

SUPPLEMENTAL MATERIAL

Serum metabolomic signatures can predict sub-clinical atherosclerosis in patients with SLE

Leda Coelewij MSc^{1,2}, Kirsty E Waddington PhD#^{1,2}, George A Robinson PhD#^{2,3}, Elvira Chocano MSc^{1,2}, Thomas McDonnell PhD², Filipa Farinha MD², Junjie Peng MSc^{2,3}, Pierre Dönnes PhD^{2,4}, Edward Smith PhD², Sara Croca PhD², Jyoti Bakshi MBBS², Maura Griffin PhD⁵, Andrew Nicolaides PhD^{5,6}, Anisur Rahman PhD FRCP², Elizabeth C Jury PhD^{2†*}. Ines Pineda-Torra PhD^{1†*}

contributed equally; †Share senior authorship

¹Centre for Cardiometabolic and Vascular Science, Department of Medicine, University College London, London W1CE 6JF, U.K.

²Centre for Rheumatology Research and ³Centre for Adolescent Rheumatology Versus Arthritis, Department of Medicine, University College London, London W1CE 6JF, U.K.

⁴Scicross AB, Skövde, Sweden

⁵Vascular Screening and Diagnostic Centre, Weymouth Street, London, UK

⁶St Georges London/Nicosia Medical School, University of Nicosia, Cyprus

*Corresponding authors:

Liz Jury (e.jury@ucl.ac.uk), Ines Pineda-Torra (i.torra@ucl.ac.uk),

Major Resources Table

In order to allow validation and replication of experiments, all essential research materials listed in the Methods should be included in the Major Resources Table below. Authors are encouraged to use public repositories for protocols, data, code, and other materials and provide persistent identifiers and/or links to repositories when available. Authors may add or delete rows as needed.

Data & Code Availability

Description	Source / Repository	Persistent ID / URL
mixOmics: Omics Data Integration Project mixOmics offers a wide range of multivariate methods for the exploration and integration of biological datasets with a particular focus on variable selection. The package proposes several sparse multivariate models we have developed to identify the key variables that are highly correlated, and/or explain the biological outcome of interest.	Bioconductor	http://www.bioconductor.org/packages/release/bioc/html/mixOmics.html DOI: 10.18129/B9.bioc.mixOmics
caret: Classification and Regression Training Misc functions for training and plotting classification and regression models.	Comprehensive R Archive Network (CRAN)	https://CRAN.R-project.org/package=caret
ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics A system for 'declaratively' creating graphics, based on "The Grammar of Graphics". You provide the data, tell 'ggplot2' how to map variables to aesthetics, what graphical primitives to use, and it takes care of the details.	CRAN	https://CRAN.R-project.org/package=ggplot2
kernlab: Kernel-Based Machine Learning Lab Kernel-based machine learning methods for classification, regression, clustering, novelty detection, quantile regression and dimensionality reduction. Among other methods 'kernlab' includes Support Vector Machines, Spectral Clustering, Kernel PCA, Gaussian Processes and a QP solver.	CRAN	https://CRAN.R-project.org/package=kernlab
MLeval: Machine Learning Model Evaluation Straightforward and detailed evaluation of machine learning models. 'MLeval' can produce receiver operating characteristic (ROC) curves, precision-recall (PR) curves, calibration curves, and PR gain curves. 'MLeval' accepts a data frame of class probabilities and ground truth labels, or, it can automatically interpret the Caret train function results from repeated cross validation, then	CRAN	https://CRAN.R-project.org/package=MLeval

select the best model and analyse the results. 'MLeval' produces a range of evaluation metrics with confidence intervals.		
randomForest: Breiman and Cutler's Random Forests for Classification and Regression Classification and regression based on a forest of trees using random inputs, based on Breiman (2001) <DOI:10.1023/A:1010933404324>.	CRAN	https://CRAN.R-project.org/package=randomForest
Package ‘forestplotNMR’ An R package to plot, in a one- or two-column forestplot layout, associations of epidemiological analysis on NMR metabolomics biomarker data. The associations may be linear (linear regression), odds ratios (logistic regression) or hazard ratios (e.g. Cox Proportional hazards).	GitHub	https://github.com/mariakalimeri/forestplot
Subclinical atherosclerosis in SLE – Metabolomics data.	Mendeley Data	http://dx.doi.org/10.17632/fmygdybj2h.1

Supplemental Methods

Ultrasound scanning: Detailed vascular scanning methods

All scans were performed by an experienced vascular scientist (MG) and strict scanning protocols were observed. The methodology used and equipment settings has been published elsewhere [1, 2]. Briefly, the Philips iU22 ultrasound system (Philips Ultrasound, Bothell, USA) was used with a linear array L9-3 MHz transducer. Technical ultrasound parameters were pre-set to the most linear post-processing curve available, maximum dynamic range (170 db), persistence (low) and frame rate (high), which remained constant.

Each carotid bifurcation was examined transversely, and then longitudinally to ensure optimal demonstration of the intima–media complex of both the near and far walls of the common carotid artery 1.5–2.0 cm proximal to the carotid bulb. All participants were examined in the supine position with the neck slightly extended. IMT measurements were performed on-screen using the QLAB IMT plug-in. The mean of the measurements from both carotid arteries was used in the analysis (IMTcc).

Both carotid and common femoral bifurcations were also examined for presence of plaque. An arterial bifurcation was classified as being affected by plaque if there was a focal thickening of greater than 1.2 mm [3, 4]. Careful transverse and longitudinal scanning with the use of colour flow allowed better appreciation of the geometry and borders of the plaque, thereby allowing a more accurate assessment and optimal image capture. The image was magnified so that the whole plaque was included along with a clearly visualised section of adventitia within the vicinity of the plaque. The ultrasound images were stored as DICOM files and transferred via DVD to a dedicated PC for application of imaging software program.

The “Carotid Plaque Texture Analysis” dedicated software for ultrasonic arterial wall and atherosclerotic plaque measurements was used (LifeQ Medical Ltd, www.lifeqmedical.com).

Total plaque area (TPA) was defined as the sum of the cross-sectional areas of all plaques seen in longitudinal images. Echogenicity of plaque was quantified in terms of the Grey Scale Median (GSM). Images were normalized using linear scaling with two reference points as follows; blood was defined as having grey scale = 0 and adventitia was defined as having grey scale = 190. Thus the echogenicity of any point within the image can be quantified on a scale between 0 and 190 and the median grey scale of all plaques within an individual subject can be calculated. The scale in mm present on the side of each image was entered into the software and the plaque was outlined manually. Plaque area in mm^2 and GSM were calculated by the software.

Data analysis

Logistic regression

To assess the association of plaque with NMR metabolomic biomarker data, adjusted logistic regressions were performed for each individual serum metabolite, adjusted for ethnicity, age, mean arterial pressure (MAP), global BILAG score, and treatments at the time of the scan. Results were visualized in a forest plot using R package forestplotNMR [5][6].

Predictive models

The data analysis pipeline used in shown in [Figure 1](#). RStudio (The R Foundation, Vienna, Austria) (R Core Team, 2019)[7] and Orange 3.24.1 (Bioinformatics Lab, University of Ljubljana, Slovenia)[8] were used for ML analysis as we have shown previously [9]. Five different supervised

learning algorithms were implemented: support vector machine (SVM), logistic regression with and without interactions, decision trees, and random forest classification. The outcome of the learning algorithms was to predict whether an SLE patient is likely to have plaque. Predictive models were generated from metabolite concentrations at the time of the scan.

Missing data: Features with >10% missing data were excluded (pyruvate). Remaining missing values were imputed using k-nearest neighbors with k=5. Imputed data: SLE vs HCs n=154 and for SLEp vs SLE-NP n=5.

Homology reduction: To reduce homology, if two features had a correlation co-efficient >0.95 then the feature with the greatest mean absolute correlation with the remaining features was removed. This left 124 metabolites.

Data scaling: Metabolite concentrations were centered on the mean and scaled to the standard deviation.

Predictors: The independent variables included in these models were the homology reduced dataset (124 metabolites) and cohort information (age, ethnicity, mean arterial blood pressure, disease duration, BILAG-2004 score, treatments at the time of first scan). Sex was not considered as all participants were female. Full lists of the predictors contributing to each model are included in [Supplementary Data File V](#).

Support vector machine (SVM): A supervised classification method which creates a hyperplane to optimally separate data into two classes[10]. As this data set was not linearly separable, the radial basis function kernel was used. Values for C, epsilon, and gamma were tuned using the R Package e1071[11]. The parameters were set to C=4.0, epsilon=0.1, gamma=0.01.

Decision tree: A supervised ML method which classifies incidents according to their features. Decision trees are built using forms of impurity measures, such as information gain and entropy[12]. To prevent overfitting, decision trees were limited to a depth of 4 and subsets of 5 or less were not split further.

Random Forest (RF): A ML algorithm for classification, consisting of numerous decision trees which can increase model accuracy without the risk of model overfitting [13] which is often a problem when analysing data with small sample size. Importance of each feature was quantified by the Gini index, which represents the total variance across the two classes, the purity of each node and the quality of each split. The values for mtry and ntree were tuned using the R Package caret[14]. The parameters were set to mtry=10, ntree=10,000.

Logistic regression with/without interactions (LR/LR+I): The least absolute shrinkage and selection operator (lasso) method uses the absolute value of the co-efficient as a penalty to shrink less important features to zero. The strength of shrinkage is determined by tuning the regularization variable lambda (λ). $\ln(\lambda)$ was optimized using the R package [13] and set to lambda=0.059 and lambda=0.065 for logistic regression with and without interaction respectively.

Model performance: Ten-fold cross-validation was used to evaluate model performance. Validation was performed in Orange for the decision tree and in R for all other models, using a balanced splitting. This balanced splitting was done as the original sample set had an unbalanced NP:P (30:50) ratio. The balanced method was applied in the bootstrap dataset construction, wherein the bootstrap dataset was first selected from the minority class (NP, n=30) whilst randomly drawing the exact number (n=30) from the majority class (P). Data were randomly partitioned into 10 groups with almost equal size. Nine groups were used as training data for model construction and the remaining group was used as validation data. The process was repeated for all 10 folds until each observation in the data is used for validation purpose once. The average performance of the 10 models was used as the result of the 10-fold cross validation. The following performance metrics were calculated from the confusion matrices: (i) F1 score - a weighted average of precision

(positive predictive value) and recall (sensitivity), (ii) specificity - the true negative rate, and (iii) classification accuracy (CA) - the proportion of correctly classified cases.

Partial Least Squares Discriminant Analysis (sPLS-DA): A sPLS-DA is a supervised clustering ML approach that combines parameter selection and classification into one operation. This analysis was performed with the R package mixOmics [15] using the NMR metabolomic data for the SLE vs HC comparison ([Supplementary Table III](#)) and the features including metabolites selected for the lasso LR+I model in the SLE-P vs SLE-NP comparison ([Supplementary Table V](#)). Ten-fold cross-validation with 50 repetitions was applied to prevent model overfitting. sPLS-DA models with different component numbers were assessed by 10-fold cross-validation, using the overall error rate to evaluate model performance. For the SLE vs HC and the SLE-P vs SLE-NP comparisons, sPLS-DA with four components gave the lowest overall estimation error rate (14.3% and 20.6% respectively) and were selected as optimal, giving the best discriminatory performance for further analysis. The separation of SLE vs HC and SLE-P vs SLE-NP samples were presented by projecting the samples into the subspace constructed of component 1 and component 2. The top weighted features were selected and presented by variable loading plots (50 for SLE vs HC and 28 for SLE-P vs SLE-NP). sPLS-DA results were visualised using the R package ggplot2 [16].

Supplementary Tables

Supplementary Table I. Demographic characteristics of SLE and HC.

		SLE Total (n=80)	HC (n=39)	P-value
Sex n (%)	Female	80 (100)	39 (100)	n/a
	Male	0 (0)	0 (0)	
Age years	Median (IQR)	47.0 (17.75)	32 (25)	0.0007 ¹
Ethnicity n (%)	Asian	9 (11.3)	2 (5.1)	0.0020 ²
	Caucasian	42 (52.5)	34 (87.2)	
	Black/Caribbean	21 (26.3)	2 (5.1)	
	Other	8 (10.0)	1 (2.6)	

Supplementary Table I. Demographic characteristics of SLE and HC. Demographic and clinical characteristics at the time of the first scan were compared between SLE patients and healthy controls (HC). Statistical comparisons were made using ¹Mann-Whitney U, ²Fisher's exact. Age and ethnicity were accounted for during analysis. Abbreviations: IQR – interquartile range.

Supplementary Table II. Cohort Characteristics.

	SLE-NP (n=50)	SLE-P (n=30)	P-value
Sex n (%) Female	50 (100)	30 (100)	n/a
Age years Median (IQR)	39.5 (18.0)	54.0 (11.5)	<0.0001 ²
Ethnicity n (%)			
Asian	7 (14)	2 (6.7)	
Caucasian	24 (48)	18 (60.0)	
Black/Caribbean	14 (28)	7 (23.3)	0.6671 ¹
Other	5 (10)	3 (10)	
Disease Duration (years), Median (IQR)	12 (10)	20 (20.25)	0.0116 ²
BILAG-2004 Median (IQR)	2 (8)	1 (7.25)	0.2201 ²
SLICC damage index Median (IQR)	0 (0-1)	1 (0-2)	0.0192 ²
Mean arterial BP (mmHg) Mean (SD), NR=65-110	91.79 (10.5)	94.7 (10.3)	0.2310 ³
ESR (mm/hr) Median (IQR). NR=0-20	14 (16.25)	14 (33.5)	0.5313 ²
CRP (mg/dl) Median (IQR), NR<5	1.4 (2.8)	2.2 (4.1)	0.2477 ²
Creatinine (umol/L) Median (IQR), NR=50–120	62.5 (53-73)	65.5 (54-78.25)	0.6402 ³
eGFR (ml/min/1.73m²) Median (IQR), NR>90	97.5 (81-131.8)	89.5 (77-107.3)	0.1737 ³
C3 (g/L) Median (IQR), NR=0.9-1.8	0.99 (0.79-1.20)	0.96 (0.87-1.22)	0.4774 ³
dsDNA (IU/mL) Median (IQR), NR<50	32.5 (10-107)	12.5 (8-72.5)	0.4554 ³
Treatments at the time of scan, n (%)			
HCQ	32 (64)	19 (63.3)	
Statins	6 (12)	5 (16.7)	
Ace Inhibitor	21 (42)	10 (33.3)	
Immunosuppressive	26 (52)	11 (36.7)	0.7807 ¹
Rituximab	18 (36)	8 (26.7)	
Aspirin	4 (8)	5 (16.7)	
Prednisolone	37 (74)	16 (53.3)	
No Treatment	2 (4)	2 (6.7)	
Time since last dose of rituximab (years) Mean (range)	3 (1-4.25)	1.5 (0.25-2.75)	0.9952 ²
Number of rituximab cycles* Mean (range)	2 (1.75-3)	2 (1.25-3.5)	0.1545 ²
Prednisolone Dosage (mg per day)	5.0 (7.5)	3.4 (5.0)	0.0794 ²
Serum cytokine levels (pg/ml) Median (IQR)	SLE-NP N=19	SLE-P N=17	
IL-6	12.44 (5.64-22.16)	13.24 (6.66-30.40)	0.546 ²
Interferon-γ	0.0 (0.0-17.0)	0.0 (0.0-11.24)	0.489 ²
TNF-α	0.0 (0.0-8.96)	0.0 (0-14.3)	0.367 ²
IL-10	4.08 (0.84-11.04)	2.72 (0-6.86)	0.367 ²

Supplementary Table II. Cohort Characteristics. Demographic and clinical characteristics at the time of the first scan were compared between SLE patients with plaque (SLE-P) or those with no

plaques (SLE-NP). Statistical comparisons were made using ¹chi-squared, ²Mann-Whitney U, or ³unpaired t-test. Mean arterial pressure (2 x diastolic-BP + systolic-BP divided by 3).

*Rituximab cycles comprise two 1000mg infusions each given two weeks apart per treatment cycle in our centre.

Serum cytokine levels (IL6, IL10, IFN- γ and TNF- α) were measured in a subset of patients (SLE-NP n=19 and SLE-P n=17) from serum taken at the time of the scan by Cytometric Bead Array (CBA, BD Bioscience): according to manufacturer's instructions.

Abbreviations: BILAG, British Isles Lupus Assessment Group disease activity score; BP, blood pressure; C3, complement component 3; CRP, C reactive protein; dsDNA, anti-double stranded DNA antibodies; eGFR, estimated glomerular filtration rate; ESR, erythrocyte sedimentation rate; IL, interleukin; IQR, interquartile range; SLICC, Systemic Lupus International Collaborating Clinics damage index; TNF, tumour necrosis factor.

Supplementary Table III: Metabolite list.

Abbreviation	Full description
AcAce	Acetoacetate
Ace	Acetate
Ala	Alanine
Alb	Albumin
ApoA1	Apolipoprotein A-I
ApoB	Apolipoprotein B
ApoB/ApoA1	Ratio of apolipoprotein B to apolipoprotein A-I
bOHBut	3-hydroxybutyrate
Cit	Citrate
Crea	Creatinine
DHA	22:6, docosahexaenoic acid
DHA/FA	Ratio of 22:6, docosahexaenoic acids to total fatty acids
EstC	Esterified cholesterol
FAw3	Omega-3 fatty acids
FAw3/FA	Ratio of omega-3 fatty acids to total fatty acids
FAw6	Omega-6 fatty acids
FAw6/FA	Ratio of omega-6 fatty acids to total fatty acids
FreeC	Free cholesterol
Glc	Glucose
Gln	Glutamine
Glol	Glycerol
Gly	Glycine
Gp	Glycoprotein acetylation
HDL2-C	Total cholesterol in HDL2
HDL3-C	Total cholesterol in HDL3
HDL-C	Total cholesterol in HDL
HDL-D	Mean diameter for HDL particles
HDL-TG	Triglycerides in HDL
His	Histidine
IDL-C	Total cholesterol in IDL particles
IDL-C_%	Percentage of cholesterol in IDL particles
IDL-CE	Cholesterol esters in IDL particles
IDL-CE_%	Percentage of cholesterol esters in IDL particles
IDL-FC	Free cholesterol in IDL particles
IDL-FC_%	Percentage of free cholesterol in IDL particles
IDL-L	Total lipids in IDL particles
IDL-P	Concentration of IDL particles

IDL-PL	Phospholipids in IDL particles
IDL-PL_%	Percentage of phospholipids in IDL particles
IDL-TG	Triglycerides in IDL particles
IDL-TG_%	Percentage of triglycerides in IDL particles
Ile	Isoleucine
LA	18:2, linoleic acid
LA/FA	Ratio of 18:2, linoleic acids to total fatty acids
Lac	Lactate
LDL-C	Total cholesterol in LDL
LDL-D	Mean diameter for LDL particles
LDL-TG	Triglycerides in LDL
Leu	Leucine
L-HDL-C	Total cholesterol in large HDL particles
L-HDL-C_%	Percentage of cholesterol in large HDL particles
L-HDL-CE	Cholesterol esters in large HDL particles
L-HDL-CE_%	Percentage of cholesterol esters in large HDL particles
L-HDL-FC	Free cholesterol in large HDL particles
L-HDL-FC_%	Percentage of free cholesterol in large HDL particles
L-HDL-L	Total lipids in large HDL particles
L-HDL-P	Concentration of large HDL particles
L-HDL-PL	Phospholipids in large HDL particles
L-HDL-PL_%	Percentage of phospholipids in large HDL particles
L-HDL-TG	Triglycerides in large HDL particles
L-HDL-TG_%	Percentage of triglycerides in large HDL particles
L-LDL-C	Total cholesterol in large LDL particles
L-LDL-C_%	Percentage of cholesterol in large LDL particles
L-LDL-CE	Cholesterol esters in large LDL particles
L-LDL-CE_%	Percentage of cholesterol esters in large LDL particles
L-LDL-FC	Free cholesterol in large LDL particles
L-LDL-FC_%	Percentage of free cholesterol in large LDL particles
L-LDL-L	Total lipids in large LDL particles
L-LDL-P	Concentration of large LDL particles
L-LDL-PL	Phospholipids in large LDL particles
L-LDL-PL_%	Percentage of phospholipids in large LDL particles
L-LDL-TG	Triglycerides in large LDL particles
L-LDL-TG_%	Percentage of triglycerides in large LDL particles
L-VLDL-C	Total cholesterol in large VLDL particles
L-VLDL-C_%	Percentage of cholesterol in large VLDL particles
L-VLDL-CE	Cholesterol esters in large VLDL particles

L-VLDL-CE_%	Percentage of cholesterol esters in large VLDL particles
L-VLDL-FC	Free cholesterol in large VLDL particles
L-VLDL-FC_%	Percentage of free cholesterol in large VLDL particles
L-VLDL-L	Total lipids in large VLDL particles
L-VLDL-P	Concentration of large VLDL particles
L-VLDL-PL	Phospholipids in large VLDL particles
L-VLDL-PL_%	Percentage of phospholipids in large VLDL particles
L-VLDL-TG	Triglycerides in large VLDL particles
L-VLDL-TG_%	Percentage of triglycerides in large VLDL particles
M-HDL-C	Total cholesterol in medium HDL particles
M-HDL-C_%	Percentage of cholesterol in medium HDL particles
M-HDL-CE	Cholesterol esters in medium HDL particles
M-HDL-CE_%	Percentage of cholesterol esters in medium HDL particles
M-HDL-FC	Free cholesterol in medium HDL particles
M-HDL-FC_%	Percentage of free cholesterol in medium HDL particles
M-HDL-L	Total lipids in medium HDL particles
M-HDL-P	Concentration of medium HDL particles
M-HDL-PL	Phospholipids in medium HDL particles
M-HDL-PL_%	Percentage of phospholipids in medium HDL particles
M-HDL-TG	Triglycerides in medium HDL particles
M-HDL-TG_%	Percentage of triglycerides in medium HDL particles
M-LDL-C	Total cholesterol in medium LDL particles
M-LDL-C_%	Percentage of cholesterol in medium LDL particles
M-LDL-CE	Cholesterol esters in medium LDL particles
M-LDL-CE_%	Percentage of cholesterol esters in medium LDL particles
M-LDL-FC	Free cholesterol in medium LDL particles
M-LDL-FC_%	Percentage of free cholesterol in medium LDL particles
M-LDL-L	Total lipids in medium LDL particles
M-LDL-P	Concentration of medium LDL particles
M-LDL-PL	Phospholipids in medium LDL particles
M-LDL-PL_%	Percentage of phospholipids in medium LDL particles
M-LDL-TG	Triglycerides in medium LDL particles
M-LDL-TG_%	Percentage of triglycerides in medium LDL particles
MUFA	Monounsaturated fatty acids
MUFA/FA	Ratio of monounsaturated fatty acids to total fatty acids
M-VLDL-C	Total cholesterol in medium VLDL particles
M-VLDL-C_%	Percentage of cholesterol in medium VLDL particles
M-VLDL-CE	Cholesterol esters in medium VLDL particles
M-VLDL-CE_%	Percentage of cholesterol esters in medium VLDL particles

M-VLDL-FC	Free cholesterol in medium VLDL particles
M-VLDL-FC_%	Percentage of free cholesterol in medium VLDL particles
M-VLDL-L	Total lipids in medium VLDL particles
M-VLDL-P	Concentration of medium VLDL particles
M-VLDL-PL	Phospholipids in medium VLDL particles
M-VLDL-PL_%	Percentage of phospholipids in medium VLDL particles
M-VLDL-TG	Triglycerides in medium VLDL particles
M-VLDL-TG_%	Percentage of triglycerides in medium VLDL particles
PC	Phosphatidylcholine and other cholines
Phe	Phenylalanine
PUFA	Polyunsaturated fatty acids
PUFA/FA	Ratio of polyunsaturated fatty acids to total fatty acids
Remnant-C	Remnant cholesterol (non-HDL, non-LDL cholesterol)
Serum-C	Serum total cholesterol
Serum-TG	Serum total triglycerides
SFA	Saturated fatty acids
SFA/FA	Ratio of saturated fatty acids to total fatty acids
S-HDL-C	Total cholesterol in small HDL particles
S-HDL-C_%	Percentage of cholesterol in small HDL particles
S-HDL-CE	Cholesterol esters in small HDL particles
S-HDL-CE_%	Percentage of cholesterol esters in small HDL particles
S-HDL-FC	Free cholesterol in small HDL particles
S-HDL-FC_%	Percentage of free cholesterol in small HDL particles
S-HDL-L	Total lipids in small HDL particles
S-HDL-P	Concentration of small HDL particles
S-HDL-PL	Phospholipids in small HDL particles
S-HDL-PL_%	Percentage of phospholipids in small HDL particles
S-HDL-TG	Triglycerides in small HDL particles
S-HDL-TG_%	Percentage of triglycerides in small HDL particles
S-LDL-C	Total cholesterol in small LDL particles
S-LDL-C_%	Percentage of cholesterol in small LDL particles
S-LDL-CE	Cholesterol esters in small LDL particles
S-LDL-CE_%	Percentage of cholesterol esters in small LDL particles
S-LDL-FC	Free cholesterol in small LDL particles
S-LDL-FC_%	Percentage of free cholesterol in small LDL particles
S-LDL-L	Total lipids in small LDL particles
S-LDL-P	Concentration of small LDL particles
S-LDL-PL	Phospholipids in small LDL particles
S-LDL-PL_%	Percentage of phospholipids in small LDL particles

S-LDL-TG	Triglycerides in small LDL particles
S-LDL-TG_%	Percentage of triglycerides in small LDL particles
SM	Sphingomyelins
S-VLDL-C	Total cholesterol in small VLDL particles
S-VLDL-C_%	Percentage of cholesterol in small VLDL particles
S-VLDL-CE	Cholesterol esters in small VLDL particles
S-VLDL-CE_%	Percentage of cholesterol esters in small VLDL particles
S-VLDL-FC	Free cholesterol in small VLDL particles
S-VLDL-FC_%	Percentage of free cholesterol in small VLDL particles
S-VLDL-L	Total lipids in small VLDL particles
S-VLDL-P	Concentration of small VLDL particles
S-VLDL-PL	Phospholipids in small VLDL particles
S-VLDL-PL_%	Percentage of phospholipids in small VLDL particles
S-VLDL-TG	Triglycerides in small VLDL particles
S-VLDL-TG_%	Percentage of triglycerides in small VLDL particles
TG/PG	Ratio of triglycerides to phosphoglycerides
TotCho	Total cholines
TotFA	Total fatty acids
TotPG	Total phosphoglycerides
Tyr	Tyrosine
UnSat	Estimated degree of unsaturation
Val	Valine
VLDL-C	Total cholesterol in VLDL
VLDL-D	Mean diameter for VLDL particles
VLDL-TG	Triglycerides in VLDL
XL-HDL-C	Total cholesterol in very large HDL particles
XL-HDL-C_%	Percentage of cholesterol in very large HDL particles
XL-HDL-CE	Cholesterol esters in very large HDL particles
XL-HDL-CE_%	Percentage of cholesterol esters in very large HDL particles
XL-HDL-FC	Free cholesterol in very large HDL particles
XL-HDL-FC_%	Percentage of free cholesterol in very large HDL particles
XL-HDL-L	Total lipids in very large HDL particles
XL-HDL-P	Concentration of very large HDL particles
XL-HDL-PL	Phospholipids in very large HDL particles
XL-HDL-PL_%	Percentage of phospholipids in very large HDL particles
XL-HDL-TG	Triglycerides in very large HDL particles
XL-HDL-TG_%	Percentage of triglycerides in very large HDL particles
XL-VLDL-C	Total cholesterol in very large VLDL particles
XL-VLDL-C_%	Percentage of cholesterol in very large VLDL particles

XL-VLDL-CE	Cholesterol esters in very large VLDL particles
XL-VLDL-CE_%	Percentage of cholesterol esters in very large VLDL particles
XL-VLDL-FC	Free cholesterol in very large VLDL particles
XL-VLDL-FC_%	Percentage of free cholesterol in very large VLDL particles
XL-VLDL-L	Total lipids in very large VLDL particles
XL-VLDL-P	Concentration of very large VLDL particles
XL-VLDL-PL	Phospholipids in very large VLDL particles
XL-VLDL-PL_%	Percentage of phospholipids in very large VLDL particles
XL-VLDL-TG	Triglycerides in very large VLDL particles
XL-VLDL-TG_%	Percentage of triglycerides in very large VLDL particles
XS-VLDL-C	Total cholesterol in very small VLDL particles
XS-VLDL-C_%	Percentage of cholesterol in very small VLDL particles
XS-VLDL-CE	Cholesterol esters in very small VLDL particles
XS-VLDL-CE_%	Percentage of cholesterol esters in very small VLDL particles
XS-VLDL-FC	Free cholesterol in very small VLDL particles
XS-VLDL-FC_%	Percentage of free cholesterol in very small VLDL particles
XS-VLDL-L	Total lipids in very small VLDL particles
XS-VLDL-P	Concentration of very small VLDL particles
XS-VLDL-PL	Phospholipids in very small VLDL particles
XS-VLDL-PL_%	Percentage of phospholipids in very small VLDL particles
XS-VLDL-TG	Triglycerides in very small VLDL particles
XS-VLDL-TG_%	Percentage of triglycerides in very small VLDL particles
XXL-VLDL-C	Total cholesterol in chylomicrons and extremely large VLDL particles
XXL-VLDL-C_%	% of cholesterol in chylomicrons and extremely large VLDL particles
XXL-VLDL-CE	Cholesterol esters in chylomicrons and extremely large VLDL particles
XXL-VLDL-CE_%	Percentage of cholesterol esters in chylomicrons and extremely large VLDL particles
XXL-VLDL-FC	Free cholesterol in chylomicrons and extremely large VLDL particles
XXL-VLDL-FC_%	Percentage of free cholesterol in chylomicrons and extremely large VLDL particles
XXL-VLDL-L	Total lipids in chylomicrons and extremely large VLDL particles
XXL-VLDL-P	Concentration of chylomicrons and extremely large VLDL particles
XXL-VLDL-PL	Phospholipids in chylomicrons and extremely large VLDL particles
XXL-VLDL-PL_%	% phospholipids in chylomicrons and extremely large VLDL particles
XXL-VLDL-TG	Triglycerides in chylomicrons and extremely large VLDL particles
XXL-VLDL-TG_%	% of triglycerides in chylomicrons and extremely large VLDL particles

Supplementary Table III: Metabolite list. Metabolites (n=228) analysed by Nightingale Health (see methods) used for analysis, including full names and abbreviations used throughout the main text. VLDL – very low density lipoprotein, IDL – intermediate density lipoprotein, LDL – low density lipoprotein and HDL – high density lipoprotein.

Supplementary Table IV: Clinical lipid profile of SLE patients.

		SLE Patients (n=80)	Normal Range
Total Cholesterol	Mean (95% CI)	4.77 (4.54–5.01)	< 5
Triglycerides	Median (Quartiles)	1.00 (0.70–1.40)	< 3
HDL Concentration	Mean (95% CI)	1.65 (1.53–1.76)	> 1
LDL Concentration	Mean (95% CI)	2.61 (2.41–2.81)	< 3
Cholesterol : HDL Ratio	Mean (95% CI)	3.10 (2.90–3.29)	< 4

Supplementary Table IV: Clinical lipid profile of SLE patients. Clinical lipid profile of SLE patients were compared to the normal range. Tests for normality were done using the D'Agostino and Pearson test. Where the lipid measurements were normally distributed, the mean and 95% confidence intervals (CI) were reported, otherwise the median and lower & upper quartiles (Quartiles) were reported.

Supplementary Table V. Predictors for lasso logistic regression with interactions.

Features Included	Beta Coefficient
Intercept	-1.17411
S-LDL-CE : Rituximab	-0.68492
S-HDL-FC : XXL-VLDL-CE_%	-0.49192
Gly : His	-0.46405
Statins : ACE Inhibitor	-0.37603
Lac : Cit	-0.37161
HCQ : ACE Inhibitor	-0.23705
XL-HDL-TG : Glol	-0.15531
IDL-CE : Glol	-0.12971
His : XS-VLDL-CE_%	-0.09162
L-VLDL-TG : Lac	-0.08858
Glol : XXL-VLDL-CE_%	-0.07212
Gly : XS-VLDL-CE_%	-0.05998
Glol : His	-0.03934
DHA/FA : Age	0.019318
DHA/FA : Tyr	0.019945
L-HDL-PL_% : Age	0.022831
L-VLDL-TG : HCQ	0.034663
Ethnicity (Caucasian) : Disease Duration	0.035969
XXL-VLDL-PL_% : Mean Arterial BP	0.076793
DHA : Aspirin	0.089295
DHA : HCQ	0.096398
HCQ : Prednisolone Dose	0.129875
Age : Mean arterial BP	0.130915
SFA/FA : M-VLDL-FC_%	0.193219
Tyr : Disease Duration	0.194325
M-VLDL-FC_% : L-LDL-FC_%	0.239355
HCQ1 : Rituximab	0.282773
IDL-TG_% : S-LDL-C_%	0.292232
SM : Tyr	0.333199
Leu : Mean arterial BP	0.401734
Gln : Ace	0.464306
SM : Leu	0.597411

Ethnicity (Black / Caribbean) : HCQ	0.755513
L-VLDL-FC_% : Age	1.056669
IDL-TG_% : Age	1.133333

Supplementary Table V. Predictors for lasso logistic regression with interactions. List of features selected by the lasso regressions with interactions. Interacting features are separated by a colon (:). Interactions are listed in order of the absolute size of the beta coefficient, where the top ten metabolite interactions will have the largest effect on the model classification. Where predictors are categorical (ethnicity) the specific category is shown in brackets. Abbreviations: BP, blood pressure; HCQ, hydroxychloroquine. See [Supplementary Table III](#) for list of metabolite abbreviations.

Supplementary Table VI: Significant correlations between metabolites and clinical scores.

	Disease Duration	GSM	BILAG	Plaque Number	TPT	TPA
L.VLDL.TG:Lac	NA	NA	NA	NA	NA	0.37101844
IDL.CE:GloI	NA	0.4110131	NA	NA	NA	NA
XL.HDL.TG:GloI	NA	0.4388473	NA	NA	NA	NA
DHA.FA:Age	NA	NA	-0.38637	NA	NA	NA
Gly:XS.VLDL.CE_.	NA	-0.44402	NA	NA	NA	NA
His:XS.VLDL.CE_.	NA	-0.398371	NA	-0.3772877	-0.392239	NA
Tyr:Dis_Dur	0.91297079	NA	NA	NA	NA	NA
M.VLDL.FC_.:L.LDL.FC_.	NA	NA	0.369133	NA	NA	NA
LDL-D	NA	-0.4386889	NA	NA	NA	NA
XS-VLDL-TG	NA	0.4788697	NA	NA	NA	NA
S-HDL-TG	NA	0.49307917	NA	NA	NA	NA
M-VLDL-FC_%	NA	NA	0.3839743	NA	NA	NA

Supplementary Table VI: Significant correlations between metabolites and clinical scores.

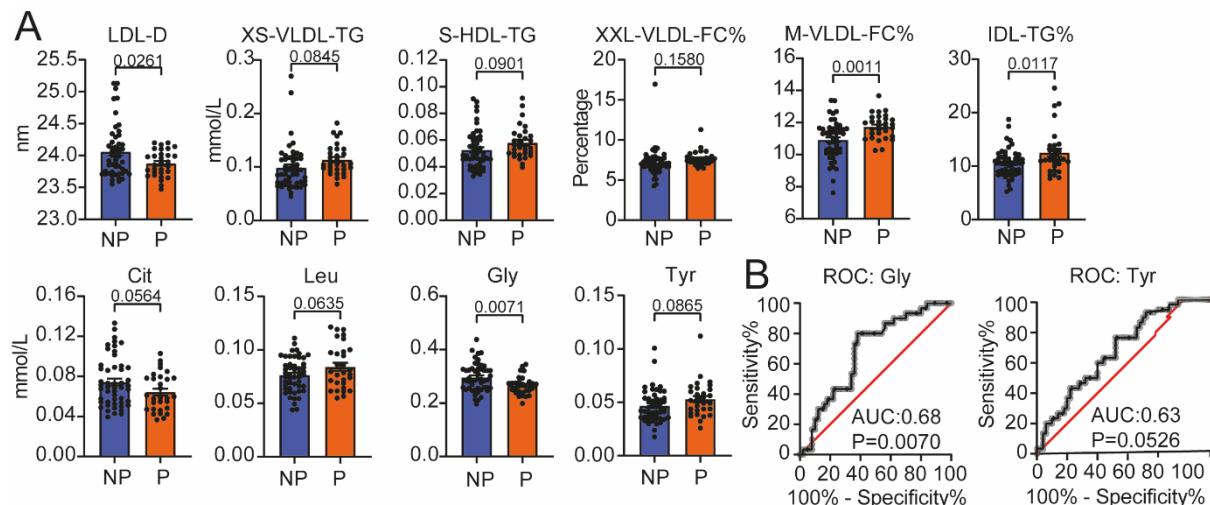
Correlations between metabolites and patient clinical characteristics ([Figure 4](#)). Pearson's product moment correlation coefficients and p values in brackets. Only correlations with p value below 0.05 are shown. GSM, grey scale median, BILAG, British Isles Lupus Group Index disease activity score, TPT, total plaque thickness, TPA, Total Plaque area, NA, not applicable (see [Supplementary Table III](#) for metabolite abbreviations).

Supplementary Figures

		Logistic Regression			Logistic Regression + Interactions				
		Predicted			Predicted				
Actual		NP	P	Σ	NP	P	Σ		
	NP	43	7	50	NP	44	6	50	
	P	13	17	30	P	10	20	30	
Σ		56	24	80	Σ		54	26	80

		SVM			Random Forest			Decision Tree						
		Predicted			Predicted			Predicted						
Actual		NP	P	Σ	NP	P	Σ	NP	P	Σ				
	NP	44	6	50	NP	45	5	50	NP	37	13	50		
	P	16	14	30	P	17	13	30	P	15	15	30		
Σ		60	20	80	Σ		62	18	80	Σ		52	28	80

Supplementary Figure I: Confusion Matrices. The confusion matrix shows the number of correct (blue squares) and incorrect (orange squares) classifications for each model. The sum (Σ) of each row and column is given. The algorithms used were lasso logistic regression (LR), logistic regression with interactions (LR+I), support vector machine (SVM), random forest (RF), and decision tree (Tree). Performance metrics in [Table 1](#) for all five models are based on these confusion matrices.



Supplementary Figure II. Identification of important metabolites separating patients with SLE-P from SLE-NP. **(A)** Bar graphs showing metabolite levels between SLE-P and SLE-NP for each metabolite identified by univariate logistic regression, random forest (RF), Lasso logistic regression (LR) analysis ([see Figure 3A-C](#)). Mean, T tests and p values shown. **(B)**. Metabolites which featured in the top ten of LR and RF were further analysed in a ROC plot ([See Figure 3B-D](#)).

References

1. Griffin, M., et al., *Carotid and femoral arterial wall changes and the prevalence of clinical cardiovascular disease*. Vascular medicine (London, England), 2009. **14**(3): p. 227-232.
2. Nicolaides, A.N., et al., *Asymptomatic internal carotid artery stenosis and cerebrovascular risk stratification*. J Vasc Surg, 2010. **52**(6): p. 1486-1496.e1-5.
3. Ebrahim, S., et al., *Carotid plaque, intima media thickness, cardiovascular risk factors, and prevalent cardiovascular disease in men and women: the British Regional Heart Study*. Stroke, 1999. **30**(4): p. 841-50.
4. Griffin, M., et al., *Cardiovascular risk assessment using ultrasound: the value of arterial wall changes including the presence, severity and character of plaques*. Pathophysiology of haemostasis and thrombosis, 2002. **32**(5-6): p. 367-370.
5. Ahola-Olli, A.V., et al., *Circulating metabolites and the risk of type 2 diabetes: a prospective study of 11,896 young adults from four Finnish cohorts*. Diabetologia, 2019. **62**(12): p. 2298-2309.
6. Kalimeri, M. & Qin, W. forestplotNMR: Create Custom Forestplots for NMR Metabolomics Data. (2019)
7. RStudioTeam. (2016). RStudio: integrated development for R. Retrieved from <http://www.rstudio.com/>
8. Demšar, J., et al., *Orange: data mining toolbox in python*. J. Mach. Learn. Res., 2013. **14**(1): p. 2349–2353.
9. Waddington, K.E., et al., *LXR alters CD4⁺ T cell function through direct regulation of glycosphingolipid synthesis*. bioRxiv, 2020: p. 721050.
10. Yu, W., et al., *Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes*. BMC medical informatics and decision making, 2010. **10**: p. 16-16.
11. Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F. e1071: Misc Functions of the Department of Statistics, Probablility Theory Group (formerly: E1071), TU Wien. R package version 1.7-3 (2019). Available online at: <https://cran.rproject.org/web/packages/e1071/e1071.pdf> (accessed Feb 26, 2020)
12. Kingsford, C. and S.L. Salzberg, *What are decision trees?* Nat Biotechnol, 2008. **26**(9): p. 1011-3.
13. Ho, T.K., *Random decision forests*, in *Proceedings of the Third International Conference on Document Analysis and Recognition (Volume 1) - Volume 1*. 1995, IEEE Computer Society. p. 278.
14. Kuhn, M., Wing, J., Weston, St., Williams, A., Jeefer, C., Engelhardt, A et al. R Core Team. (2019). caret: Classification and Regression Training. Retrieved from <https://cran.r-project.org/package=caret>
15. Rohart, F., et al., *mixOmics: An R package for 'omics feature selection and multiple data integration*. PLoS Comput Biol, 2017. **13**(11): p. e1005752.
16. Wickham, H., *ggplot2: elegant graphics for data analysis*. 2016: springer.

1 Supplementary Data Files

Supplementary Data File I

Forest plot visualization of the results of the logistic regression for individual metabolites in SLE vs HC. P-values <0.05 are given in coloured in circles.

Supplementary Data File II

Results of univariate logistic regression for individual metabolites in SLE vs HC, accompanying Supplementary Data File I. P-values <0.05 are given in red text.

Supplementary Data File III

Forest plot visualization of the results of the logistic regression for individual metabolites in SLE-P vs SLE-NP. P-values <0.05 are given in coloured in circles.

Supplementary Data File IV

Results of logistic regression for individual metabolites in SLE-P vs SLE-NP, accompanying

Supplementary Data File V

Lists of predictors for the support vector machine (SVM), random forest (RF), and logistic regression (LR) models. These lists include both metabolic and clinical/demographic predictors. For LR models negative beta-coefficients indicate a reduced likelihood of plaque development.

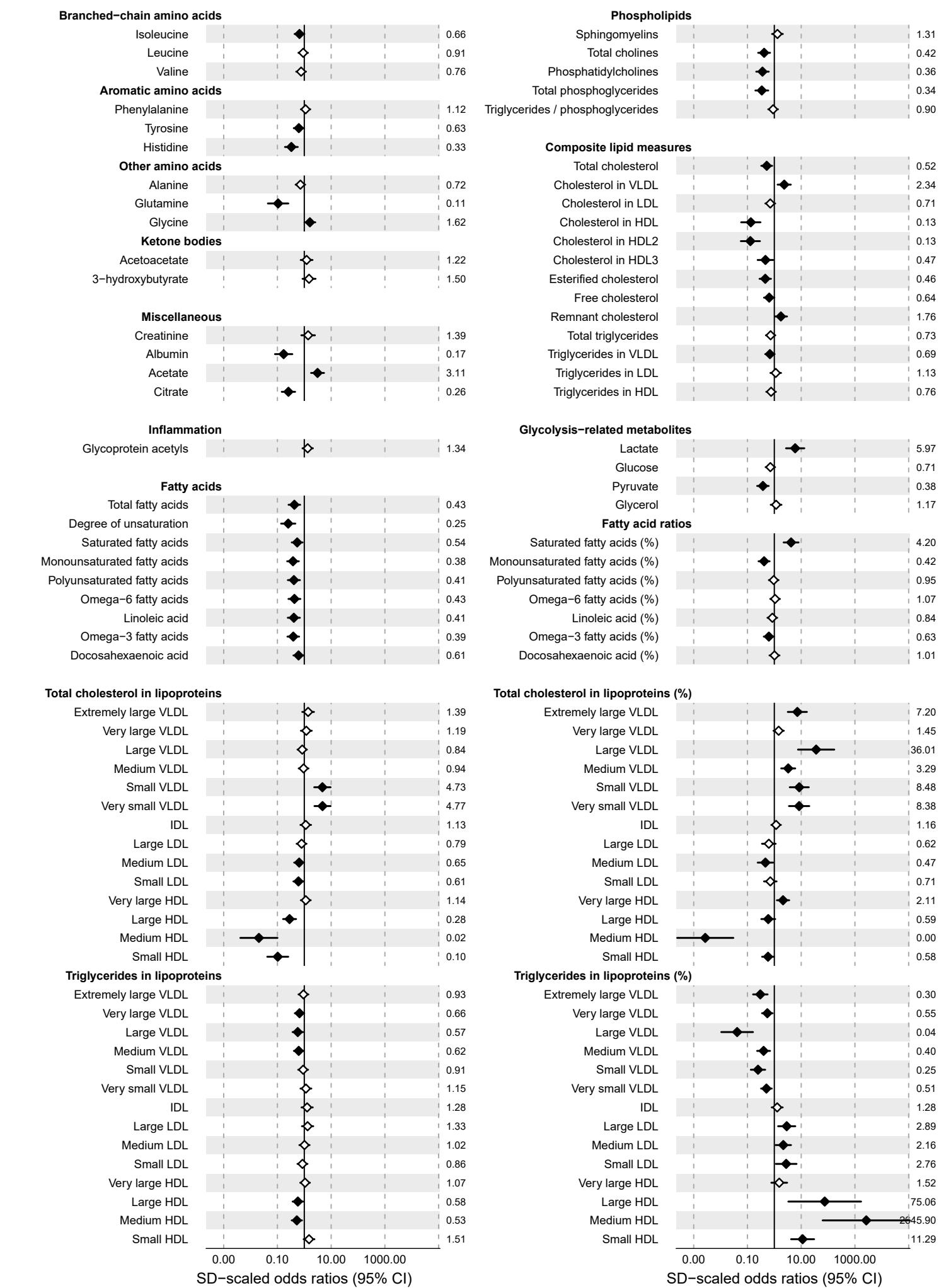
Supplementary Data File VI

Forest plot visualization of the results of the logistic regression for individual metabolites in SLE vs Age. P-values <0.05 are given in coloured in circles.

Supplementary Data File I

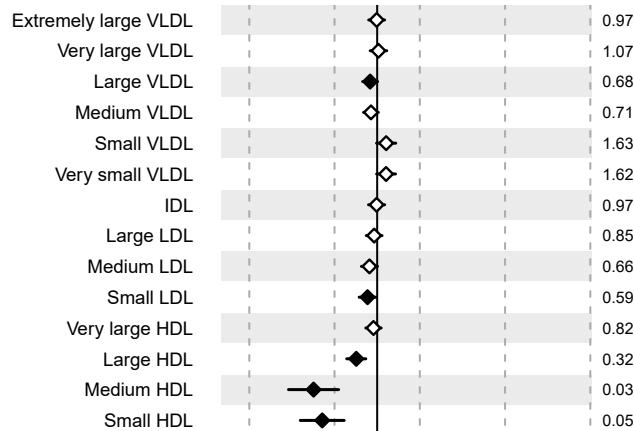
Forest plot visualization of the results of the logistic regression for individual metabolites in SLE vs HC.

P-values <0.05 are given in coloured in circles. SLE vs Healthy control

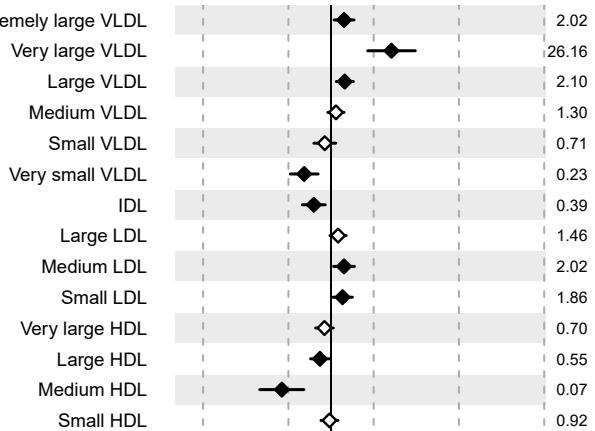


SLE vs Healthy control

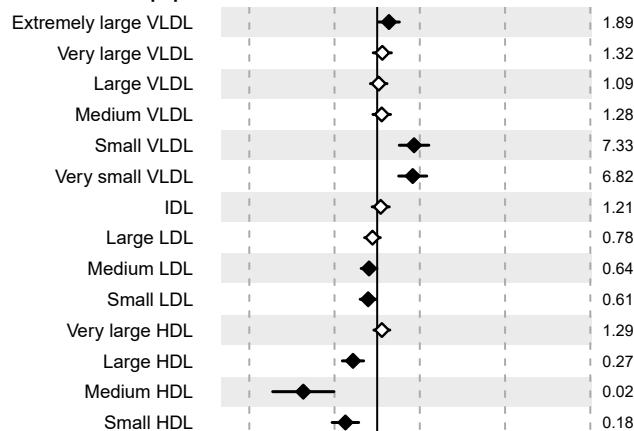
Free cholesterol in lipoproteins



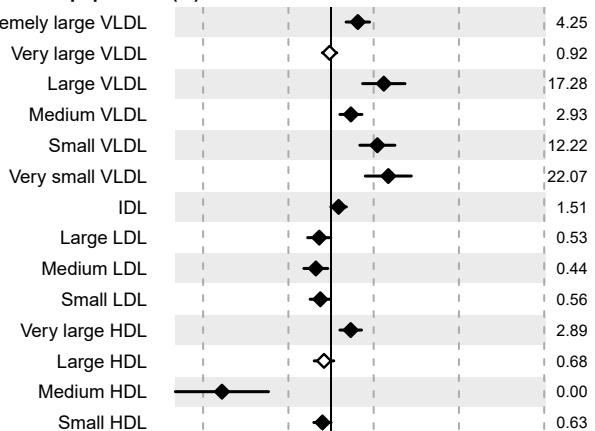
Free cholesterol in lipoproteins (%)



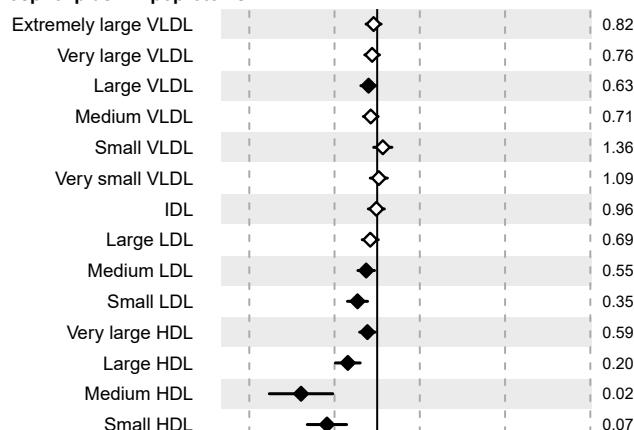
Cholesterol esters in lipoproteins



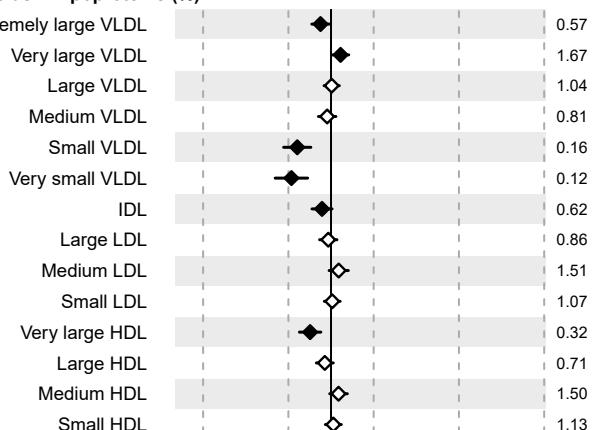
Cholesterol esters in lipoproteins (%)



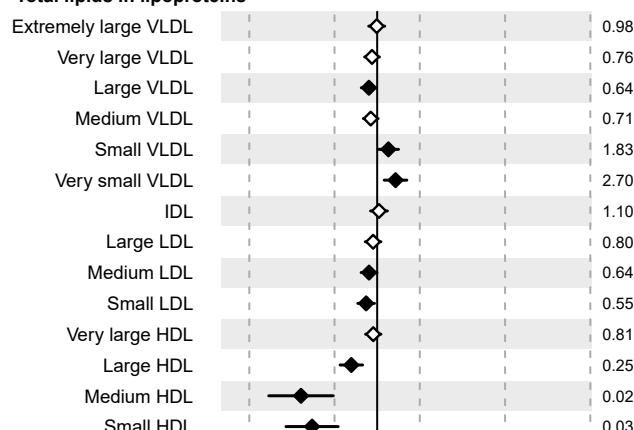
Phospholipids in lipoproteins



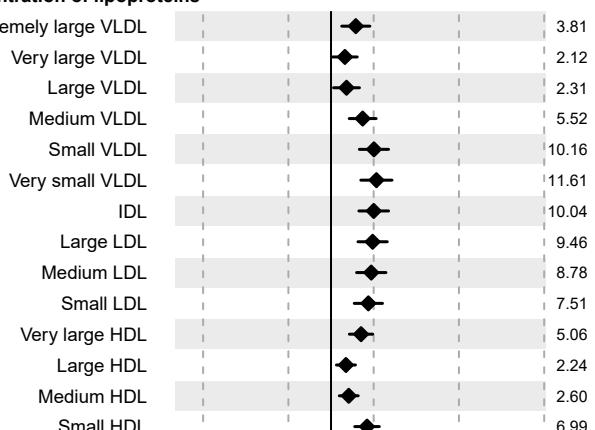
Phospholipids in lipoproteins (%)



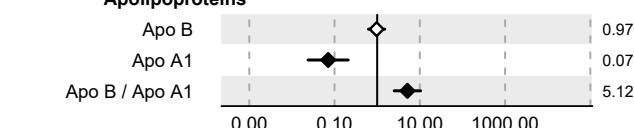
Total lipids in lipoproteins



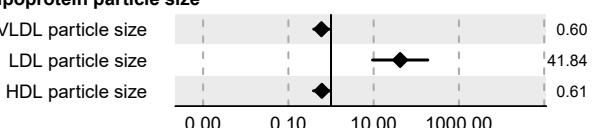
Concentration of lipoproteins



Apolipoproteins



Lipoprotein particle size



SD-scaled odds ratios (95% CI)

SD-scaled odds ratios (95% CI)

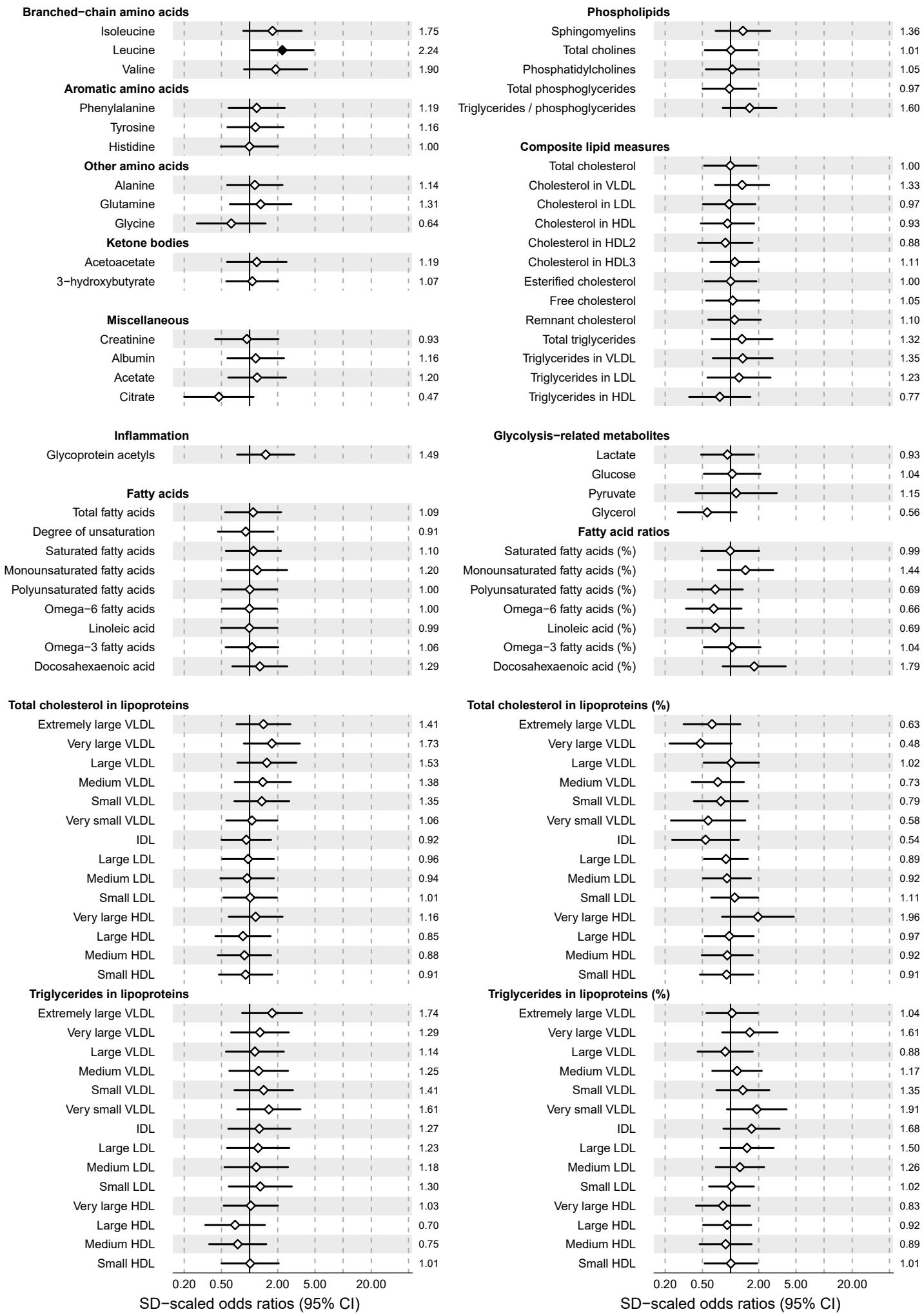
◆ SLEvsHC

L-HDL-CE_%	-0.378884	0.269815	0.684625	1.309722	0.118255	S-VLDL-PL	0.0429	0.005925	0.539565	0.068599
XL-HDL-TG_%	0.421271	0.345743	1.523898	1.413039	0.141149	S-VLDL-PL_%	0.055046	0.004576	0.744191	0.033022
XL-VLDL-L	-0.279180	0.203318	0.756404	1.225462	0.153261	S-VLDL-TG	0.049775	0.001362	0.584428	0.047005
HDL-TG	-0.277141	0.210291	0.757947	1.234037	0.165244	S-VLDL-TG_%	0.04628	0.008427	0.491466	0.127999
S-LDL-C_%	-0.343679	0.275338	0.709157	1.316976	0.165520	Serum-C	0.06443	9.83E-05	0.725628	0.022277
XL-VLDL-PL	-0.271536	0.204074	0.762208	1.226389	0.167792	Serum-TG	0.055375	0.000513	0.605138	0.042462
XXL-VLDL-C	0.328668	0.255778	1.389116	1.291466	0.173138	SFA	0.063628	0.00012	0.604421	0.047724
Gp	0.292688	0.227134	1.340025	1.254999	0.178455	SFA/FA	0.0541	0.002282	1.178707	0.001888
Val	-0.270879	0.213341	0.762709	1.237806	0.192164	SM	0.042242	0.008136	0.579883	0.049724
M-VLDL-FC_%	0.266159	0.223989	1.304942	1.251057	0.204393	TG/PG	0.049671	0.001323	0.580792	0.048748
S-VLDL-PL	0.304531	0.251743	1.355989	1.286265	0.208812	TotCho	0.068958	4.99E-05	0.66481	0.036697
L-LDL-TG	0.283536	0.246061	1.327817	1.278978	0.225951	TotFA	0.070263	3.70E-05	0.71264	0.027875
L-LDL-CE	-0.253296	0.215773	0.776238	1.240820	0.232462	TotPG	0.07595	2.03E-05	0.695985	0.035236
Crea	0.330345	0.296987	1.391448	1.345797	0.235059	Tyr	0.056159	0.000311	0.595708	0.049654
XL-HDL-CE	0.254231	0.219960	1.289470	1.246026	0.235786	UnSat	0.050517	0.003499	0.954668	0.007155
XL-VLDL-CE	0.278377	0.250054	1.320984	1.284094	0.240566	Val	0.05065	0.001003	0.626587	0.034361
SM	0.266581	0.235909	1.305493	1.266060	0.241841	VLDL-C	0.034786	0.029917	0.550744	0.065546
S-VLDL-FC_%	-0.338418	0.301786	0.712897	1.352272	0.268963	VLDL-D	0.052877	0.000783	0.536587	0.077529
L-LDL-C	-0.231401	0.215994	0.793421	1.241096	0.276720	VLDL-TG	0.056137	0.000432	0.591931	0.047816
IDL-TG_%	0.246889	0.240031	1.280037	1.271289	0.279290	XL-HDL-C	0.046372	0.002883	0.571033	0.051572
M-VLDL-CE	0.246595	0.2424404	1.279660	1.274309	0.291921	XL-HDL-C_%	0.04181	0.008031	0.679141	0.029938
IDL-TG	0.244234	0.245545	1.276643	1.278318	0.303070	XL-HDL-CE	0.044369	0.004619	0.5664	0.054517
L-LDL-L	-0.219151	0.216434	0.803200	1.241641	0.304556	XL-HDL-CE_%	0.040908	0.011563	0.742462	0.023305
XL-HDL-L	-0.210984	0.211348	0.809787	1.235342	0.311300	XL-HDL-FC	0.051049	0.000951	0.589428	0.046799
XL-HDL-FC	-0.199881	0.212653	0.818828	1.236955	0.339626	XL-HDL-FC_%	0.050941	0.000891	0.57495	0.054369
XXL-VLDL-PL	-0.193442	0.208342	0.824118	1.231634	0.340667	XL-HDL-L	0.051152	0.000915	0.595644	0.044298
M-VLDL-PL_%	-0.208279	0.241181	0.811980	1.272751	0.378607	XL-HDL-P	0.048307	0.006738	0.556824	0.1223
L-VLDL-C	-0.177364	0.206680	0.837475	1.229589	0.380157	XL-HDL-PL	0.054587	0.000472	0.627467	0.038825
AcAce	0.200290	0.258628	1.221757	1.295152	0.397246	XL-HDL-PL_%	0.038875	0.017211	0.733271	0.022843
IDL-CE	0.191132	0.235812	1.210619	1.265936	0.407112	XL-HDL-TG	0.047402	0.002383	0.568408	0.052536
LA/FA	-0.169574	0.213525	0.844024	1.238035	0.416675	XL-HDL-TG_%	0.046012	0.00257	0.578994	0.048004
XL-VLDL-C	0.175301	0.239343	1.191605	1.270414	0.445783	XL-VLDL-C	0.045052	0.003704	0.573213	0.049839
L-LDL-FC	-0.161617	0.217166	0.850767	1.242550	0.451266	XL-VLDL-C_%	0.052621	0.000649	0.561397	0.061317
IDL-C_%	0.152100	0.208767	1.164276	1.232158	0.462852	XL-VLDL-CE	0.044071	0.00422	0.578543	0.047881
S-LDL-TG	-0.155385	0.219462	0.856085	1.245407	0.472684	XL-VLDL-CE_%	0.047182	0.00242	0.582785	0.04685
GloI	0.152840	0.234755	1.165138	1.264598	0.505296	XL-VLDL-FC	0.046716	0.002864	0.573922	0.049997
L-LDL-PL_%	-0.148065	0.239175	0.862375	1.270201	0.532987	XL-VLDL-FC_%	0.032995	0.135268	0.932847	0.058086
XS-VLDL-TG	0.140050	0.238286	1.150331	1.269071	0.548224	XL-VLDL-L	0.054384	0.000616	0.579707	0.052428
XL-HDL-C	0.127371	0.216819	1.135838	1.242119	0.549256	XL-VLDL-P	0.042761	0.005873	0.543503	0.067499
S-HDL-PL_%	0.124678	0.218793	1.132783	1.244573	0.561526	XL-VLDL-PL	0.054473	0.000623	0.579018	0.052483
IDL-C	0.124408	0.232516	1.132478	1.261770	0.585697	XL-VLDL-PL_%	0.037318	0.023766	0.669038	0.025924
Phe	0.109412	0.207246	1.115622	1.230285	0.587457	XL-VLDL-TG	0.057302	0.000346	0.573264	0.057971
LDL-TG	0.119646	0.231777	1.127098	1.260838	0.595183	XL-VLDL-TG_%	0.05125	0.001033	0.583356	0.055258
TG/PG	-0.103459	0.212979	0.901713	1.237359	0.622309	XS-VLDL-C	0.033201	0.055621	0.587039	0.06945
Leu	-0.092923	0.209072	0.911264	1.232534	0.647598	XS-VLDL-C_%	0.065824	0.000714	0.709003	0.052403
S-VLDL-TG	-0.097340	0.216137	0.907247	1.241273	0.647952	XS-VLDL-CE	0.03493	0.055985	0.622713	0.066979
IDL-L	0.097356	0.231819	1.102253	1.260891	0.668916	XS-VLDL-CE_%	0.088601	8.21E-05	0.894415	0.029416
XL-VLDL-CE_%	-0.078587	0.200912	0.924422	1.222517	0.681190	XS-VLDL-FC	0.040912	0.008823	0.539884	0.068383
XS-VLDL-PL	0.088753	0.236320	1.092811	1.266579	0.702683	XS-VLDL-FC_%	0.059716	0.000393	0.674823	0.039153
S-HDL-FC_%	-0.084789	0.233411	0.918706	1.262900	0.708478	XS-VLDL-L	0.033573	0.038682	0.557166	0.066186
L-VLDL-CE	0.082519	0.228432	1.086019	1.256628	0.710966	XS-VLDL-P	0.05004	0.029281	0.643177	0.113836
XXL-VLDL-TG	-0.071198	0.216519	0.931278	1.241746	0.736424	XS-VLDL-PL	0.046642	0.002922	0.554573	0.059655
XL-HDL-TG	0.067324	0.211883	1.069642	1.236003	0.739474	XS-VLDL-PL_%	0.074056	0.000159	0.824409	0.028786
FAw6/FA	0.064840	0.211213	1.066988	1.235176	0.754703	XS-VLDL-TG	0.04564	0.003509	0.557629	0.05798
XL-VLDL-FC	0.069151	0.228349	1.071598	1.256523	0.755231	XS-VLDL-TG_%	0.053618	0.000761	0.542967	0.070353
S-LDL-PL_%	0.064851	0.217652	1.067000	1.243154	0.758226	XXL-VLDL-C	0.042895	0.005719	0.563374	0.054297
M-VLDL-C	-0.060852	0.215241	0.940963	1.240161	0.774044	XXL-VLDL-C_%	0.04868	0.005644	0.594249	0.076953
PUFA/FA	-0.051538	0.212672	0.949768	1.236979	0.804894	XXL-VLDL-CE	0.040364	0.009343	0.569152	0.053507
IDL-PL	-0.044805	0.220647	0.956183	1.246883	0.837194	XXL-VLDL-CE_%	0.056838	0.000824	0.541014	0.099765
L-VLDL-PL_%	0.034592	0.210497	1.035197	1.234291	0.865892	XXL-VLDL-FC	0.048668	0.001976	0.578965	0.048744
IDL-FC	-0.031748	0.222779	0.968751	1.249545	0.885335	XXL-VLDL-FC_%	0.03925	0.017251	0.55521	0.065283
ApoB	-0.031053	0.224579	0.969424	1.251796	0.889062	XXL-VLDL-L	0.048526	0.001982	0.578054	0.049042
XXL-VLDL-FC	-0.027474	0.219567	0.972900	1.245537	0.898087	XXL-VLDL-P	0.039938	0.012996	0.566976	0.065595
XXL-VLDL-L	-0.023570	0.220801	0.976706	1.247075	0.912918	XXL-VLDL-PL	0.052673	0.000923	0.578154	0.05094
M-LDL-TG	0.018018	0.224106	1.018181	1.251203	0.934636	XXL-VLDL-PL_%	0.061236	0.000151	0.598956	0.052598
DHA/FA	0.013792	0.218648	1.013887	1.244393	0.948835	XXL-VLDL-TG	0.049536	0.00163	0.577289	0.04985
						XXL-VLDL-TG_%	0.037633	0.024422	0.599861	0.055081

Supplementary Data File III

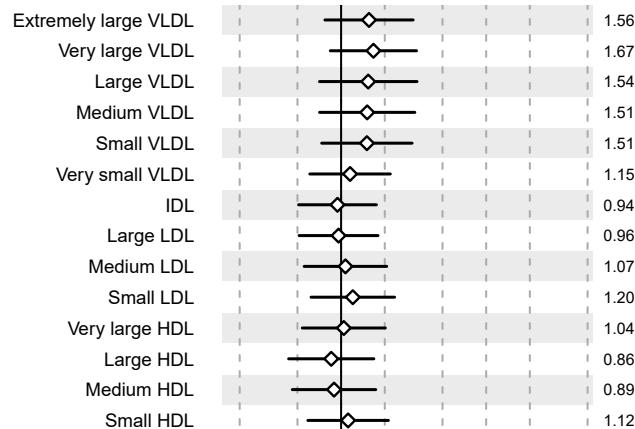
Forest plot visualization of the results of the logistic regression for individual metabolites in SLE-P vs SLE-NP. P-values <0.05 are given in circles.

SLE Plaque vs No Plaque

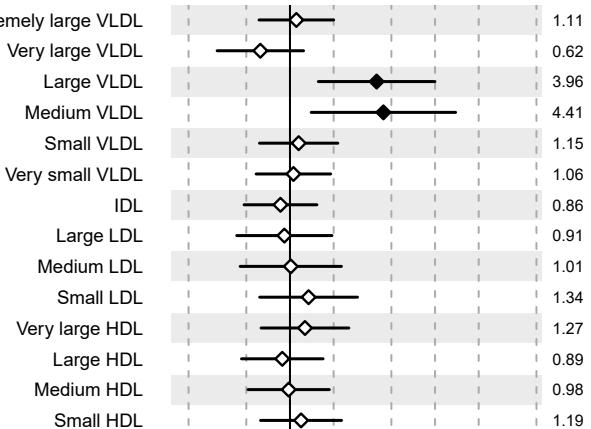


SLE Plaque vs No Plaque

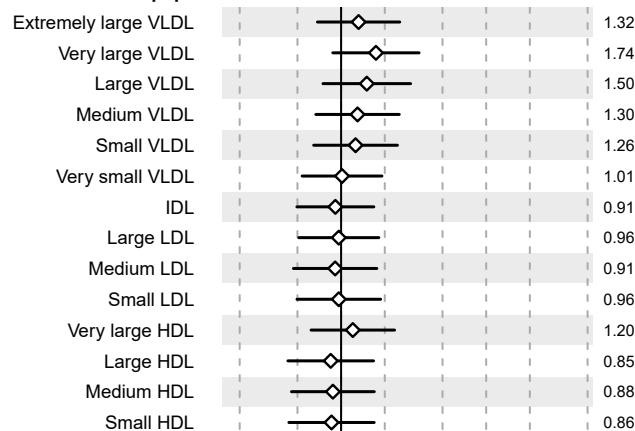
Free cholesterol in lipoproteins



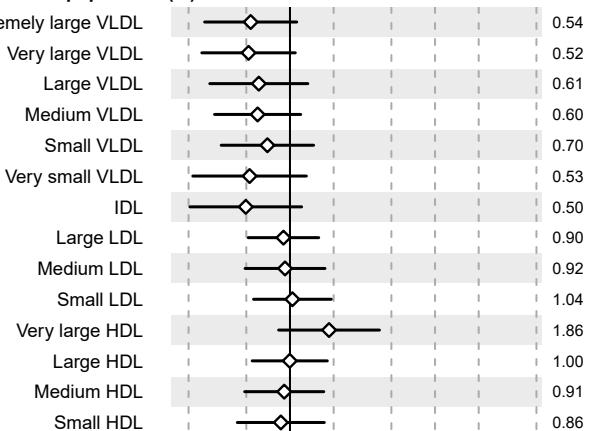
Free cholesterol in lipoproteins (%)



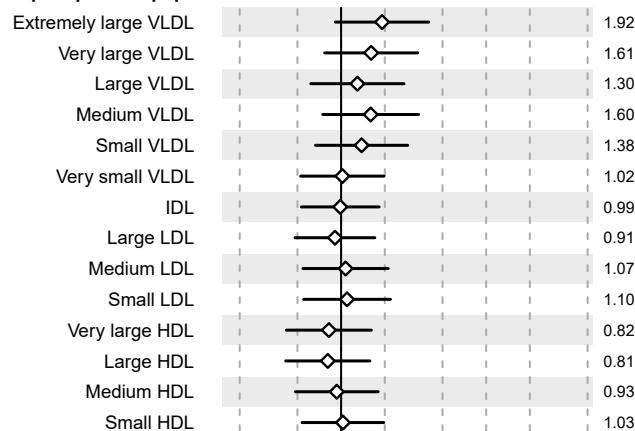
Cholesterol esters in lipoproteins



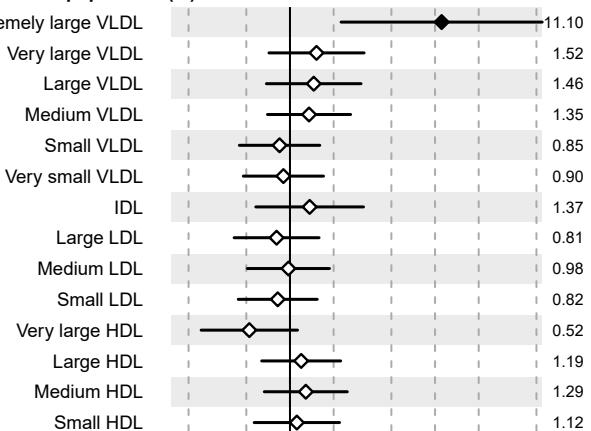
Cholesterol esters in lipoproteins (%)



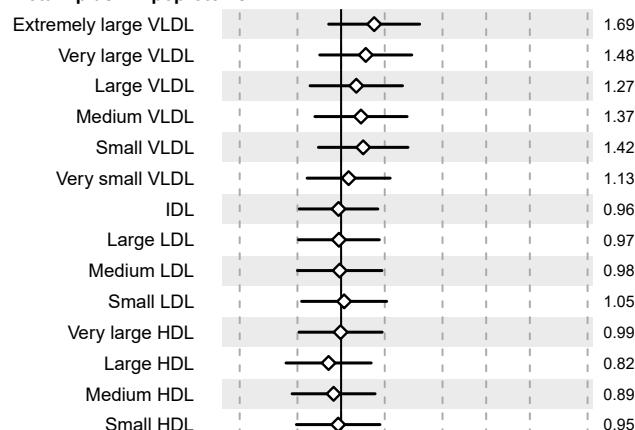
Phospholipids in lipoproteins



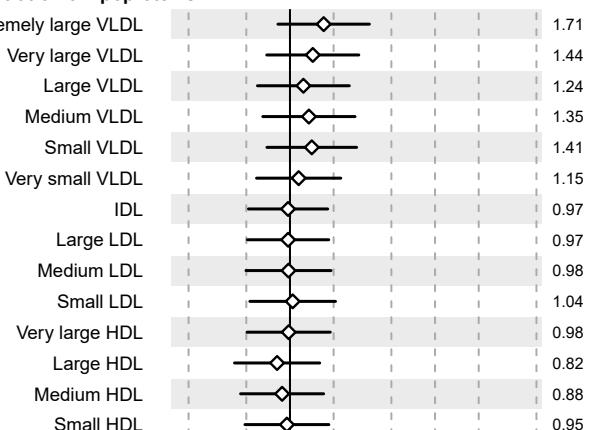
Phospholipids in lipoproteins (%)



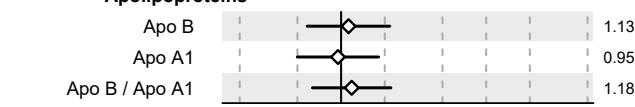
Total lipids in lipoproteins



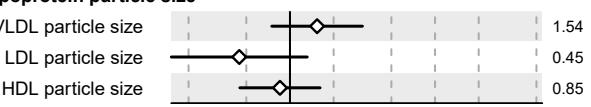
Concentration of lipoproteins



Apolipoproteins



Lipoprotein particle size



SD-scaled odds ratios (95% CI)

SD-scaled odds ratios (95% CI)

◆ Odds Ratio

Supplementary Data File V

Lists of predictors for the support vector machine (SVM), random forest (RF), and logistic regression (LR) models. These lists include both metabolic and clinical/demographic predictors.

For LR models negative beta-coefficients indicate a reduced likelihood of plaque development.

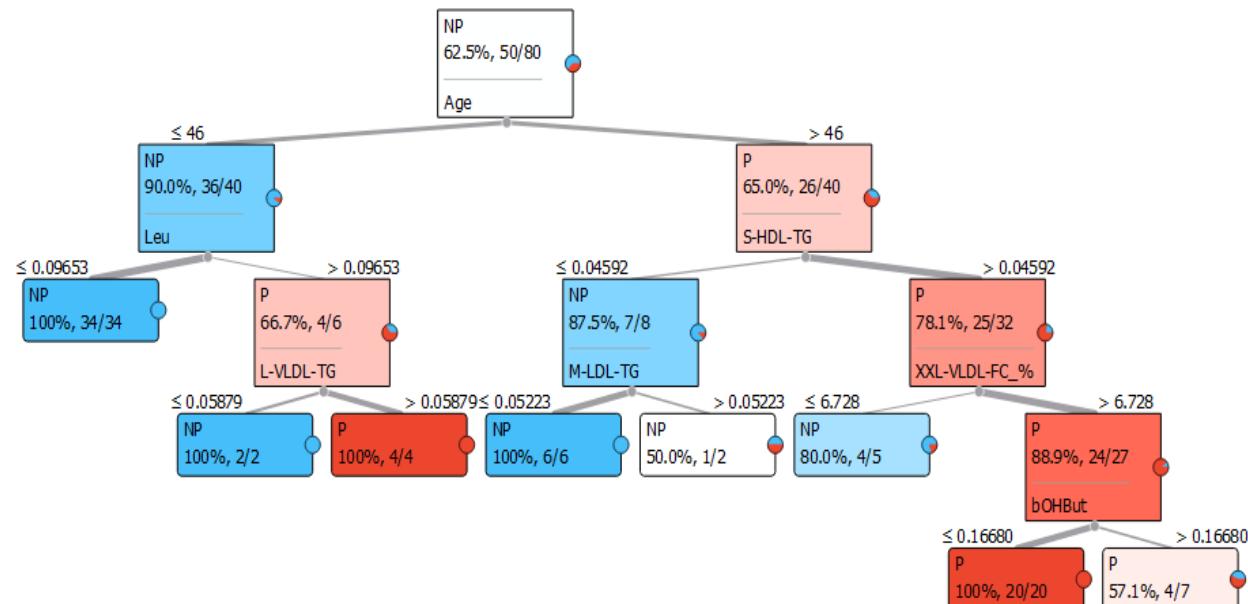
LR

Feature	Beta
Intercept	-0.74001
Cit	-0.09267
Gly	-0.08807
LDL-D	-0.08464
XXL-VLDL-CE_%	-0.0271
Tyr	0.03254
Aspirin	0.04446
Leu	0.04875
XXL-VLDL-PL_%	0.0805
IDL-TG_%	0.24108
L-VLDL-FC_%	0.42302
Age	0.9163

LR+I

Feature	Beta
Intercept	-1.17411
S-LDL-CE : RTX1	-0.68492
S-HDL-FC : XXL-VLDL-CE_%	-0.49192
Gly : His	-0.46405
Statins1 : ACE_Inh1	-0.37603
Lac : Cit	-0.37161
HCQ1 : ACE_Inh1	-0.23705
XL-HDL-TG : Glol	-0.15531
IDL-CE : Glol	-0.12971
His : XS-VLDL-CE_%	-0.09162
L-VLDL-TG : Lac	-0.08858
Glol : XXL-VLDL-CE_%	-0.07212
Gly : XS-VLDL-CE_%	-0.05998
Glol : His	-0.03934
DHA/FA : Age	0.019318
DHA/FA : Tyr	0.019945
L-HDL-PL_% : Age	0.022831
L-VLDL-TG : HCQ1	0.034663
Ethnicity1 : Dis_Dur	0.035969
XXL-VLDL-PL_% : Mean_BP	0.076793
DHA : Aspirin1	0.089295
DHA : HCQ1	0.096398
HCQ1 : Pred	0.129875
Age : Mean_BP	0.130915
SFA/FA : M-VLDL-FC_%	0.193219
Tyr : Dis_Dur	0.194325
M-VLDL-FC_% : L-LDL-FC_%	0.239355
HCQ1 : RTX1	0.282773
IDL-TG_% : S-LDL-C_%	0.292232
SM : Tyr	0.333199
Leu : Mean_BP	0.401734
Gln : Ace	0.464306
SM : Leu	0.597411
Ethnicity2 : HCQ1	0.755513
L-VLDL-FC_% : Age	1.056669
IDL-TG_% : Age	1.133333

TREE



SVM

Feature	Info gain	Gain ratio	Gini
Age	0.265962	0.133077	0.156848565
XXL-VLDL-PL_%	0.141398	0.070716	0.088892349
Dis_Dur	0.135487	0.067824	0.087452853
Serum-TG	0.134705	0.067368	0.072495588
M-VLDL-FC_%	0.131009	0.06552	0.07843351
L-VLDL-FC_%	0.117331	0.058679	0.07431745
Gly	0.10858	0.054303	0.067839718
XXL-VLDL-FC_%	0.107878	0.053952	0.063858611
XXL-VLDL-TG_%	0.105814	0.05292	0.061227032
XL-HDL-TG	0.10413	0.052077	0.060079934
XS-VLDL-TG	0.096361	0.048192	0.056031351
L-VLDL-TG	0.095729	0.047876	0.062778989
FAw6/FA	0.093352	0.046687	0.058190595
XS-VLDL-TG_%	0.090397	0.045209	0.057583307
M-HDL-PL_%	0.088215	0.044118	0.055491539
S-VLDL-CE_%	0.079157	0.039588	0.048406519
S-VLDL-TG_%	0.077822	0.03892	0.04820409
S-VLDL-C_%	0.077524	0.038771	0.047394374
M-VLDL-CE_%	0.074069	0.037043	0.047934185
L-LDL-PL_%	0.071928	0.035973	0.04354822
Val	0.066868	0.033442	0.042805979
XL-VLDL-C_%	0.066164	0.03309	0.039567113
S-HDL-TG	0.061336	0.030675	0.037137963
S-HDL-PL	0.060631	0.030323	0.039297208
XL-VLDL-CE_%	0.059298	0.029656	0.037677774
XS-VLDL-C_%	0.058991	0.029502	0.038150109
S-VLDL-TG	0.057799	0.028907	0.034708813
M-LDL-TG	0.057799	0.028907	0.034708813
HDL3-C	0.05693	0.028472	0.037070487
S-HDL-FC_%	0.056698	0.028356	0.036125818
S-HDL-TG_%	0.052717	0.026365	0.033899097
XXL-VLDL-TG	0.052487	0.02625	0.03477629
TG/PG	0.051165	0.025589	0.03308938
MUFA	0.051165	0.025589	0.03308938
LA/FA	0.051165	0.025589	0.03308938
DHA	0.051063	0.025538	0.033899097
M-HDL-CE_%	0.04719	0.0236	0.029580608
MUFA/FA	0.047121	0.023566	0.030052943
VLDL-C	0.047044	0.023528	0.030592754
UnSat	0.045896	0.022953	0.030120419
XL-VLDL-TG	0.045578	0.022795	0.029040797
Lac	0.045578	0.022795	0.029040797
XL-VLDL-TG_%	0.045578	0.022795	0.029040797
L-LDL-FC_%	0.044872	0.022441	0.028028652
ApoB/ApoA1	0.043339	0.021675	0.028028652
His	0.041357	0.020683	0.026071836
S-VLDL-PL	0.041357	0.020683	0.026071836
S-VLDL-FC	0.041357	0.020683	0.026071836
LDL-D	0.041357	0.020683	0.026071836
IDL-C_%	0.041039	0.020524	0.026611647
S-LDL-PL_%	0.041039	0.020524	0.026611647
Cit	0.04096	0.020485	0.025329596
XS-VLDL-CE_%	0.04096	0.020485	0.025329596
Gp	0.038978	0.019493	0.025532025
IDL-TG_%	0.038978	0.019493	0.025532025
BILAG	0.038679	0.020087	0.025140662
XL-HDL-CE_%	0.03813	0.019069	0.023912592
Pred	0.037255	0.019735	0.023328143
Ace	0.03601	0.018009	0.023642687
SM	0.035812	0.01791	0.022360635
Tyr	0.035812	0.01791	0.022360635
SFA	0.035066	0.017537	0.02283297

RF

Feature	Mean Decrease Gini
Age	2.080349527
XXL-VLDL-PL_%	0.888955783
Gly	0.649600356
L-VLDL-FC_%	0.572424738
M-VLDL-FC_%	0.56437281
XXL-VLDL-FC_%	0.559499324
Dis_Dur	0.519396339
XS-VLDL-TG	0.482499301
Tyr	0.455444397
S-HDL-TG	0.409914365
Serum-TG	0.391439372
FAw6/FA	0.388192969
TG-PG	0.378892389
XL-VLDL-CE_%	0.354219555
Gp	0.347350601
L-LDL-PL_%	0.3381853
Ace	0.335649782
SM	0.323290594
UnSat	0.320063214
Cit	0.31696656
LA/FA	0.29825867
M-LDL-TG	0.293115717
S-VLDL-TG	0.289791795
XXL-VLDL-CE_%	0.288773243
LDL-D	0.287921483
M-VLDL-CE_%	0.287400898
S-VLDL-FC	0.28642413
His	0.281679384
XL-VLDL-TG	0.264888864
GloI	0.259445587
DHA/FA	0.254323191
Leu	0.25151431
HDL3-C	0.247766072
IDL-CE_%	0.24617907
S-LDL-TG_%	0.234767637
MUFA/FA	0.234680429
S-VLDL-CE_%	0.23355315
XL-HDL-CE_%	0.232841191
L-VLDL-TG	0.231797139
M-HDL-PL_%	0.221490059
XXL-VLDL-TG_%	0.218981077
AcAce	0.218562189
XXL-VLDL-TG	0.216210364
Lac	0.215909427
M-LDL-TG_%	0.214327816
XS-VLDL-TG_%	0.213153171
FAw3/FA	0.212628503
Val	0.212366914
XL-VLDL-C_%	0.207888402
IDL-TG_%	0.206819083
S-HDL-FC_%	0.206101516
Ile	0.205528915
Phe	0.203912959
S-VLDL-PL	0.203648763
Alb	0.203272668
IDL-C_%	0.202099548
S-VLDL-C_%	0.199158967
Mean_BP	0.192382803
VLDL-C	0.189286867
L-VLDL-PL_%	0.187932941
L-LDL-FC_%	0.18655097

S-LDL-TG_%	0.035066	0.017537	0.02283297	S-HDL-TG_%	0.186142404
IDL-CE_%	0.034535	0.017272	0.022293159	M-VLDL-CE	0.184329337
Alb	0.033752	0.01688	0.020943631	MUFA	0.183467373
XL-HDL-PL	0.033359	0.016683	0.021955777	SFA/FA	0.180430792
Leu	0.033359	0.016683	0.021955777	L-VLDL-CE_%	0.179526582
AcAce	0.033213	0.01661	0.021550919	XL-HDL-TG	0.179265882
S-VLDL-CE	0.030262	0.015135	0.020133915	S-VLDL-CE	0.178285719
HDL-D	0.030262	0.015135	0.020133915	ApoB/ApoA1	0.173921148
DHA/FA	0.030262	0.015135	0.020133915	SFA	0.171487924
Crea	0.029875	0.014941	0.019324198	Crea	0.171078789
M-LDL-FC_%	0.029875	0.014941	0.019324198	XL-VLDL-TG_%	0.168618497
L-HDL-CE	0.029875	0.014941	0.019324198	M-HDL-CE_%	0.167522505
XXL-VLDL-CE_%	0.029875	0.014941	0.019324198	bOHBut	0.166752916
L-HDL-TG_%	0.028838	0.014422	0.019256722	XS-VLDL-CE_%	0.166317566
Ile	0.028562	0.014284	0.018514481	Ala	0.165464505
L-VLDL-PL_%	0.028562	0.014284	0.018514481	M-HDL-TG	0.164478424
XL-HDL-PL_%	0.026339	0.013173	0.016895048	S-VLDL-TG_%	0.163027411
L-HDL-FC_%	0.026339	0.013173	0.016895048	M-HDL-TG_%	0.162244755
Mean_BP	0.026339	0.013173	0.016895048	XL-HDL-PL_%	0.162070045
XS-VLDL-CE	0.024279	0.012143	0.015478044	VLDL-D	0.158505286
PC	0.024032	0.012019	0.015882903	XL-HDL-PL	0.154161062
VLDL-D	0.023592	0.011799	0.015275615	Glc	0.15099357
Phe	0.023592	0.011799	0.015275615	M-LDL-C_%	0.147870911
IDL-FC_%	0.023087	0.011546	0.015073186	XS-VLDL-FC_%	0.147646167
S-LDL-CE	0.022049	0.011027	0.014263469	M-LDL-FC_%	0.14692517
M-HDL-CE	0.022049	0.011027	0.014263469	XS-VLDL-C_%	0.145672529
S-HDL-C	0.021961	0.010983	0.013926087	S-HDL-PL	0.143805069
L-VLDL-CE_%	0.021961	0.010983	0.013926087	Gln	0.141645738
XL-HDL-TG_%	0.021533	0.010769	0.013858611	S-LDL-PL_%	0.140493412
M-LDL-TG_%	0.020495	0.01025	0.013453753	S-HDL-PL_%	0.140348659
M-HDL-FC_%	0.020495	0.01025	0.013453753	XL-VLDL-PL_%	0.140097918
M-VLDL-CE	0.018436	0.00922	0.012036749	XL-HDL-FC	0.138911268
bOHBut	0.018436	0.00922	0.012036749	L-HDL-FC_%	0.137831414
GloI	0.018306	0.009155	0.011699367	XL-HDL-FC_%	0.137166571
L-HDL-PL_%	0.018306	0.009155	0.011699367	IDL-CE	0.136617569
XL-VLDL-PL_%	0.017528	0.008766	0.011564414	L-VLDL-C_%	0.135089153
Immuno	0.016468	0.016539	0.010715068	L-LDL-C_%	0.13413523
XS-VLDL-FC	0.015988	0.007996	0.01014741	HDL-TG	0.132150074
XS-VLDL-PL_%	0.015988	0.007996	0.01014741	XS-VLDL-FC	0.130240893
Glc	0.014877	0.00744	0.009877504	M-HDL-CE	0.126042911
IDL-CE	0.013785	0.006894	0.009000311	HDL-D	0.125999038
S-HDL-PL_%	0.013466	0.006735	0.008797882	IDL-FC_%	0.125851309
L-VLDL-C_%	0.013466	0.006735	0.008797882	L-HDL-CE_%	0.125775259
Ethnicity	0.012917	0.007728	0.00826438	Pred	0.125222865
S-HDL-P	0.011923	0.005963	0.007785737	LA	0.125093113
S-LDL-FC_%	0.011923	0.005963	0.007785737	L-HDL-TG_%	0.124671801
S-VLDL-PL_%	0.011923	0.005963	0.007785737	PC	0.124197554
Aspirin	0.011282	0.021866	0.00762275	XL-VLDL-FC_%	0.124117955
HDL-TG	0.010499	0.005251	0.006908544	S-HDL-C	0.121782498
FAw3/FA	0.01024	0.005121	0.006638638	IDL-PL_%	0.12129939
IDL-PL_%	0.01024	0.005121	0.006638638	M-VLDL-PL_%	0.119900347
RTX	0.009167	0.009983	0.00591716	S-LDL-CE	0.119616117
S-HDL-FC	0.008815	0.004409	0.005761445	XS-VLDL-CE	0.1193615
XL-VLDL-FC_%	0.007921	0.003962	0.005086681	XS-VLDL-PL_%	0.117746304
M-VLDL-PL_%	0.007921	0.003962	0.005086681	L-HDL-TG	0.116757966
XL-HDL-FC	0.007848	0.003925	0.005221634	XL-HDL-TG_%	0.112665581
L-HDL-TG	0.007848	0.003925	0.005221634	M-HDL-FC_%	0.112570042
LA	0.007848	0.003925	0.005221634	S-HDL-P	0.112045597
ACE_Inh	0.00781	0.008056	0.005076493	L-HDL-CE	0.109711837
XS-VLDL-FC_%	0.007272	0.003637	0.004749299	M-LDL-PL_%	0.107640264
XL-HDL-FC_%	0.007272	0.003637	0.004749299	BILAG	0.106445198
S-VLDL-FC_%	0.004954	0.002478	0.003197342	L-HDL-PL_%	0.105003418
L-HDL-CE_%	0.004621	0.002311	0.00306239	S-HDL-FC	0.104869687
Statins	0.002409	0.004104	0.001605742	S-LDL-C_%	0.104495242
M-HDL-TG	0.002303	0.001152	0.001510433	S-LDL-FC_%	0.103489707
SFA/FA	0.002303	0.001152	0.001510433	S-VLDL-PL_%	0.10252821
Ala	0.002303	0.001152	0.001510433	S-VLDL-FC_%	0.084703227

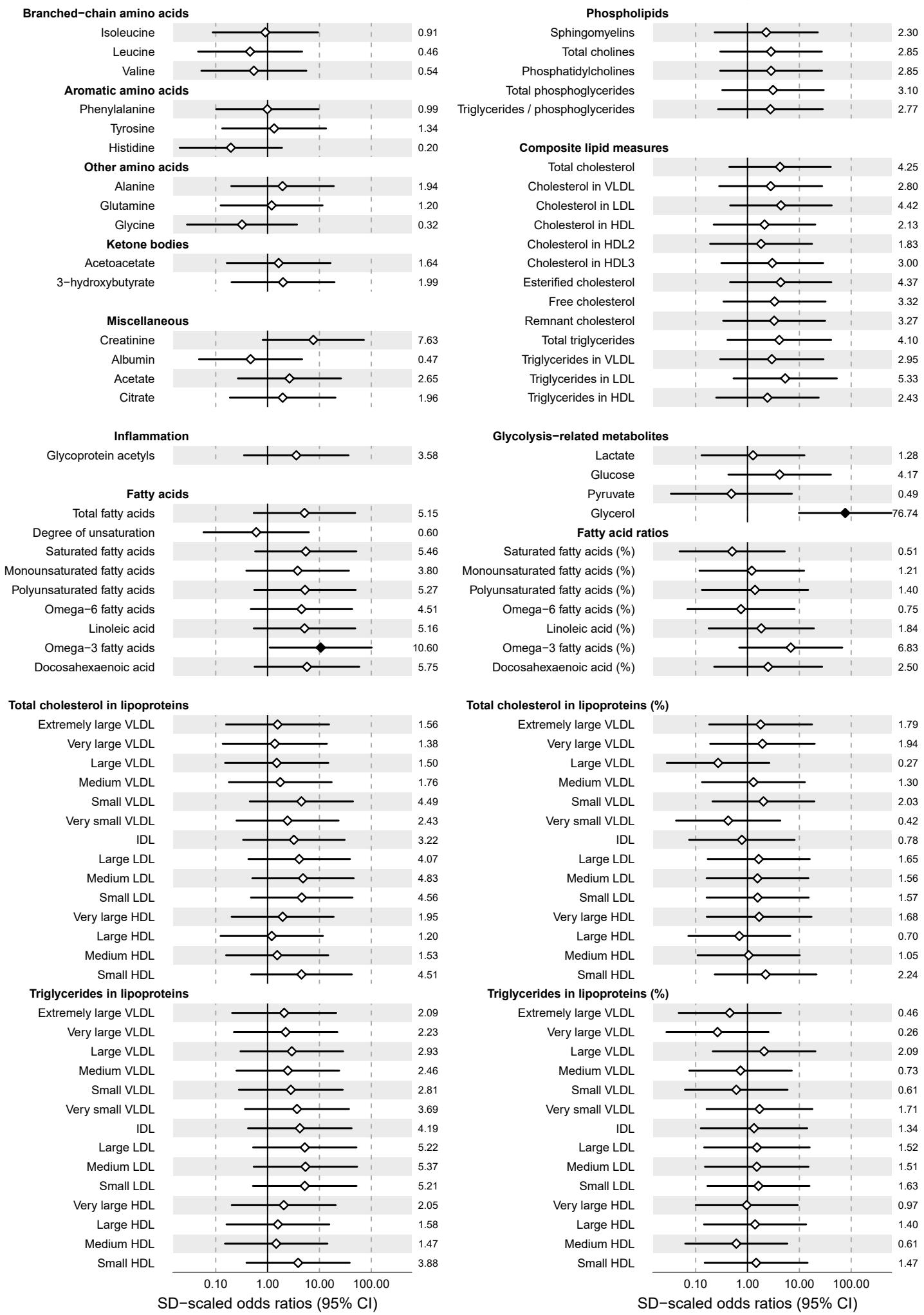
Gln	0.002303	0.001152	0.001510433
M-LDL-PL_%	0.002303	0.001152	0.001510433
S-LDL-C_%	0.002303	0.001152	0.001510433
M-HDL-TG_%	0.002303	0.001152	0.001510433
L-LDL-C_%	0.00076	0.00038	0.000498287
M-LDL-C_%	0.00076	0.00038	0.000498287
HCQ	5.08E-05	5.34E-05	3.33E-05

Ethnicity	0.051719038
RTX	0.03927328
Immuno	0.02321842
HCQ	0.022207051
ACE_Inh	0.020052247
Aspirin	0.019785466
Statins	0.017050286

Supplementary Data File VI

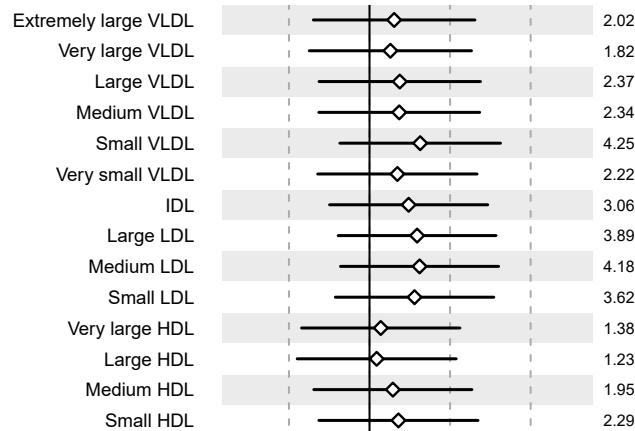
Forest plot visualization of the results of the logistic regression for individual metabolites in SLE vs Age. P-values <0.05 are given in coloured in circles.

Age Effect on Metabolites

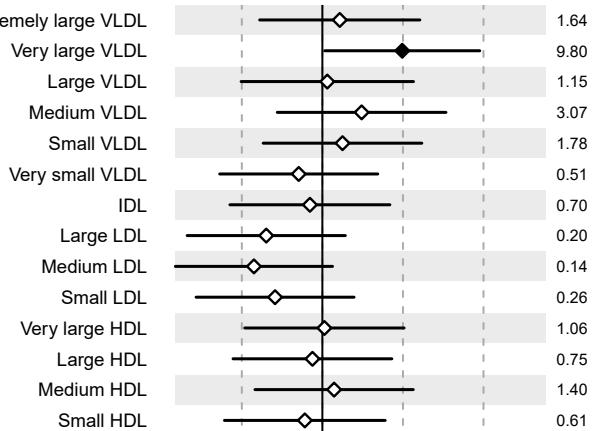


Age Effect on Metabolites

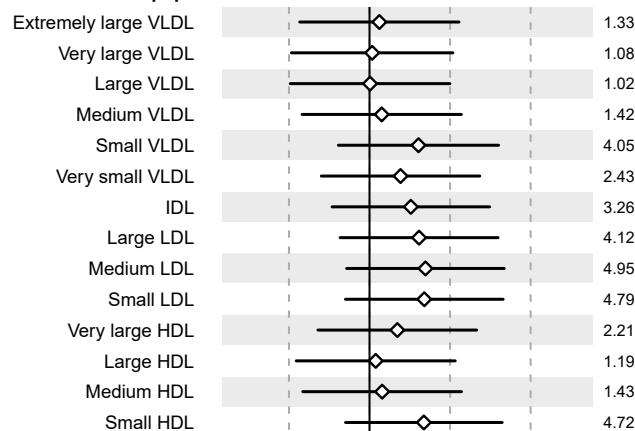
Free cholesterol in lipoproteins



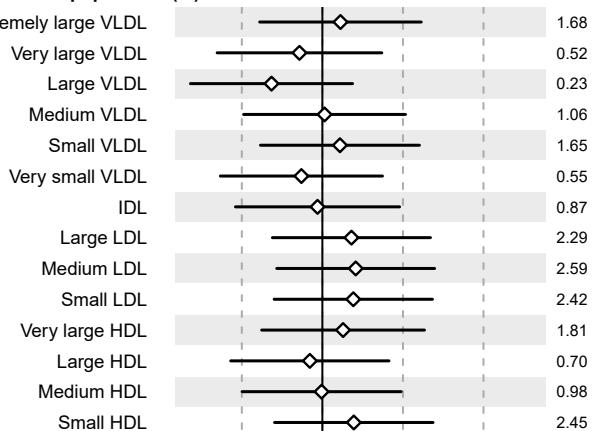
Free cholesterol in lipoproteins (%)



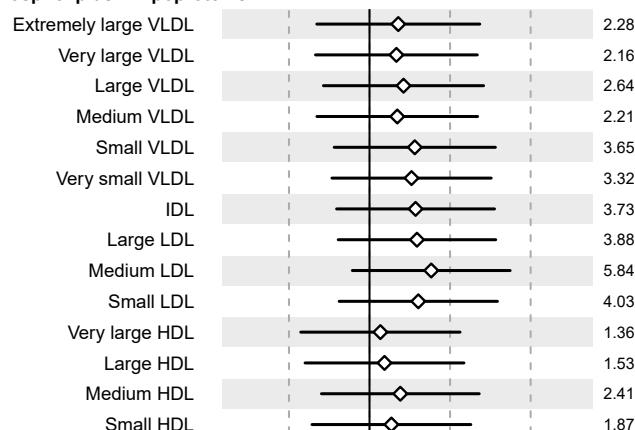
Cholesterol esters in lipoproteins



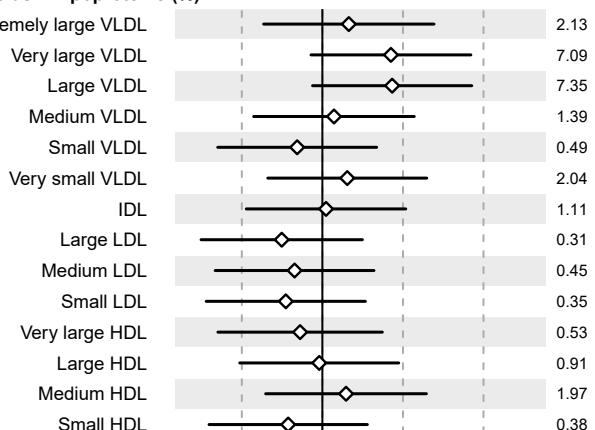
Cholesterol esters in lipoproteins (%)



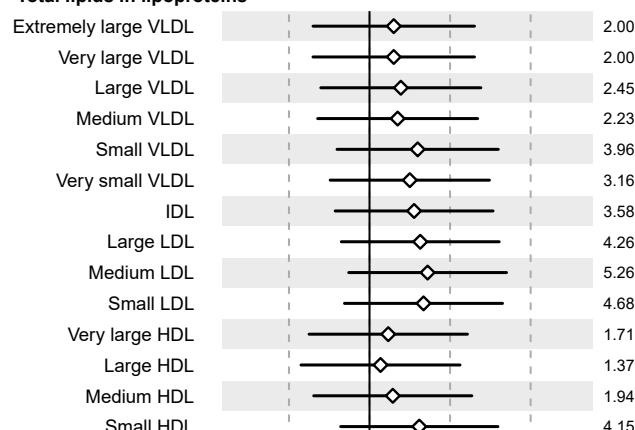
Phospholipids in lipoproteins



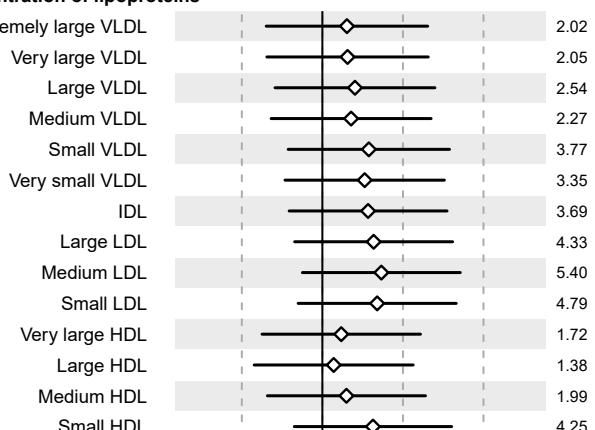
Phospholipids in lipoproteins (%)



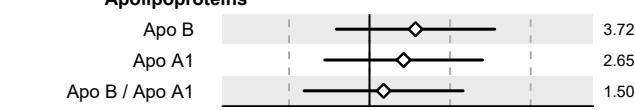
Total lipids in lipoproteins



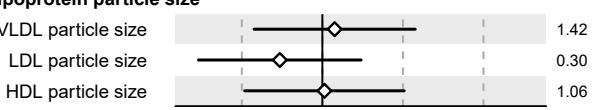
Concentration of lipoproteins



Apolipoproteins



Lipoprotein particle size



◆ Odds Ratio