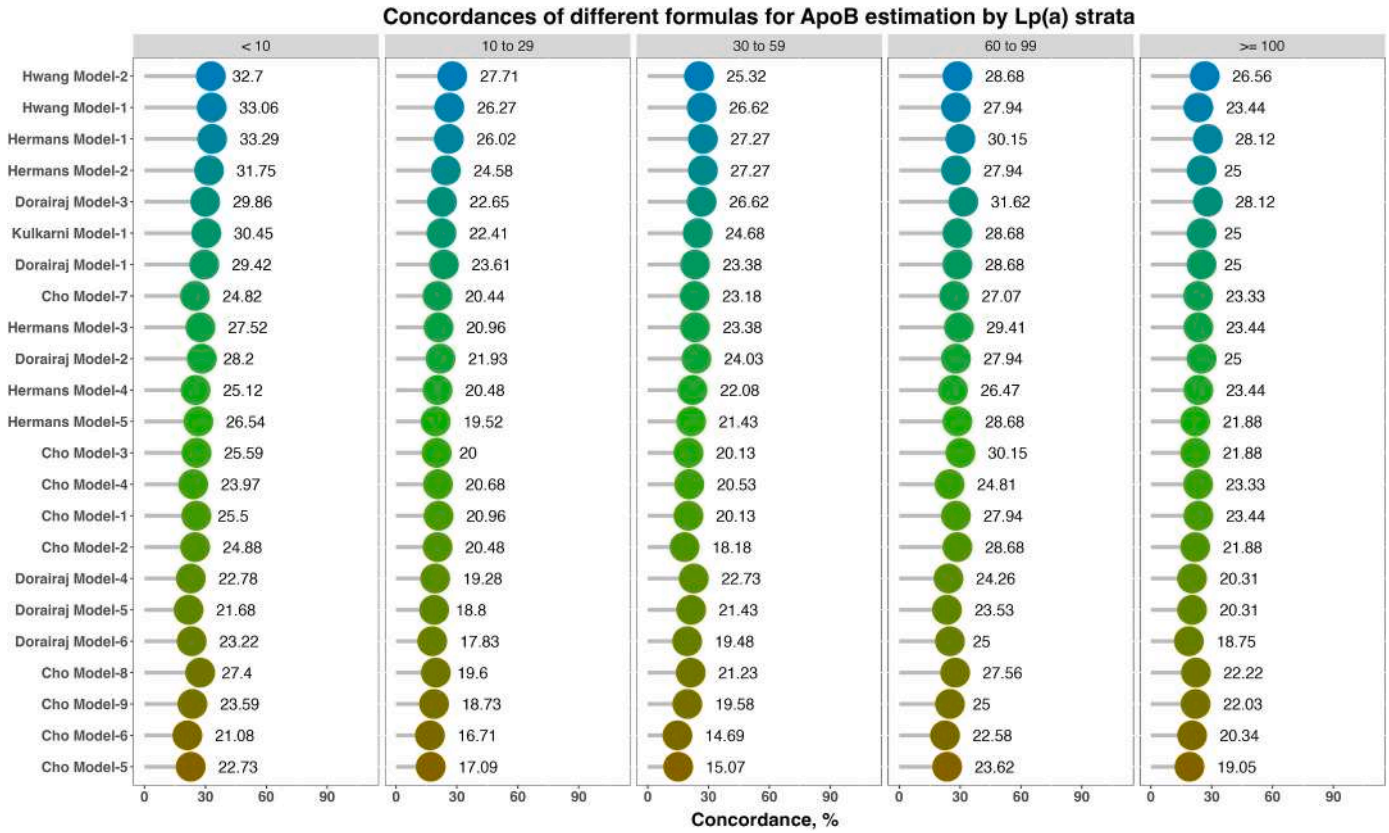

Fig. 8. Concordance between measured and estimated ApoB levels using different formulas across Lp(a) level. The clustered lollipop chart displays the concordances of various ApoB estimation formulas for each Lp(a) groups, based on directly measured ApoB levels.

the formulas (r = 0.65**). Additionally, models incorporating LDL-C and TG as variables, such as the Hwang Model-1 and Dorairaj Models 1, 2, and 4, also exhibited very high correlations with non-HDL-C ($r \approx 1$).

The residual error plots in Fig. 7 illustrate how the bias between ApoB estimations derived from the 23 tested formulas and directly measured ApoB levels changes with increasing non-HDL-C levels. In nearly all formulas, as the non-HDL-C level increased, ApoB levels were overestimated. Compared to other models, the figure shows that the difference between directly measured and calculated ApoB values in the Dorairaj models was close to zero, indicating the lowest R^2 and mean absolute deviation (MAD) scores. This suggests that the Dorairaj models are the least affected by non-HDL-C levels. In contrast, the highest MAD score was observed in the Cho Model-4.

Fig. 8 demonstrates the overall concordances of the various ApoB-estimating formulas based on Lp(a) levels. It becomes evident that the performance of the formulas decreases as Lp(a) level rise until they reach 60 mg/dL. The performance slightly improves between 60 and 100 mg/dL, but then declines again once Lp(a) levels exceed 100 mg/dL. These findings suggest that the performance of the formulas is somewhat influenced by Lp(a) levels, with low performance observed in certain value ranges that could be explained by a combination of biological variability and technical difficulties in estimating Lp(a) across the different strata.

The residual error plots in Fig. 9 show how the bias between ApoB estimations derived from the 23 tested formulas and directly measured ApoB levels changes with increasing Lp(a) levels. It was observed that all formulas consistently overestimated ApoB levels, regardless of Lp(a) levels and without any identifiable trend in the observations. Dorairaj Model-3 had the lowest MAD, indicating the smallest error among the models. In contrast, Dorairaj Model-6 had the lowest R^2 value, suggesting it was the model least influenced by Lp(a) levels.

4. Discussion

ApoB can be routinely measured on automated clinical chemistry analyzers using immunoturbidimetry or immunonephelometry, being an highly reliable assessment not significantly influenced by fasting status [20]. In fact, ApoB assays available today detect both ApoB-48 and ApoB-100 [21]. Although there are relatively few ApoB-48 particles, even in the postprandial phase, total ApoB reflects the sum of VLDL, LDL, and Lp(a), which are among the most atherogenic particles in the bloodstream [21]. This explains why ApoB could be a more relevant predictor of cardiovascular events than other lipid fractions.

A recent study by Sniderman et al., which included 293,876 adults from the UK Biobank (ages 40–73 years, 42 % men), free of cardiovascular disease at baseline and with a median follow-up of 11 years for new-onset cardiovascular disease, found that the high variability of ApoB at individual levels of LDL-C, non-HDL-C, and TG -along with significant differences in 10-year cardiovascular disease rates and the valuable residual information provided by ApoB for predicting new cardiovascular events-demonstrates that these lipid parameters are inadequate proxies for ApoB in clinical practice [22]. However, direct measurement of ApoB is not always accessible in low-income countries, and it remains relatively expensive. In this context, identifying a population-specific formula to estimate ApoB levels in large cohorts could be highly valuable. Furthermore, applying such formulas to existing datasets, where residual blood samples are unavailable, would enable the reanalysis of previous data, offering a better understanding of the relationship between ApoB levels and cardiovascular health biomarkers or disease events.

In our study, we observed that in a large population sample, plasma ApoB levels were best predicted by the Hwang [16] and Hermans [18] formulas, based on concordance ratios, independent of ApoB, non-HDL-C, and Lp(a) plasma levels. Despite this, the highest

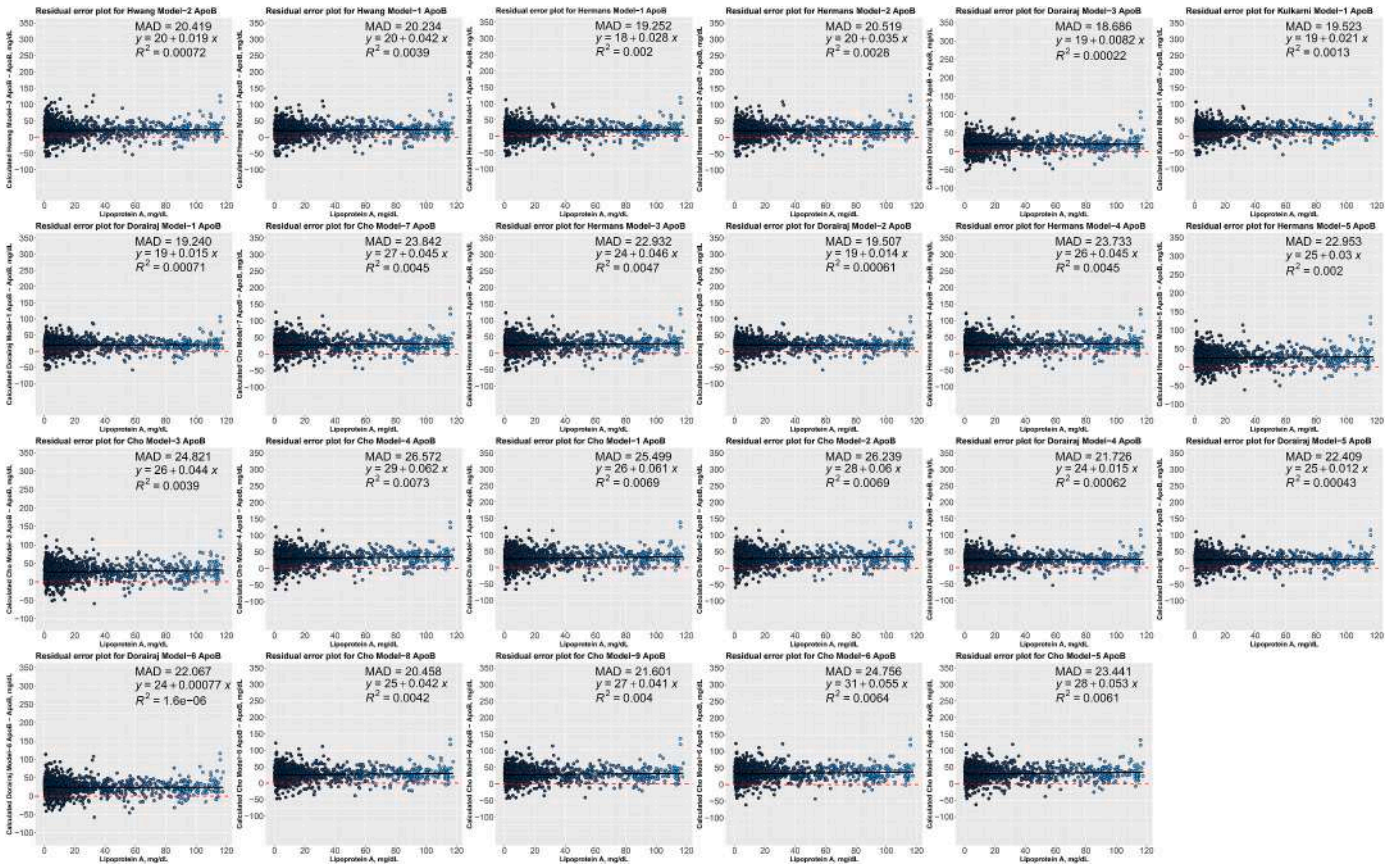

Fig. 9. Ordinary least squares linear regression analyses comparing estimated and measured ApoB values using different formulas across Lp(a) levels. The x-axis represents Lp(a) levels, and the y-axis shows the difference between estimated ApoB (calculated using 23 formulas) and measured ApoB values (obtained through a standardized method). The mean absolute deviation (MAD) for each dataset is provided in each panel.

concordance observed was only 41 %, indicating the need for the development of new formulas. Remarkably, the correlation coefficient between plasma LDL-C concentrations calculated by the Friedewald formula and those measured by β-quantification has been shown to be 73 % [23], suggesting that a concordance of >70 % would be considered adequate for a useful equation in the context of risk prediction and treatment decisions. Therefore, while the concordance observed in our study is lower than this threshold, it highlights the necessity for further refinement in the development of ApoB estimation methods for clinical applications.

In our study, we found that the Dorairaj models exhibited the highest R^2 values and were least influenced by non-HDL-C and Lp(a) levels. Notably, the LDL-C values used in Dorairaj Model-3 were calculated using the Martin-Hopkins formula. This highlights the critical importance of the methods used to measure and calculate lipid variables for ApoB estimation, as these factors significantly impact the performance of predictive formulas.

Among individuals using statins, the performance of ApoB estimation formulas was lower compared to non-users, although the Cho models showed slightly better performance in this group. The inclusion of demographic variables such as age and BMI in certain Cho models [17] suggests that these factors may be particularly influential in predicting ApoB levels, especially in individuals undergoing statin therapy. Interestingly, the Hwang formula (both Model-1 and Model-2), which provided the best performance in each group, is similar to those used for LDL-C estimation (e.g. Friedewald [8], Martin-Hopkins [10], Chen [24], Vujovic [25] etc.), as they utilize the same variables and form a linear combination of these variables. The literature suggests that formulas designed for LDL-C estimation can be adapted for ApoB estimation. Moreover, using machine learning methods to model ApoB estimation, while accounting for factors like population diversity and non-lipid parameters, will significantly enhance its accuracy.

The Hwang model has already been demonstrated to correlate with surrogate markers of cardiovascular risk – such as high-sensitivity C-reactive protein (hs-CPR) levels, microalbuminuria, coronary artery calcium (CAC) score, and Framingham risk score-just as directly measured ApoB does, in a huge Korean population sample (n = 78,125) [26]. It has also been demonstrated to be superior to estimated LDL-C in predicting cardiovascular events in another large sample of the Korean general population (8713 individuals with a mean age of 52.2 years, followed biannually for an average of 8.1 years) [27]. Furthermore, in high-risk patients with a history of coronary heart disease enrolled in the Treating to New Targets (TNT) and Incremental Decrease in Endpoints through Aggressive Lipid lowering (IDEAL) trials, ApoB levels estimated using the Hwang model provided cardiovascular disease risk predictions comparable to those from directly measured ApoB [28]. However, in both the TNT and IDEAL trials, South-European individuals were relatively underrepresented.

It is important to note that our study has several limitations. First, while our population sample is heterogeneous with respect to clinical characteristics, it is large and representative of the general population. Second, the sample was limited to Caucasian individuals to exclude the potential influence of ethnic variations or distinct dietary patterns on estimated ApoB levels. On the other hand, we conducted our analysis on a large cohort of South European individuals, with a wide range of non-HDL-C and Lp(a) levels. Furthermore, while a relatively large number of studies have previously compared the reliability of various formulas for estimating LDL-C levels [29–31], there are no similar studies focused on ApoB estimation. It should also be acknowledged that, since we did not include repeated ApoB measurements for each patient, we could not

assess the impact of biological variability on the ApoB estimates. This could introduce further uncertainty beyond measurement error, which we could not account for. In this context, future studies that include repeated ApoB measurements would allow for more accurate quantification of both measurement error and biological variability. Finally, the absence of cardiovascular event data in our current dataset prevented us from evaluating the performance of estimated ApoB in predicting cardiovascular disease. This limitation highlights a significant gap in our understanding, emphasizing the need for future research to assess the predictive power of estimated ApoB compared to directly measured ApoB in identifying individuals at high risk of cardiovascular disease. Future studies should also clarify whether estimated ApoB is superior to non-HDL-C in predicting cardiovascular risk. The reliance of the Hermans model on non-HDL-C stems from the high correlation between ApoB and non-HDL-C, their strong metabolic association, and evidence suggesting that these two markers can be used interchangeably in certain contexts [15]. For instance, the discrimination power of non-HDL-C is similar to that of ApoB in ranking diabetic patients according to their atherogenic cholesterol and lipoprotein burden [15]. Similarly, the Kulkarni model [14] demonstrated high concordance ratios in our study, further reinforcing the strong correlation between ApoB and non-HDL-C. This finding aligns with the consensus recommendations issued by the American Diabetes Association (ADA) and the American College of Cardiology (ACC), which emphasize the importance of focusing on non-HDL-C and ApoB levels in patients likely to have small LDL particles, such as diabetic patients [16]. In addition, the Adult Treatment Panel III (ATP III) recommended that, in individuals with elevated TG, non-HDL-C should serve as a secondary therapeutic target after LDL-C, as it appears to be a more reliable predictor of cardiovascular risk than LDL-C, particularly in statin-treated patients [18].

While recent evidence has increasingly positioned ApoB as the most reliable predictor of atherosclerotic cardiovascular disease, studies specifically investigating the comparability between estimated and measured ApoB remain limited [20,25,32,33]. Validating the clinical utility of estimated ApoB is crucial, particularly given its potential as a widely accessible and cost-effective alternative in settings where directly measured ApoB is unavailable or prohibitively expensive. Moreover, there is a pressing need to develop more robust estimation formulas that show improved concordance with measured ApoB, especially since the highest concordance rate achieved by currently available formulas is only 41 %.

In conclusion, simple formulas based on low-cost and widely available parameters can predict ApoB levels independently of ApoB, non-HDL-C, and Lp(a) plasma levels. This approach could be particularly useful for estimating ApoB plasma levels in low-income countries.

Financial support

This research was supported by the Institutional Fundings by the University of Bologna (RFO2022 and RFO2023).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to express our sincere gratitude to the Reviewers for their invaluable feedback and constructive comments, which have significantly contributed to the improvement of the quality of our manuscript.

References

[1] S. Bilgic, A.D. Sniderman, Low-density lipoprotein cholesterol, non-high-density lipoprotein cholesterol and apolipoprotein B for cardiovascular care, Curr. Opin. Cardiol. 39 (1) (2024) 49–53, https://doi.org/10.1097/HCO.0000000000001100.
[2] A.D. Sniderman, S. Islam, S. Yusuf, M.J. McQueen, Discordance analysis of apolipoprotein B and non-high density lipoprotein cholesterol as markers of cardiovascular risk in the INTERHEART study, Atherosclerosis 225 (2) (2012) 444–449, https://doi.org/10.1016/j.atherosclerosis.2012.08.039.
[3] M.Q. Yan, Y. Huang, X.C. Liu, C.L. Chen, D. Zhou, Y.Q. Huang, Y.Q. Feng, Association of apolipoprotein B with all-cause and cardiovascular mortality among adults: results from the national health and nutrition examination surveys, Am. J. Med. Sci. 366 (5) (2023) 367–373, https://doi.org/10.1016/j.amjms.2023.07.012.
[4] G. Yang, A.M. Mason, A.M. Wood, C.M. Schooling, S. Burgess, Dose-response associations of lipid traits with coronary artery disease and mortality, JAMA Netw. Open 7 (1) (2024) e2352572, https://doi.org/10.1001/jamanetworkopen.2023.52572.
[5] C.D.L. Johannesen, M.B. Mortensen, A. Langsted, B.G. Nordestgaard, Apolipoprotein B and non-HDL cholesterol better reflect residual risk than LDL cholesterol in statin-treated patients, J. Am. Coll. Cardiol. 77 (11) (2021) 1439–1450, https://doi.org/10.1016/j.jacc.2021.01.027.
[6] S.U. Khan, M.U. Khan, S. Valavoor, M.S. Khan, V. Okunrintemi, M.A. Mamas, T. M. Leucker, M.J. Blaha, E.D. Michos, Association of lowering apolipoprotein B with cardiovascular outcomes across various lipid-lowering therapies: systematic review and meta-analysis of trials, Eur J Prev Cardiol 27 (12) (2020) 1255–1268, https://doi.org/10.1177/2047487319871733.
[7] A.F.G. Cicero, F. Fogacci, M. Giovannini, E. Rizzoli, E. Grandi, S. D'Addato, C. Borghi, The effect of dietary supplementation with plant sterols on total and LDL-cholesterol in plasma is affected by adherence to mediterranean diet: insights from the DESCO randomized clinical study, Nutrients 15 (21) (2023) 4555, https://doi.org/10.3390/nu15214555.
[8] 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.
[9] M. Sampson, C. Ling, Q. Sun, R. Harb, M. Ashmaig, R. Warnick, A. Sethi, J. K. Fleming, J.D. Otvos, J.W. Meeusen, S.R. Delaney, A.S. Jaffe, R. Shamburek, M. Amar, A.T. Remaley, A new equation for calculation of low-density lipoprotein cholesterol in patients with normolipidemia and/or hypertriglyceridemia, JAMA Cardiol 5 (5) (2020) 540–548, https://doi.org/10.1001/jamacardio.2020.0013.
[10] 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 310 (19) (2013) 2061–2068, https://doi.org/10.1001/jama.2013.280532.
[11] Cobas 6000 Analyzer Operator's Manual. Roche Diagnostics Inc., 9115 Hague Road Indianapolis, IN 46250-0457. Version 5.0.
[12] Roche APOBT reagent package insert. Version 2012-02, V8. Roche Diagnostics Inc., 9115 Hague Road Indianapolis, IN 46250-0457.
[13] M.D. Carroll, M.E. Mussolino, M. Wolz, P.R. Srinivas, Trends in apolipoprotein B, non-high-density lipoprotein, and low-density lipoprotein for adults 60 Years and older by use of lipid-lowering medications: United States, 2005 to 2006 through 2013 to 2014, Circulation 138 (2) (2018) 208–210, https://doi.org/10.1161/CIRCULATIONAHA.117.031982.
[14] K. Kulkarni, ApoB measurement system and method, Age (Chester): DLA PIPER US LLP; IPC8 Class: AG01N3392FI; USPC Class: (2008) 43652. http://www.faqs.org/patents/app/20080261313.
[15] M.P. Hermans, F.M. Sacks, S.A. Ahn, M.F. Rousseau, Non-HDL-cholesterol as valid surrogate to apolipoprotein B100 measurement in diabetes: discriminant Ratio and unbiased equivalence, Cardiovasc. Diabetol. 10 (2011) 20, https://doi.org/10.1186/1475-2840-10-20.
[16] Y.C. Hwang, H.Y. Ahn, W.J. Lee, C.Y. Park, S.W. Park, An equation to estimate the concentration of serum apolipoprotein B, PLoS One 7 (12) (2012) e51607, https://doi.org/10.1371/journal.pone.0051607.
[17] D.S. Cho, S. Woo, S. Kim, C.D. Byrne, J.H. Kong, K.C. Sung, Estimation of plasma apolipoprotein B concentration using routinely measured lipid biochemical tests in apparently healthy Asian adults, Cardiovasc. Diabetol. 11 (2012) 55, https://doi.org/10.1186/1475-2840-11-55.
[18] M.P. Hermans, S.A. Ahn, M.F. Rousseau, Discriminant ratio and biometrical equivalence of measured vs. calculated apolipoprotein B100 in patients with T2DM, Cardiovasc. Diabetol. 12 (2013) 39, https://doi.org/10.1186/1475-2840-12-39.
[19] P. Dorairaj, S. Manickam, S. Raju, A. Chandrasekhar, An android app "apolipoprotein B calculator" calculates apolipoprotein B using regression analysis and neural network–using the Friedewald equation is the same as directly measured low-density lipoprotein cholesterol and better at low-density lipoprotein levels, J Indian Coll Cardiol 13 (2) (2023) 69–75, https://doi.org/10.4103/jicc.jicc_34_22.
[20] D.E. Soffer, N.A. Marston, K.C. Maki, T.A. Jacobson, V.A. Bittner, J.M. Peña, G. Thanassoulis, S.S. Martin, C.F. Kirkpatrick, S.S. Virani, D.L. Dixon, C. M. Ballantyne, A.T. Remaley, 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 18 (5) (2024 Sep-Oct) e647–e663, https://doi.org/10.1016/j.jacl.2024.08.013.
[21] A.D. Sniderman, G. Thanassoulis, T. Glavinovic, A.M. Navar, M. Pencina, A. Catapano, B.A. Ference, Apolipoprotein B particles and cardiovascular disease: a

narrative review, JAMA Cardiol 4 (12) (2019) 1287–1295, https://doi.org/10.1001/jamacardio.2019.3780.
[22] A.D. Sniderman, L. Dufresne, K.M. Pencina, S. Bilgic, G. Thanassoulis, M. J. Pencina, Discordance among apoB, non-high-density lipoprotein cholesterol, and triglycerides: implications for cardiovascular prevention, Eur. Heart J. 45 (27) (2024 Jul 12) 2410–2418, https://doi.org/10.1093/eurheartj/ehae258.
[23] A.J. Tremblay, H. Morrissette, J.M. Gagné, J. Bergeron, C. Gagné, P. Couture, Validation of the Friedewald formula for the determination of low-density lipoprotein cholesterol compared with beta-quantification in a large population, Clin. Biochem. 37 (9) (2004 Sep) 785–790, https://doi.org/10.1016/j.clinbiochem.2004.03.008.
[24] Y. Chen, X. Zhang, B. Pan, X. Jin, H. Yao, B. Chen, Y. Zou, J. Ge, H. Chen, A modified formula for calculating low-density lipoprotein cholesterol values, Lipids Health Dis. 9 (2010) 52, https://doi.org/10.1186/1476-511X-9-52.
[25] A. Vujovic, J. Kotur-Stevuljevic, S. Spasic, N. Bujisic, J. Martinovic, M. Vujovic, V. Spasojevic-Kalimanovska, A. Zeljkovic, D. Pajic, Evaluation of different formulas for LDL-C calculation, Lipids Health Dis. 9 (2010) 27, https://doi.org/10.1186/1476-511X-9-27.
[26] J.D. Kim, Y.C. Hwang, H.Y. Ahn, C.Y. Park, Validation of a newly developed equation for estimating serum apolipoprotein B: associations with cardiovascular disease surrogate markers in Koreans, Yonsei Med. J. 58 (5) (2017) 975–980, https://doi.org/10.3349/ymj.2017.58.5.975.
[27] Y.C. Hwang, C.Y. Park, H.Y. Ahn, N.H. Cho, Prediction of future development of cardiovascular disease with an equation to estimate apolipoprotein B: a community-based cohort study, Medicine 95 (24) (2016) e3644, https://doi.org/10.1097/MD.0000000000003644.
[28] Y.C. Hwang, H.Y. Ahn, K.H. Han, S.W. Park, C.Y. Park, Prediction of future cardiovascular disease with an equation to estimate apolipoprotein B in patients with high cardiovascular risk: an analysis from the TNT and IDEAL study, Lipids Health Dis. 16 (1) (2017) 158, https://doi.org/10.1186/s12944-017-0549-8.
[29] A.F.G. Cicero, F. Fogacci, D. Patrono, R. Mancini, E. Ramazzotti, C. Borghi, S. D'Addato, BLIP Study group, Application of the Sampson equation to estimate LDL-C in children: comparison with LDL direct measurement and Friedewald equation in the BLIP study, Nutr. Metabol. Cardiovasc. Dis. 31 (6) (2021) 1911–1915, https://doi.org/10.1016/j.numecd.2021.02.034.
[30] S. Bolat, G. Ertürk Zararsız, K. Doğan, N. Kochan, S.I. Yerlitaş, A. Cephe, G. Zararsız, A.F.G. Cicero, Concordance of LDL-C estimating equations with direct enzymatic measurement in diabetic and prediabetic subjects, J. Clin. Med. 12 (10) (2023) 3570, https://doi.org/10.3390/jcm12103570.
[31] J. Ramesh, S. Selvarajan, S. Krishnamurthy, S. Kopula Sathyamoorthy, D. Senthil Kumar, Evaluation of 13 formulae for calculated LDL-C using direct homogenous assay in a South Indian population, J Appl Lab Med 9 (5) (2024 Sep 3) 963–977, https://doi.org/10.1093/jalm/jfae048.
[32] X. Kong, Y. Cai, Y. Li, P. Wang, Causal relationship between apolipoprotein B and risk of atherosclerotic cardiovascular disease: a mendelian randomization analysis, Health Inf. Sci. Syst. 13 (2025) 13, https://doi.org/10.1007/s13755-024-00323-5.
[33] J. Behbodikhah, S. Ahmed, A. Elyasi, L.J. Kasselman, J. De Leon, A.D. Glass, A. B. Reiss, Apolipoprotein B and cardiovascular disease: biomarker and potential therapeutic target, Metabolites 11 (10) (2021) 690, https://doi.org/10.3390/metabo11100690.