Development and validation of cardiometabolic risk predictive models based on LDL oxidation and candidate geromarkers from the MARK-AGE data

Valeanu, Andrei; Margina, Denisa; Weber, Daniela; Stuetz, Wolfgang; Moreno-Villanueva, María; Dollé, Martijn E.T.; Jansen, Eugène HJM; Gonos, Efstathios S.; Bernhardt, Jürgen; Grubeck-Loebenstein, Beatrix; Weinberger, Birgit; Fiegl, Simone; Sikora, Ewa; Mosieniak, Grazyna; Toussaint, Olivier; Debacq-Chainiaux, Florence; Capri, Miriam; Garagnani, Paolo; Pirazzini, Chiara; Bacalini, Maria Giulia; Hervonen, Antti; Slagboom, P. Eline; Talbot, Duncan; Breusing, Nicolle; Frank, Jan; Bürkle, Alexander; Franceschi, Claudio; Grune, Tilman; Gradinaru, Daniela

Published in:
Mechanisms of Ageing and Development

DOI:
10.1016/j.mad.2024.111987

Publication date:
2024

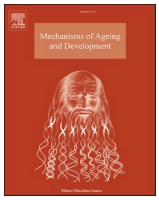
Document Version
Publisher's PDF, also known as Version of record

Mechanisms of Ageing and Development 222 (2024) 111987

Contents lists available at ScienceDirect

Mechanisms of Ageing and Development

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



Development and validation of cardiometabolic risk predictive models based on LDL oxidation and candidate geromarkers from the MARK-AGE data

Andrei Valeanu [a,*], Denisa Margina [a], Daniela Weber [b,c], Wolfgang Stuetz [d], María Moreno-Villanueva [e,f], Martijn E.T. Dollé [g], Eugène HJM Jansen [g], Efstathios S. Gonos [h], Jürgen Bernhardt [i], Beatrix Grubeck-Loebenstein [j], Birgit Weinberger [j], Simone Fiegl [k], Ewa Sikora [l], Grazyna Mosieniak [l], Olivier Toussaint [m], Florence Debacq-Chainiaux [m], Miriam Capri [n,o], Paolo Garagnani [n,p], Chiara Pirazzini [n], Maria Giulia Bacalini [q], Antti Hervonen [r], P. Eline Slagboom [s], Duncan Talbot [t], Nicolle Breusing [u], Jan Frank [d], Alexander Bürkle [e], Claudio Franceschi [n,v], Tilman Grune [b,c,w,x,y,z], Daniela Gradinaru [a,aa]

[a] *Carol Davila University of Medicine and Pharmacy, Faculty of Pharmacy, 6 Traian Vuia St., Bucharest 020956, Romania*
[b] *Department of Molecular Toxicology, German Institute of Human Nutrition, Potsdam-Rehbrücke, Nuthetal 14558, Germany*
[c] *NutriAct-Competence Cluster Nutrition Research Berlin-Potsdam, Nuthetal 14458, Germany*
[d] *Department of Food Biofunctionality, Institute of Nutritional Sciences (140), University of Hohenheim, Stuttgart 70599, Germany*
[e] *Molecular Toxicology Group, Department of Biology, University of Konstanz, Konstanz 78457, Germany*
[f] *Human Performance Research Centre, Department of Sport Science, University of Konstanz, Konstanz 78457, Germany*
[g] *Centre for Health Protection, National Institute for Public Health and the Environment, PO Box 1, Bilthoven 3720 BA, the Netherlands*
[h] *National Hellenic Research Foundation, Institute of Biology, Medicinal Chemistry and Biotechnology, Athens, Greece*
[i] *BioTeSys GmbH, Schelztorstr. 54-56, Esslingen 73728, Germany*
[j] *Research Institute for Biomedical Aging Research, University of Innsbruck, Rennweg, 10, Innsbruck 6020, Austria*
[k] *UMIT TIROL - Private University for Health Sciences, Medical Informatics and Technology, Hall in Tyrol 6060, Austria*
[l] *Laboratory of the Molecular Bases of Ageing, Nencki Institute of Experimental Biology, Polish Academy of Sciences, 3 Pasteur street, Warsaw 02-093, Poland*
[m] *URBC-NARILIS, University of Namur, Rue de Bruxelles, 61, Namur, Belgium*
[n] *Department of Medical and Surgical Sciences (DIMEC), Alma Mater Studiorum, University of Bologna, Bologna 40126, Italy*
[o] *Alma Mater Research Institute on Global Challenges and Climate Change (Alma Climate), University of Bologna, Bologna 40126, Italy*
[p] *IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy*
[q] *IRCCS Istituto delle Scienze Neurologiche di Bologna, Bologna, Italy*
[r] *Medical School, University of Tampere, Tampere 33014, Finland*
[s] *Section of Molecular Epidemiology, Leiden University Medical Centre, Leiden, the Netherlands*
[t] *Department of Unilever Science and Technology, Beauty and Personal Care, Sharnbrook, UK*
[u] *Department of Applied Nutritional Science/Dietetics, Institute of Nutritional Medicine, University of Hohenheim, Stuttgart 70599, Germany*
[v] *Laboratory of Systems Medicine of Healthy Aging, Institute of Biology and Biomedicine and Institute of Information Technology, Mathematics and Mechanics, Department of Applied Mathematics, N. I. Lobachevsky State University, Nizhny Novgorod 603005, Russia*
[w] *German Center for Diabetes Research (DZD), München-Neuherberg 85764, Germany*
[x] *German Center for Cardiovascular Research (DZHK), Partner Site Berlin, Berlin 13347, Germany*
[y] *University of Potsdam, Institute of Nutritional Science, Nuthetal 14458, Germany*
[z] *University of Vienna, Department of Physiological Chemistry, Faculty of Chemistry, Vienna 1090, Austria*
[aa] *Ana Aslan National Institute of Gerontology and Geriatrics, Bucharest, Romania*

* Corresponding author.
E-mail addresses: andrei.valeanu@umfcd.ro (A. Valeanu), denisa.margina@umfcd.ro (D. Margina), Daniela.Weber@dife.de (D. Weber), wolfgang.stuetz@uni-hohenheim.de (W. Stuetz), maria.moreno-villanueva@uni-konstanz.de (M. Moreno-Villanueva), Martijn.Dolle@rivm.nl (M.E.T. Dollé), eugenejansen99@gmail.com (E.H. Jansen), sgonos@eie.gr (E.S. Gonos), j.bernhardt@biotesys.de (J. Bernhardt), Beatrix.Grubeck-Loebenstein@uibk.ac.at (B. Grubeck-Loebenstein), birgit.weinberger@uibk.ac.at (B. Weinberger), simone.fiegl@umit-tirol.at (S. Fiegl), e.sikora@nencki.edu.pl (E. Sikora), g.mosieniak@nencki.edu.pl (G. Mosieniak), florence.chainiaux@unamur.be (F. Debacq-Chainiaux), miriam.capri@unibo.it (M. Capri), paolo.garagnani2@unibo.it (P. Garagnani), chiara.pirazzini5@unibo.it (C. Pirazzini), mariagiulia.bacalini@isnb.it (M.G. Bacalini), antti.hervonen@gmail.com (A. Hervonen), p.slagboom@lumc.nl (P.E. Slagboom), scoutwell.supplies@btinternet.com (D. Talbot), nicolle_breusing@yahoo.de (N. Breusing), jan.frank@uni-hohenheim.de (J. Frank), alexander.buerkle@uni-konstanz.de (A. Bürkle), claudio.franceschi@unibo.it (C. Franceschi), scientific.director@dife.de (T. Grune), daniela.gradinaru@umfcd.ro (D. Gradinaru).

https://doi.org/10.1016/j.mad.2024.111987
Received 25 July 2024; Received in revised form 27 August 2024; Accepted 5 September 2024
Available online 14 September 2024

ARTICLE INFO

Keywords:
LDL oxidation
MARK-AGE
Vascular ageing
Cardiometabolic risk
Machine learning

ABSTRACT

The predictive value of the susceptibility to oxidation of LDL particles (LDLox) in cardiometabolic risk assessment is incompletely understood. The main objective of the current study was to assess its relationship with other relevant biomarkers and cardiometabolic risk factors from MARK-AGE data. A cross-sectional observational study was carried out on 1089 subjects (528 men and 561 women), aged 40–75 years old, randomly recruited age- and sex-stratified individuals from the general population. A correlation analysis exploring the relationships between LDLox and relevant biomarkers was undertaken, as well as the development and validation of several machine learning algorithms, for estimating the risk of the combined status of high blood pressure and obesity for the MARK-AGE subjects. The machine learning models yielded Area Under the Receiver Operating Characteristic Curve Score ranging 0.783–0.839 for the internal validation, while the external validation resulted in an Under the Receiver Operating Characteristic Curve Score between 0.648 and 0.787, with the variables based on LDLox reaching significant importance within the obtained predictions. The current study offers novel insights regarding the combined effects of LDL oxidation and other ageing markers on cardiometabolic risk. Future studies might be extended on larger patient cohorts, in order to obtain reproducible clinical assessment models.

1. Introduction

Oxidative stress is one of the important drivers of multifactorial disease and can be quantified by various measures and clinical parameters (Ghezzi et al., 2017). Of them, of notable importance is the lipoprotein oxidation process, characterized *in vivo* by circulating oxidized low-density lipoproteins (oxLDL) and *in vitro* by a latent state, the susceptibility to oxidation of LDL (LDL oxidizability - LDLox). While oxLDL has been extensively studied as an important risk factor of cardiometabolic risk and pro-inflammatory response, LDLox remains largely unstudied with respect to its predictive value in such conditions, with some preliminary studies showing its implication in the atherosclerosis process (Liu et al., 2002; Aoki et al., 2012; Gradinaru et al., 2015).

The European Project FP-7 MARK-AGE was a large scale populational study which took place between 2009 and 2013; its main objective was to identify a specific combination of ageing biomarkers which could optimally explain the biological age, better than any marker considered individually (Bürkle et al., 2015; Capri et al., 2015). Even though several studies describing detailed analyses of the collected data were already published (Baur et al., 2015a, 2015b; Bürkle et al., 2015; Capri et al., 2015; Giampieri et al., 2015; Moreno-Villanueva et al., 2015a, 2015b; Weber et al., 2017; Moreno-Villanueva and Bürkle, 2019; Pinchuk et al., 2019, 2021; Kananen et al., 2021, 2023; Giacconi et al., 2023), no such analysis focused on examining the relationship between LDL oxidizability and cardiometabolic risk; moreover, to our knowledge, at the moment no specific machine learning algorithms were developed and validated based on MARK-AGE data, with the notable exception of research focused on creating general recommendations for machine learning driven data curation (Baur et al., 2015a; Giampieri et al., 2015).

Therefore, the main aim of the current study was to evaluate the MARK-AGE database with respect to the most relevant links between LDLox and other MARK-AGE parameters, exploring the cumulative effects of the analyzed relationships on the cardiometabolic risk of the randomly recruited age- and sex-stratified individuals from the general population (RASIG), quantified through blood pressure, body mass index and waist-to-hip ratio.

2. Subjects and methods

2.1. Study design and participants

The cross-sectional observational study was carried out on a relevant population sample of 1089 subjects: 528 men and 561 women, aged between 40 and 75 years old, selected among the participants included in the MARK-AGE group of randomly recruited age- and sex-stratified individuals from the general population (RASIG). According to the recommendations from the updated clinical guidelines, the cardiovascular risk assessment is undertaken for subjects over 40 years old (Williams et al., 2018; Mach et al., 2020). Only MARK-AGE subjects with complete data for all studied parameters were included in the present study. Participants from MARK-AGE cohort were enrolled, through the media, from seven European countries: Austria, Belgium, Finland, Germany, Greece, Italy and Poland. Subjects who reported seropositivity for HIV or hepatitis (HBV, HCV), whose blood was tested positive for HBV or HCV, and who were being treated for cancer or receiving glucocorticoids were excluded from the study (Bürkle et al., 2015; Capri et al., 2015).

The biological samples (fasting blood) collected from participants were processed and stored within the MARK-AGE consortium, according to rigorous Standard Operating Procedures and quality control measures, as described in recent articles (Moreno-Villanueva et al., 2015a; Jansen et al., 2015). Briefly, the double-coded blood samples (plasma and serum) were centrally stored in a Biobank and distributed to each MARK-AGE partner for the independent measurement of the specific candidate biomarkers. All of the subject's clinical and biochemical data obtained from each partner were uploaded to a central Database containing also the demographic and anthropometric data. This phenotypic database could be accessed and analyzed only at the end of the MARK-AGE project (Baur et al., 2015-a,b; Moreno-Villanueva et al., 2015b; Giampieri et al., 2015).

Standard demographic (age, sex) and anthropometric data (height, weight, waist circumference, body mass index, waist-to-hip ratio) were obtained from each participant. The resting systolic and diastolic blood pressures (SBP and DBP, mmHg) were recorded for all subjects, as well as whether there is currently a diagnosed high blood pressure problem. Participants completed a comprehensive questionnaire that included information on self-reported past and present diseases, hormone therapy (women), self-rated health status, as well as lifestyle characteristics (such as smoking status -never, former, current smoker, number of years of smoking (smoking years), alcohol and other beverage consumption -whether the subjects never drank beer, wine, juice or cola beverages, nutritional status -the quantitative consumption of meat, fish, eggs, bread, rice, fruits, vegetables, salty snacks and sweets, educational background, marital status and information about residence -house, apartment and whether the subjects lived with children) (Bürkle et al., 2015; Capri et al., 2015).

2.2. Laboratory methods

Blood glycated hemoglobin (Hb1AC), serum glucose, total cholesterol, triglycerides, γ-glutamyl transferase, uric acid, urea and creatinine, were measured on the clinical auto-analyzer (LX20-Pro, Beckman–Coulter, Woerden, TheNetherlands). Insulin and ferritin were measured with an immuno-analyzer (Access-2, Beckman–Coulter, Woerden, The Netherlands), as previously described (Jansen et al., 2015).

The cholesterol and triglyceride content of the serum lipoproteins fractions (HDL, LDL, VLDL) and subfractions (HDL2, LDL2, VLDL2) were determined using nuclear magnetic resonance (NMR) spectroscopy (Bruker Biospin), as described previously (Heijmans et al., 2006; Vaarhorst et al., 2011). Each measurement produces the signal amplitudes of lipoprotein subclasses that allows the estimation of the total lipoprotein particle concentration as well as their subclasses, including small particles (HDL2, LDL2, VLDL2) (Giacconi et al., 2019).

Insulin resistance was evaluated using as surrogate markers the Homeostasis Model Assessment-Insulin Resistance (HOMA-IR) and the Triglyceride Glucose (TyG) Index. HOMA-IR was calculated as follows: [fasting insulin (mU/L) x fasting glucose (mg/dL)]/405 (Matthews et al., 1985). TyG index is based on the levels of triglycerides and fasting plasma glucose,and was calculated as follows: Ln [fasting triglycerides (mg/dL) × fasting glucose (mg/dL)/2] (Ramdas Nayak et al., 2022).

The plasma atherogenic index (AIP) was calculated by the logarithmically transformed ratio of triglycerides on HDL-cholesterol, according to equation: log [fasting triglycerides (mg/dL)/HDL-cholesterol (mg/dL)] (Dobiásová and Frohlich, 2001).

C-reactive protein (CRP), fibrinogen and adiponectin were measured as markers of inflammatory status (Jansen et al., 2015). In particular, the assays for high-sensitive CRP (HS-CRP), fibrinogen were turbidimetric assays (immunoprecipitation). The HS-CRP were obtained from Beckman Coulter (Woerden, The Netherlands) and fibrinogen were obtained from Dialab (Neudorf, Austria). Adiponectin levels were assessed by time resolved fluorescent immunoassay utilizing R&D systems monoclonal antibodies (Mab 10651&BAM1065) on an AutoDELFIA® automated immunoassay system (PerkinElmer, Turku, Finland) (Yates et al., 2012).

LDL oxidation susceptibility (LDL oxidizability, LDLox) was assessed *in vitro* using serum LDL isolated by selective precipitation with buffered heparin (0.01 M citric acid / tri-sodium citrate pH 5.12, containing 100 000 IU of heparin) (Ahotupa et al., 1996; Scoccia et al., 2001; Gradinaru et al., 2009). Insoluble LDL were then sedimented by centrifugation at 1000 x g for 15 minutes. The sediment obtained was suspended in a saline phosphate buffer solution (1 vol of 0.1 M KH_2PO_4 / K_2HPO_4 in 9 volumes of 0.15 M NaCl, pH 7.4), containing 0.1 g % Triton X-100. The LDL lipid peroxidation was non-enzymatically induced by incubating the suspended LDL, at 37 °C in the presence of a pro-oxidant system ($FeSO_4$ and ascorbic acid, 1:1 molar ratio, 0.6 mM). Reaction kinetics showed a three-phase pattern - latency, propagation and decomposition phases, and in 5 selected samples the propagation phase reached a maximum at about 120 minutes. Based on these results, the incubation period was established at 120 minutes. The oxidative reaction was stopped with 1 mM EDTA and lipid peroxidation products, mainly the malondialdehyde (MDA), were evaluated as thiobarbituric acid reactive substances (TBARS) following the heating at 95 °C. Absorbance of samples was measured at 535 nm using a Lambda Bio10 spectrophotometer (Perkin-Elmer, USA). TBARS concentration in samples were determined with a calibration curve obtained by use of 1,1′,3,3′-tetra-methoxy-propane standard solution. Results were expressed as nmol malondialdehyde (MDA) equivalent content/mL serum. In our standard working conditions, the intra-assay CV was 6.5 % and the inter-assay CV was 7.4 %.

The total amount of plasma stable metabolic pathway products of NO [NOx, the sum of nitrites and nitrates (NO_2^- + NO_3^-)] was determined using the Griess reagent, following the quantitative conversion of nitrates (NO_3^-) to nitrites (NO_2^-) with nitrate reductase (kit 23479, SIGMA). The results were expressed in μmols NOx/L plasma. Intra- and inter-assay CV were below 7 % and 9 %, respectively.

Plasma malondialdehyde (MDA) was determined by RP-HPLC coupled with fluorescence detection after derivatization with thiobarbituric acid as described by Weber et al. (Weber et al., 2014, 2017).

Plasma levels of tocopherols (α-/γ-tocopherol) were simultaneously determined by HPLC and spectrophotometric detection as previously described (Weber et al., 2014; Stuetz et al., 2016).

The DNA methylation status of a number of promoter-associated 5′-cytosine-phosphate-guanine-3′ (CpG) islands in *FHL2* (Four and a half LIM domains 2) gene was assessed in peripheral blood mononuclear cells (PBMCs) using an optimized MALDI-TOF Mass Spectrometry (EpiTYPER Mass Spectrometry). This technique provides quantitative methylation for CpG sites within a target region (Garagnani et al., 2012).

2.3. Preliminary statistical analysis

In order to provide a general characterization of the studied RASIG population sample dataset, we used Python Programming Language, version 3.9.2 (Python Software Foundation, 2021). Results are presented as mean ± standard deviation, for normally distributed parameters and as median [quartile 25; quartile 75] for non-normally distributed. Normality was assessed with the Shapiro-Wilk test. Comparison between male and female groups was performed using Kolmogorov-Smirnov test. The variables included in the preliminary statistical analysis, as well as the mean (and median values when appropriate) are presented in Table 1.

Spearman's rank order (*rho*) correlation analysis was performed to examine associations of LDLox with chronological age, clinical, metabolic, oxidative stress, inflammatory and epigenetic MARK-AGE biomarkers. The level of significance was set at 0.05.

2.4. The development and validation of several combined hypertension and obesity prediction models

2.4.1. Predictive variables and outcomes

After performing the preliminary statistical analysis, several machine learning models were developed, with the aim of estimating the risk of overlapping arterial hypertension and overweight status/obesity. Arterial hypertension (high blood pressure, HBP) was considered at least stage 1, as defined by the current guidelines developed by the European Society of Cardiology: SBP of at least 140 mmHg or DBP of at least 90 mmHg, or if "problem_blood_pressure" was set to "yes" in the MARK-AGE dataset, since several subjects could already have been under treatment for arterial hypertension and hence had their blood pressure under control. The overweight /obesity status, was considered through waist-to-hip ratio (WTHR) and body mass index (BMI), according to the general risk criteria recommended by World Health Organization (WHO) and to recent criteria for high blood pressure (World Health Organization, 2008; Pedregosa et al., 2011; Williams et al., 2018). Overweight status was defined through a BMI of at least 25 kg/m^2, while obesity was defined either through BMI (at least 30 kg/m^2) or WTHR (at least 0.85 for females and at least 0.9 for males). Hence, three types of prediction models were developed: Case a (HBP+WTHR) – overlapping status of hypertension and obesity, quantified through WTHR; Case b ($HBP+BMI_{overweight}$) – overlapping status of hypertension and overweight, quantified through BMI; Case c ($HBP+BMI_{obese}$) – overlapping status of hypertension and obesity, quantified through BMI). The combined status of arterial hypertension and overweight/obesity status was chosen as outcome, since the identification of subjects at higher cardiometabolic risk was deemed more relevant than building separate machine learning models for hypertension and overweight/obesity status.

For each outcome, the predictive biomarkers were chosen based on the preliminary statistical analysis results (the biomarkers with the highest correlation with SBP, DBP, WTHR and BMI) and domain knowledge; in addition, LDLox was added as predictive biomarker, alone and in combination with HDL cholesterol and DNA methylation of FHL2 on 13–15 and 16–17 sites, since one of the main aims of the current study was to evaluate the impact of LDLox on vascular ageing and cardiometabolic risk, as well as to measure the comparative effects with other well established metabolic risk factors, such as low levels of HDL cholesterol, HOMA-IR, TyG or AIP. Moreover, several lifestyle variables were included in the model, based on their specific associations with

Table 1
Demographic, anthropometric and clinical characteristics, systemic metabolic variables and MARK-AGE candidate biomarkers for ageing, recorded in male and female subjects aged 40 – 75 years old, selected from RASIG participants (n=1089).

Variable	Male (n = 528)	Female (n = 561)	p
Demographic, anthropometric and clinical characteristics			
Age, years	58.2 ± 9.8	57.8 ± 10.3	0.418 (NS)
Systolic BP, mmHg	**139.5 ± 19.2**	**131.8 ± 19.1**	**< 0.001**
Diastolic BP, mmHg	**84.4 ± 11.1**	**79.3 ± 10.3**	**< 0.001**
BMI, kg/m^2	**27.0 ± 3.8**	**25.6 ± 5.0**	**< 0.001**
Waist–to-hip ratio	**0.95± 0.071**	**0.87 ± 0.137**	**< 0.001**
Systemic metabolism and toxicity biochemical parameters			
Fasting glucose, mmol/L	**5.5 ± 1.4**	**5.1 ± 0.9**	**< 0.001**
HbA1C, %	**6.03 ± 0.76**	**6.02 ± 0.51**	0.413 (NS)
Insulin, μIU/mL	**6.7 ± 4.8**	**5.8 ± 5.2**	**< 0.001**
Total cholesterol, mg/dL	**237.7 ± 44.0**	**246.6 ± 44.1**	**< 0.001**
Triglycerides, mg/dL	**112.9 ± 59.5**	**104.0 ± 70.3**	**< 0.001**
γ-GT, U/L	**25.9 ± 30.24**	**17.48 ± 30.21**	**< 0.001**
Uric acid, mg/L	**51.21 ± 11.29**	**39.51 ± 9.98**	**<0.001**
Urea, mmol/L	**5.87 ± 1.39**	**5.47 ± 1.40**	**< 0.001**
Creatinine, μmol/L	**81.10 ± 15.34**	**66.27 ± 12.34**	**< 0.001**
Cardiometabolic risk markers			
HOMA-IR	**1.726 ± 1.637**	**1.39 ± 1.75**	**< 0.001**
TyG	**8.61 ± 0.54**	**8.36 ± 0.48**	**< 0.001**
AIP	**0.25 ± 0.27**	**0.09 ± 0.23**	**<0.001**
Systemic inflammation parameters			
Fibrinogen, g/L	**3.4 ± 1.5**	**3.7 ± 1.4**	**< 0.001**
hsCRP, μg/L	2.25 ± 3.11	2.21 ± 3.30	0.479 (NS)
Ferritin, ng/mL	**116.53 ± 109.56**	**589.04 ± 54.22**	**< 0.001**
Adiponectin, μg/mL	**11.18 ± 5.64**	**18.16 ± 8.32**	**< 0.001**
NOx(NO_2 + NO_3), μmol/L	28.39 ± 13.54	29.22 ± 14.18	0.206(NS)
Serum lipoproteins fractions and subfractions			
HDL-cholesterol, mg/dL	**65.7 ± 15.3**	**79.3 ± 17.8**	**< 0.001**
HDL2-cholesterol, mg/dL	**38.00 ± 6.24**	**39.27 ± 6.25**	**0.002**
HDL2-triglycerides, mg/dL	**5.52 ± 2.04**	**5.23 ± 2.21**	**0.002**
LDL-cholesterol, mg/dL	126.4 ± 32.1	130.0 ± 31.4	0.09 (NS)
LDL2-cholesterol, mg/dL	**81.33 ± 20.75**	**70.50 ± 17.17**	**< 0.001**
LDL2-triglycerides, mg/dL	**7.68 ± 4.26**	**6.52 ± 3.15**	**< 0.001**
VLDL-cholesterol, mg/dL	**30.70 ± 17.13**	**23.55 ± 13.51**	**< 0.001**
VLDL2-cholesterol, mg/dL	**10.63 ± 4.53**	**9.71 ± 3.98**	**0.005**
VLDL2-triglycerides, mg/dL	**12.6 ± 5.51**	**11.52 ± 4.68**	**0.012**
MARK-AGE candidate biomarkers of ageing			
LDLox, nmol MDA/mL	18.73 ± 7.20	19.51 ± 7.7	0.344 (NS)
Malondialdehyde, μmol/L	**0.348 ± 0.19**	**0.329 ± 0.23**	**< 0.001**
α - Tocopherol, μmol/L	**28.226 ± 7.63**	**29.417± 7.58**	**0.012**
γ - Tocopherol, μmol/L	1.528 ± 0.85	1.506 ± 0.82	0.992 (NS)
*FHL2*_CpG_13,14,15	0.336 ± 0.05	0.336 ± 0.05	0.999 (NS)
*FHL2*_CpG_16,17	0.508 ± 0.08	0.51 ± 0.09	0.828 (NS)

Data are presented as means ± SD (standard deviations); n, number of subjects; NS – non-significant
RASIG: randomly recruited age-stratified individuals from the general population.
BP, blood pressure; BMI, body mass index; HOMA-IR, homeostasis model assessment of insulin resistance; hsCRP, high-sensitivity C-reactive protein; γ-GT, Gamma glutamyl transferase; HbA1C, Glycosylated haemoglobin A1C; HDL, high-density lipoprotein; LDL, low-density lipoprotein; VLDL, very low density lipoprotein; HDL1, LDL1, VLDL1 – large lipoprotein particles; HDL2, LDL2, VLDL2 – small lipoprotein particles; LDLox, LDL susceptibility to oxidation; MDA, malondialdehyde; NOx, nitric oxide metabolic pathway products; *FHL2*, Four and a half LIM domains 2 gene methylation status of promoter-associated CpG islands; AIP, plasma atherogenic index; TyG index, the product of the levels of triglycerides and fasting plasma glucose.

outcomes a, b and c. Table 1 presents the specific predictive variables for each case. It should be mentioned that the predictive criteria such as sex, age and smoking years, were the same for all cases, while the lifestyle predictors differed depending on their age adjusted correlations with each outcome (the first 10 lifestyle factors in terms of age adjusted Spearman correlation (absolute value) coefficient were selected) (Python Software Foundation, 2021).

2.4.2. *The validation process of the machine learning algorithms*

The validation of the cardiometabolic risk predictive models developed on RASIG population was performed both internally - through a 10-time repeated 5-fold cross validation technique, on RASIG and externally - by using data obtained from 2 different study groups of subjects also analyzed within the MARK-AGE project, namely descendants from long-living parents (GEHA offspring, GO; n=346 subjects) which are predicted to age at a slower rate than the average population, and spouses of GO (SGO; n=192 subjects) (Bürkle et al., 2015).

In terms of specific machine learning algorithms, 3 algorithms were evaluated for each case: Multilayer Perceptron, Support Vector Machines (through scikit-learn library) and Extreme Gradient Boosting (XGBoost Classifier - xgboost library) (Pedregosa et al., 2011; Chen and Guestrin, 2016). The algorithms were evaluated based on their ability to correctly classify the subjects in one of the two classes (0 – absent status of HBP and obesity/overweight, 1 - present status of HBP and obesity/overweight), as well estimating the probability (from 0 % to 100 %) that a patient has the overlapping status of HBP and obesity/overweight. Accuracy, precision, recall, f1 score and Matthews correlation were used to estimate the algorithms' classification performance, while the Area Under the Receiver Operating Characteristic Curve (ROC AUC) Score was used to evaluate how well the algorithms perform in terms of probability estimation and was considered the most important validation metric. With regards to model parameters, in order to provide an optimized and unbiased approach, they were chosen for each algorithm and train-test split (for both internal and external validation) by implementing a hyperparameter tuning through the RandomizedSearchCV option available in scikit-learn; the ROC AUC score was chosen as evaluation metric, for choosing the best combination of hyperparameters. The number of iterations was set to 60 and the random state was set to 42 for both the RandomizedSearchCV and the evaluated machine learning algorithms. In addition, in order to quantify the importance that each variable had on the predictions for all splits of the cross-validation, the random permutation feature importance was implemented, with ROC AUC Score as evaluation metric. From this, the average percentage with which each variable contributed was computed (Pedregosa et al., 2011). The feature importance was computed with the main goal of assessing the relative prediction contribution of the LDLox based predictive variables in relation with the other parameters.

3. Results

3.1. *Cohort characteristics*

Regarding the general characteristics of the selected MARK-AGE biomarkers assessed in the study subjects (Table 1), there was no significant difference in the age distribution between male and female groups. Systolic and diastolic blood pressure mean values were significantly higher in male subjects, as compared with the female group. Concerning the anthropometric characteristics, we noticed that BMI and WTHR mean values were significantly higher in male subjects, as compared to the female group.

With regard to the global metabolic profile, the serum glucose, insulin and HOMA-IR mean levels were significantly higher in men whereas total cholesterol levels were significantly higher in women. Men had higher levels of serum triglycerides, AIP and TyG index values, as compared with women. Systemic inflammation parameters

Table 2
The predictive variables selected from the MARK-AGE database for the machine learning prediction models.

Case a (HBP+WTHR)	Case b (HBP+$BMI_{overweight}$)	Case c (HBP+BMI_{obese})
Case-relevant lifestyle factors[a]		
consume_cake_pie	drink_other_alco_never	day_activities_bending_kneeling
drink_beer_never	day_activities_several_stairs	day_activities_walking_several_miles
consume_other_supplements	day_activities_bending_kneeling	day_activities_walking_hundred_yards
drink_other_alco_never	day_activities_walking_several_miles	day_activities_several_stairs
consume_icecream_dessert	consume_sausages	day_activities_one_stairs
ip_education_university degree	day_activities_walking_half_mile	day_activities_walking_half_mile
consume_bread_whole	drink_beer_never	day_activities_bathing_dressing_self
drink_never_juice	consume_icecream_dessert	day_activities_vigorous
consume_sausages	day_activities_walking_hundred_yards	day_activities_moderate
consume_candies_sweets	consume_meat	consume_meat
Common predictive variables for all 3 cases (description)		
LDLox serum		
Risk_HDL-C+LDLox_12 (the overlapping status: LDLox≥12 AND HDL-cholesterol<40 mg/dl (males) or <50 mg/dl (females))		
Risk_HDL-C+LDLox_16 (the overlapping status: LDLox≥16 AND HDL-cholesterol<40 mg/dl (males) or <50 mg/dl (females))		
Risk_HDL-C (HDL cholesterol<40 mg/dl (males) or <50 mg/dl (females))		
Risk_LDLox/HDL-C (the ratio $\frac{LDLox}{HDL\ cholesterol}$)		
FHL2_13_14_15_+ LDLox (the product *LDLox x FHL*213.14.15.*CpG site*):		
FHL2_16_17_+ LDLox (the ratio $\frac{LDLox}{FHL216.17.CpG\ site}$)		
Sex		
Chronological age		
HOMA-IR		
TyG		
AIP		
Adiponectin		
Uric acid		
Smoking years		

AIP - plasma atherogenic index; *FHL2* - Four and a half LIM domains 2 gene methylation status of promoter-associated CpG islands; HDL-C, HDL-cholesterol; HOMA-IR, Homeostasis Model Assessment-Insulin Resistance; TyG - Triglyceride Glucose Index.

[a] Lifestyle factors beginning with "consume" – the quantitative consumption of the specific food (frequency per month – was converted to continuous data depending on the number of times per month the specific food was consumed – daily consumption: 30; 4–6 times per week: 20; 1–3 times per week: 8; 1–3 times per month: 2; never: 0); lifestyle factors beginning with "drink_never" – specified whether the patient never drinks the specific beverage (categorical variable – yes/no); lifestyle factors beginning with "day_activities" – specified how limited was the patient with regards to the specific physical activity (categorical variable – categories: not limited, little limited, limited at a high degree); ip_education_university degree – specified whether the patient has a university degree (categorical variable – yes/no).

fibrinogen, ferritin and adiponectin were significantly higher in female subjects. There were no significant differences between men and women regarding hsCRP. Serum γ-GT, uric acid, urea and creatinine levels were significantly higher in male subjects. Important variations in the lipid metabolism biomarkers evaluated as lipoprotein fraction and subfraction concentrations, were pointed out between sexes: women had higher concentrations of cholesterol in HDL lipoprotein particles and in the small HDL2 subfraction, whereas triglycerides levels in HDL2 were significantly lower as compared with men. Although women had no significantly higher levels of LDL-cholesterol, both cholesterol and triglycerides levels in small lipoprotein subfraction LDL2 were significantly lower than in men. Also, men had higher levels of cholesterol and triglycerides in VLDL and VLDL2, as compared with women (Table 1).

Among oxidative stress biomarkers, serum malondialdehyde (MDA) levels were significantly higher in men *versus* women, in contrast to the plasma α-tocopherol which were higher in women. We couldn't notice significant sex-related variations in the LDL oxidizability nor the nitric oxide metabolic pathway products (NOx). The DNA-based MARK-AGE biomarkers of ageing measured in PBMCs, namely the *FHL2* genes methylation status, had similar levels in men and women.

In Fig. 1 are depicted the sex-differences in the distribution of BMI groups. It should be noted that the overall distribution for RASIG was 43.53 % (normal weight), 39.21 % (overweight) and 17.26 % (obese).

Overall, in the whole study group of RASIG participants, 19 % of participants were current smokers (23 % in males and 15.2 % in females), while the average number of years of smoking was 11.64 (13.97 for males and 9.44 for females). The most frequently associated CVD risk factors recorded were the HDL defined threshold (HDL cholesterol<40 mg/dl for males – 1.89 % and <50 mg/dl for females – 3.92 %; overall – 2.94 %), the abdominal obesity, quantified through the WTHR threshold (80.11 % in males and 58.29 % in females; overall – 68.87 %), and the high blood pressure (63.6 % in males and 43.1 % in females; overall – 53 %).

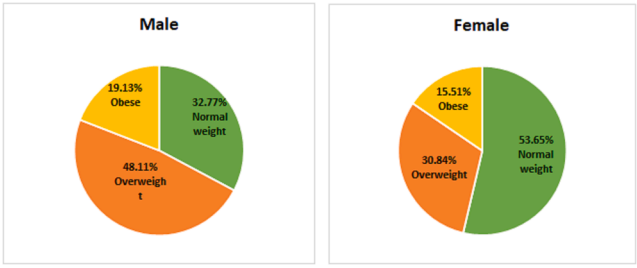


Fig. 1. Distribution of male and female participants according to BMI groups: normal weight (BMI 25 < kg/m^2), overweight (25 kg/m^2 ≥ BMI < 30 kg/m^2) and obese (BMI ≥ 30 kg/m^2).

3.2. LDLox relationships with MARK-AGE parameters

As there were reported differences between sexes with respect to different evaluated parameters, we analyzed further the associations of LDL oxidizability with different MARK-AGE biomarkers, distinctly in male and female subjects. In Fig. 2(a,b,c) are illustrated the correlation coefficients adjusted to age between LDLox and markers representative for lipid metabolism, systemic inflammation, oxidative stress, gene methylation status and cardiometabolic risk evaluation.

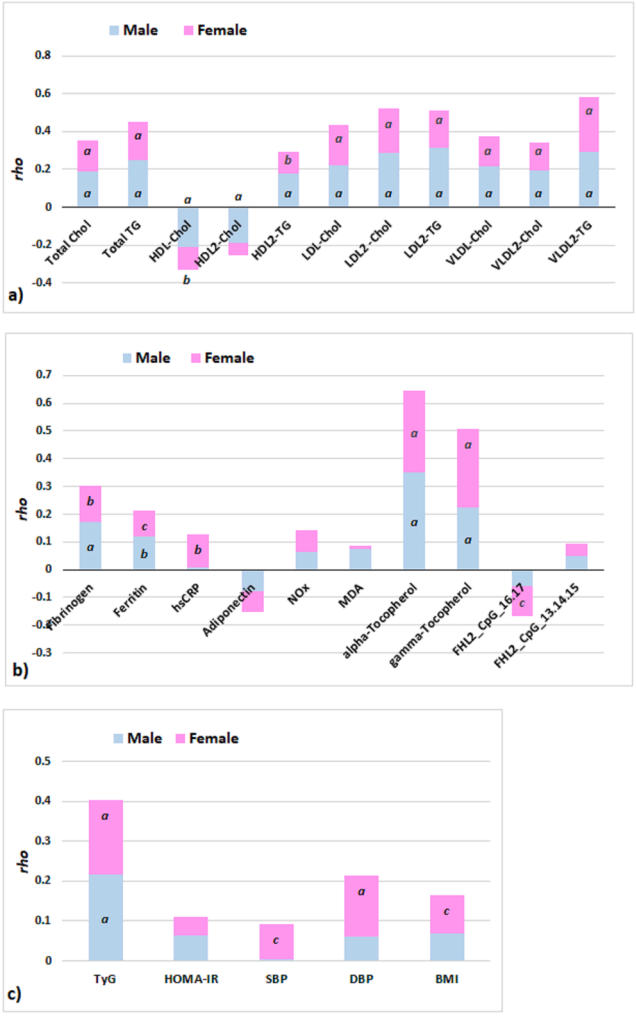


Fig. 2. Stacked bars showing the Spearman's (*rho*) correlation coefficients adjusted to age between LDL oxidizability (LDLox) and serum representative MARK-AGE biomarkers specific for the lipid metabolism - lipoproteins fractions and subfractions **(a)**, systemic inflammation and epigenetic markers **(b)**, anthropometric and clinical parameters specific for glucose homeostasis and cardiometabolic risk evaluation **(c)**, in male (n=528) and female (n=561) RASIG participants. Chol – cholesterol; TG - triglycerides; HDL, high-density lipoprotein; LDL, low-density lipoprotein; VLDL, very low density lipoprotein; HDL2, LDL2, VLDL2 – small lipoprotein particles; *FHL2* - Four and a half LIM domains 2 gene methylation status of promoter-associated CpG islands; hsCRP, high-sensitivity C-reactive protein; NOx - nitric oxide metabolic pathway products; HOMA-IR, Homeostasis Model Assessment-Insulin Resistance; SBP - Systolic Blood Pressure; DBP – Diastolic Blood Pressure; BMI – Body Mass Index; TyG - Triglyceride Glucose Index. Statistical significance: ^{a} $p < 0.001$; ^{b} $p < 0.01$ and ^{c} $p < 0.05$; In unlabeled histograms Spearman's (*rho*) values are non-significant.

We examined the relationships between LDLox and the advanced lipoprotein profile, comprising the cholesterol and triglycerides concentrations in lipoproteins fractions and subfractions (Fig. 2-a). LDLox was strongly positively associated in both men and women with total serum cholesterol, triglycerides, LDL-cholesterol and VLDL-cholesterol, as well as cholesterol and triglycerides content in small LDL and VLDL subfractions (LDL2 and VLDL2). Only in men, LDLox was significantly positively correlated with HDL2-triglycerides and negatively correlated with HDL2-cholesterol content (small HDL subfraction).

As regards the systemic inflammation parameters, in both men and women, LDLox was positively associated with increasing levels of fibrinogen, ferritin and nitric oxide metabolic pathway products (NOx) (Fig. 2-b). By contrast, LDLox was negatively correlated with adiponectin levels. Only in women, LDLox was significantly positively correlated with hsCRP. Only in men, LDLox was positively, non-significantly associated with plasma lipid peroxidation product malondialdehyde (MDA). Correlation analysis identified, equally in men and women, strong positive associations between LDLox and plasma levels of α- and γ-tocopherol.

In both sexes, LDLox levels were negatively associated with hypermethylation of FHL2_CpG_16,17 gene, measured in PBMCs but in women this correlation was significant. By contrast, in men and women, LDLox was positively but non-significantly correlated with hypermethylation of FHL2_CpG_13,14,15 gene.

The correlations between LDLox and the anthropometric and clinical parameters specific for glucose homeostasis and cardiometabolic risk evaluation pointed out significantly positively associations with TyG index in men and women. Only in women, LDLox was significantly positively correlated with systolic and diastolic blood pressure and with BMI (Fig. 2-c).

3.3. Age-related changes in LDLox and MARK-AGE parameters

It is relevant to display - comparatively in men and women - the association with chronological age of the MARK-AGE parameters which were found to be significantly correlated with LDLox. As a general feature, all age-related correlation coefficients had higher values in women, as compared to men (Table 3). Furthermore, only in female subjects' significant positive correlations were identified with lipoprotein fractions and subfractions, except the HDL-cholesterol which was significantly, positively correlated with chronological age only in men (Table 3).

In both men and women results showed increasing levels of systemic inflammation parameters with increasing chronological age, but only in women strong associations with almost all the tested biomarkers were pointed out, namely: ferritin, fibrinogen, hsCRP, adiponectin, MDA, α-tocopherol, LDLox and LDLox/apoB (Table 3). A weak increase in NOx with chronological age was evidenced only in female subjects. In men a significant age-related increase only in fibrinogen and adiponectin levels and a significant decrease in LDLox and LDLox/apoB were pointed out. The increasing methylation status of *FHL2* genes, measured in PBMCs, are among the most significant parameters that correlated with chronological age, in both sexes.

Among the most relevant parameters that increased with chronological age, in women and men, were systolic blood pressure levels (SBP), WTHR, BMI and HOMA-IR, whereas TyG index and diastolic blood pressure (DBP) was significantly increased only in women. Arterial hypertension (high blood pressure, HBP) as a condition identified within MARK-AGE subjects, was significantly increased with age in both men and women. (Table 3).

Fig. 3-a,b,c,d shows comparatively by age-groups, in female and male subjects, the age-related changes in LDL oxidizability parameters. Overall, in women an increase with age in LDLox levels, mainly from the 50–59 age group, and also in LDLox/apoB values, mainly from the 60–69 age-group was pointed out. By contrast, in men a gradual decrease in LDLox and LDLox/apoB levels throughout age-groups was recorded, possible due to the fact that younger men tend to have higher levels of oxidative stress when compared to younger women (Kander et al., 2017). Secondly, a higher variability, quantified through the number of outliers, was observed in women for both LDLox and LDLox/ApoB in the 40–49 age group, while in men the 50–59 age group exhibited a higher variability in terms of LDLox and LDLox/ApoB median values. Thirdly, both in women and men a decrease in both LDLox and LDLox/ApoB was observed from the 60–69 age group to the 70–75 age group.

Table 3
Spearman's (*rho*) correlations with chronological age of representative MARK-AGE biomarkers specific for lipid metabolism – serum lipoproteins fractions and subfractions **(a)**, systemic inflammation, oxidative stress and epigenetic markers **(b)**, anthropometric and clinical parameters specific for glucose homeostasis and cardiometabolic risk evaluation **(c)**, in male (n=528) and female (n=561) RASIG participants.

Variable	Male (n = 528)	p	Female (n = 561)	p
a) **Lipid metabolism – serum lipoproteins fractions and subfractions**				
Total cholesterol, mg/dL	- 0.082	NS	**0.167**	< 0.001
Triglycerides, mg/dL	- 0.042	NS	**0.265**	< 0.001
HDL-cholesterol, mg/dL	**0.116**	< 0.01	- 0.017	NS
HDL2-triglycerides, mg/dL	- 0.035	NS	**0.156**	< 0.001
LDL-cholesterol, mg/dL	- 0.086	NS	**0.097**	< 0.05
LDL2-cholesterol, mg/dL	- 0.041	NS	**0.205**	< 0.001
LDL2-triglycerides, mg/dL	0.041	NS	**0.197**	< 0.001
VLDL-cholesterol, mg/dL	- 0.084	NS	**0.248**	< 0.001
VLDL2-cholesterol, mg/dL	- 0.066	NS	**0.253**	< 0.001
VLDL2-triglycerides, mg/dL	- 0.05	NS	**0.290**	< 0.001
b) **Systemic inflammation, oxidative stress and epigenetic markers**				
Fibrinogen, g/L	**0.23**	< 0.001	**0.346**	< 0.001
Ferritin, ng/mL	0.076	NS	**0.463**	< 0.001
hsCRP, μg/L	0.074	NS	**0.234**	< 0.001
Adiponectin, μg/mL	**0.103**	< 0.01	**0.188**	< 0.001
NOx(NO_2^- + NO_3^-), μmol/L	0.005	NS	**0.095**	< 0.05
LDLox, nmol MDA/mL	**- 0.103**	< 0.05	**0.099**	< 0.05
LDLox/ApoB	**- 0.13**	< 0.01	0.053	NS
Malondialdehyde, μmol/L	0.005	NS	**0.167**	< 0.001
α - Tocopherol, μmol/L	0.013	NS	**0.242**	< 0.001
γ - Tocopherol, μmol/L	0.018	NS	0.07	NS
*FHL2*_CpG_13,14,15	**0.471**	< 0.001	**0.494**	< 0.001
*FHL2*_CpG_16,17	**0.457**	< 0.001	**0.552**	< 0.001
c) **Anthropometric and clinical parameters specific for glucose homeostasis and cardiometabolic risk evaluation**				
TyG index	0.052	NS	**0.314**	< 0.001
HOMA-IR	**0.133**	< 0.001	**0.249**	< 0.001
Systolic BP, mmHg	**0.29**	< 0.001	**0.395**	< 0.001
Diastolic BP, mmHg	0.02	NS	**0.142**	< 0.001
Waist–to-hip ratio	**0.167**	< 0.001	**0.22**	< 0.001
BMI, kg/m^2	**0.116**	< 0.01	**0.256**	< 0.001
High blood pressure	**0.203**	< 0.001	**0.363**	< 0.001

n, number of subjects; NS – non-significant;
RASIG: randomly recruited age-stratified individuals from the general population.
Apo B, apolipoprotein B-100; BP, blood pressure; BMI, body mass index; HOMA-IR, homeostasis model assessment of insulin resistance; hsCRP, high-sensitivity C-reactive protein; HDL, high-density lipoprotein; LDL, low-density lipoprotein; VLDL, very low density lipoprotein; HDL1, LDL1, VLDL1 – large lipoprotein particles; HDL2, LDL2, VLDL2 – small lipoprotein particles; MDA, malondialdehyde; NOx, nitric oxide metabolic pathway products; *FHL2*, Four and a half LIM domains 2 gene methylation status of promoter-associated CpG islands; TyG index, the product of the levels of triglycerides and fasting plasma glucose.

3.4. Validation results for the machine learning algorithms

Table 4 summarizes the patient distribution for the three combined hypertension and obesity outcomes which were used (Cases a, b and c), Table 5 illustrates the Spearman's age-adjusted correlations between the three outcome thresholds and the common predictive variables, for the RASIG study population. while Tables 6 and 7 present the validation results for the predictive models. All two validation types (internal and external) are shown: the internal repeated cross validation on RASIG, and the external validation on GO and SGO populations. In addition, Table 8 and Fig. 4 illustrate the most relevant feature importance data. In Table 8, for each case, the cumulative importance by averaging the results obtained after implementing all three algorithms is shown. In Fig. 4, the comparative feature importance of the combined LDLox and HDL variables *versus* the classical HDL cholesterol-based variables are presented, illustrating the situations for which the combined predictors outperformed the HDL predictors in terms of feature importance, highlighting the relative importance of LDLox based predictive variables.

4. Discussion

The novelty of the MARK-AGE project was the assessment and the identification of the most representative biomarkers of ageing in subjects recruited randomly from the European general population (RASIG), thus in a sample population representative of the "real-world", common population, and not (apparently) "healthy" population. This original approach was based on the exclusive use of biological samples, without clinical evaluation, and involved age-matched and sex-matched comparisons between groups, for the identification of parameters indicating the biological differences, namely the biological age (Bürkle et al., 2015).

Within the MARK-AGE project a broad set of parameters were assessed as candidate biomarkers of ageing, but the present study focuses on LDL oxidation as a hallmark of atherosclerosis development and a candidate biomarker of vascular ageing and cardiometabolic risk assessment (Gradinaru et al., 2015).

The current state of the art methodology used to investigate the clinical significance of LDL oxidation involves sensitive immunoassays quantifying the circulating levels of oxidized LDL using specific monoclonal antibodies directed against unique oxidation-specific epitopes (Afonso and Spickett, 2019; de de de Mello Barros Pimentel et al., 2023). Alternatively, the *in vitro* estimation of LDL "oxidizability" following its selective isolation from plasma involves the measurements of the specific products of the lipid peroxidation chain reaction (conjugated dienes, lipid hydroperoxides, and aldehydes), after the exposure of LDL particles to a standard oxidative stress inducer (Esterbauer et al., 1989; Liu et al., 2002; Ahotupa et al., 1996; Scoccia et al., 2001; Aoki et al., 2012; Suzuki-Sugihara et al., 2016).

Given that cholesterol, phospholipids, polyunsaturated fatty acids and apolipoprotein B-100 constitute the LDL substrates for oxidation, these two methodologies should indicate the quantity of oxidized LDLs and the "quality" of lipoprotein particles, as oxidation is significantly influenced by the presence of lipophilic micronutrients and antioxidants such as: α- and γ-tocopherol, β-carotene and ubiquinol-10 in the lipoprotein particle (Esterbauer et al., 1992). Also, significant individual differences in oxidized LDL (oxLDL) and in oxidation susceptibility of LDL (LDLox) were reported, and such variability greatly depended on particle size and lipid composition (Kresanov et al., 2021; Moriyama, 2020; Kollar et al., 2021). Furthermore, none of the LDL oxidation biomarkers has so far fulfilled the essential criteria to be considered as a clinical "surrogate endpoint" of cardiovascular events, but fit within the concept of early vascular ageing markers (Vlachopoulos et al., 2015).

To evaluate LDLox in the MARK-AGE samples, LDL fractions were

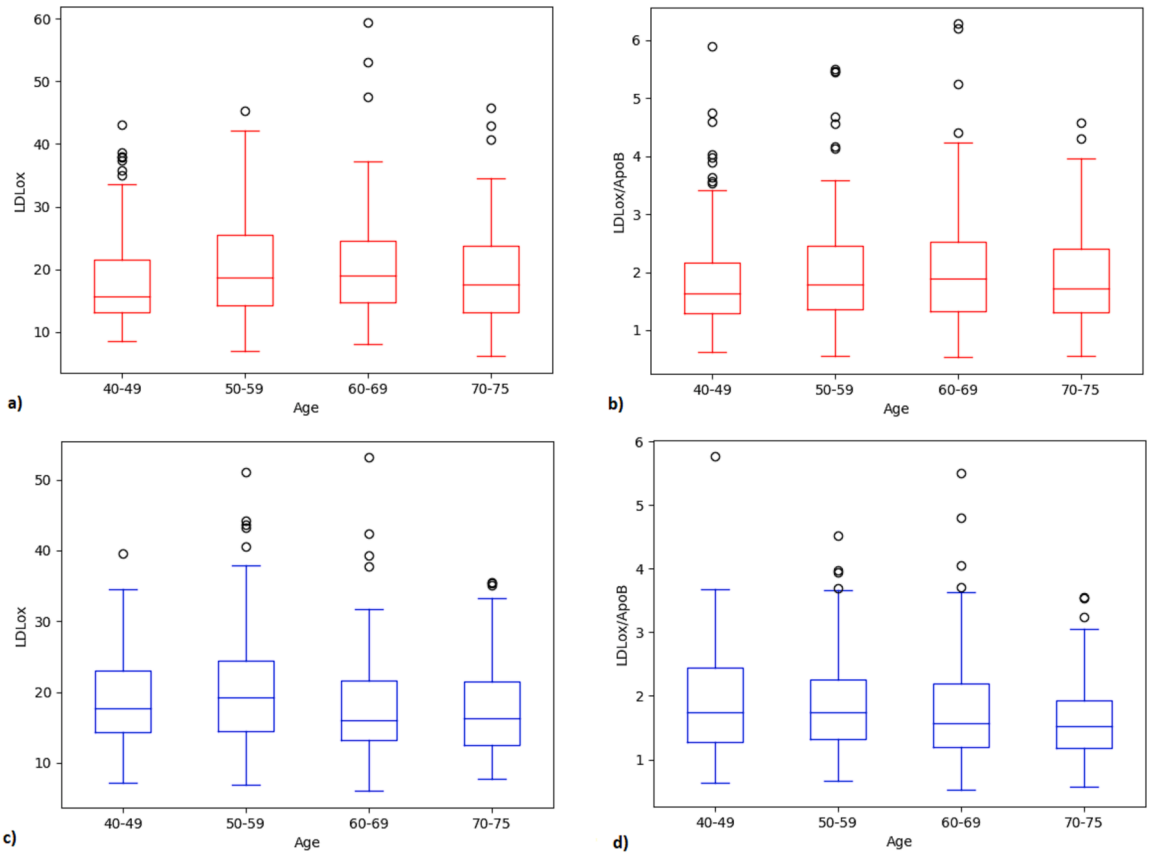


Fig. 3. Age groups related changes in LDL oxidizability, expressed as nmoles MDA/dL serum (LDLox) (**a** and **c**) and as LDLox/apoB ratio (**b** and **d**), in female **(a** and **b**, n=561) and male **(c** and **d**, n=528) RASIG subjects. The boxes show the median values, the 25th and 75th percentile. Whiskers indicate the 5th and 95th percentile. Outliers are displayed for each age-group: 40–49 (n= 161 women and n= 127 men); 50–59 (n= 151 women and n= 161 men); 60–69 (n= 154 women and n= 169 men) and 70–75 years (n= 95 women and n= 71 men).

Table 4
The patient distribution for each of the three created outcomes for the combined prediction models.

Prediction outcome	RASIG (1089 subjects)	GO (346 subjects)	SGO (192 subjects)
% HBP+WTHR (no.)	40.77 % (444)	46.24 % (160)	58.33 % (112)
% HBP+$BMI_{overweight}$ (no.)	36.27 % (395)	38.73 % (134)	53.65 % (103)
% HBP+BMI_{obese} (no.)	13.13 % (143)	15.03 % (52)	17.19 % (33)

isolated from serum samples and were subsequently challenged with a prooxidant stimulus capable of inducing an *in vitro* lipid peroxidation, which was evaluated and expressed as malondialdehyde (MDA) concentration (Ahotupa et al., 1996; Scoccia et al., 2001). As the *in vitro* oxidation of lipoproteins is often achieved by treatment with transition metal ions such as Cu^{2+} or Fe^{2+}, in the present experimental model the combination of Fe^{2+} with ascorbic acid could be consistent with the situation existing *in vivo* when in presence of transition metals, the ascorbic acid displays a pro-oxidant action (Esterbauer et al., 1989; Pinchuk and Lichtenberg, 2014; Ojo and Leake, 2021). The use of iron

Table 5
The the Spearman's (*rho*) correlation coefficients adjusted to chronological age between the three outcome thresholds (corresponding to Cases a, b and c) and the common predictive variables for all three cases (RASIG participants).

Variable	Case a (HTA+ WTHR)	p	Case b (HTA+ BMI)	p	Case c (HTA+ BMIobese)	p
LDLox	0.052	NS	0.051	NS	**0.066**	0.029
Risk_HDL-C+LDLox_12	**0.091**	0.003	**0.116**	<0.001	**0.125**	<0.001
Risk_HDL-C+LDLox_16	**0.083**	0.006	**0.107**	<0.001	**0.102**	0.001
Risk_HDL-C	**0.085**	0.005	**0.109**	<0.001	**0.119**	<0.001
Risk_LDLox/HDL-C	**0.155**	<0.001	**0.181**	<0.001	**0.161**	<0.001
FHL2_13,14,15+LDLox	0.051	NS	**0.071**	0.018	**0.076**	0.013
FHL2_16,17+LDLox	0.035	NS	0.031	NS	0.047	NS
Sex	**0.257**	<0.001	**0.209**	<0.001	**0.073**	0.015
Chronological age[a]	**0.335**	<0.001	**0.275**	<0.001	**0.111**	<0.001
HOMA-IR	**0.29**	<0.001	**0.384**	<0.001	**0.373**	<0.001
TyG	**0.288**	<0.001	**0.328**	<0.001	**0.271**	<0.001
AIP	**0.279**	<0.001	**0.332**	<0.001	**0.269**	<0.001
Adiponectin	**−0.236**	<0.001	**−0.266**	<0.001	**−0.157**	<0.001
Uric acid	**0.329**	<0.001	**0.365**	<0.001	**0.263**	<0.001
Smoking years	**0.094**	0.002	**0.089**	0.003	**0.067**	0.026

[a] The unadjusted Spearman's (*rho*) correlation coefficients are shown for chronological age; NS – non-significant. HDL-C, HDL-cholesterol

Table 6
Internal cross-validation results (RASIG population).

RASIG population – 1089 subjects			
Case a (HBP+WTHR)			
Validation measure	**MLP**	**SVM**	**XGB**
Accuracy	0.714	0.714	0.705
Precision	0.671	0.627	0.619
Recall	0.594	0.740	0.725
F1 Score	0.628	0.678	0.667
Matthews Corr.	0.401	0.430	0.411
ROC AUC Score	0.784	0.785	0.783
Case b (HBP+$BMI_{overweight}$)			
	MLP	**SVM**	**XGB**
Accuracy	0.728	0.711	0.735
Precision	0.667	0.584	0.616
Recall	0.514	0.718	0.727
F1 Score	0.577	0.643	0.665
Matthews Corr.	0.392	0.412	0.455
ROC AUC Score	0.793	0.792	0.808
Case c (HBP+BMI_{obese})			
	MLP	**SVM**	**XGB**
Accuracy	0.864	0.752	0.772
Precision	0.457	0.311	0.339
Recall	0.125	0.717	0.771
F1 Score	0.180	0.432	0.470
Matthews Corr.	0.174	0.349	0.401
ROC AUC Score	0.811	0.803	0.839

*MLP = Multilayer Perceptron, SVM = Support Vector Machines, XGB = XGBoost Classifier, Matthews Corr. = Matthews Correlation Coefficient

Table 7
External validation results (GO and SGO populations).

	GO population (346 subjects)			**SGO population (192 subjects)**		
	Case a (HBP+WTHR)			**Case a (HBP+WTHR)**		
Validation measure	**MLP**	**SVM**	**XGB**	**MLP**	**SVM**	**XGB**
Accuracy	0.630	0.630	0.618	0.609	0.656	0.661
Precision	0.614	0.584	0.566	0.664	0.672	0.677
Recall	0.538	0.694	0.750	0.670	0.804	0.804
F1 Score	0.573	0.634	0.645	0.667	0.732	0.735
Matthews Corr.	0.251	0.270	0.261	0.195	0.272	0.284
ROC AUC Score	0.693	0.705	0.673	0.659	0.654	0.648
	Case b (HBP+$BMI_{overweight}$)			**Case b (HBP+$BMI_{overweight}$)**		
	MLP	**SVM**	**XGB**	**MLP**	**SVM**	**XGB**
Accuracy	0.685	0.647	0.642	0.620	0.661	0.630
Precision	0.609	0.534	0.528	0.670	0.646	0.640
Recall	0.522	0.709	0.694	0.573	0.816	0.709
F1 Score	0.562	0.609	0.600	0.618	0.721	0.673
Matthews Corr.	0.321	0.309	0.295	0.247	0.319	0.252
ROC AUC Score	0.736	0.738	0.715	0.688	0.687	0.682
	Case c (HBP+BMI_{obese})			**Case c (HBP+BMI_{obese})**		
	MLP	**SVM**	**XGB**	**MLP**	**SVM**	**XGB**
Accuracy	0.858	0.740	0.676	0.823	0.667	0.641
Precision	0.565	0.327	0.258	0.455	0.299	0.275
Recall	0.250	0.692	0.615	0.152	0.697	0.667
F1 Score	0.347	0.444	0.364	0.227	0.418	0.389
Matthews Corr.	0.310	0.338	0.225	0.185	0.275	0.231
ROC AUC Score	0.751	0.787	0.730	0.739	0.759	0.711

*MLP = Multilayer Perceptron, SVM = Support Vector Machines, XGB = XGBoost Classifier, Matthews Corr. = Matthews Correlation Coefficient

also mimics the model for ferritin-induced oxidation of LDL (Afonso and Spickett, 2019; Martínez-Soto et al., 2021).

In the MARK-AGE study ferritin is a positive metabolic determinant for LDLox, along with fibrinogen, glucose and α-tocopherol, parameters found to be increased in ageing, particularly in women. A previous MARK-AGE study which analyzed a larger cohort (> 2200 participants) reported the positive association of α-tocopherol with chronological age, even when all covariates including cholesterol and use of vitamin supplements were included (Stuetz et al., 2016).

In the MARK-AGE population, the micronutrients α- and γ-tocopherol (the most acknowledged lipophilic antioxidants of lipoprotein particles) were found in both sexes to be strongly, positively correlated with LDLox, suggesting they could rather act as "pro-oxidant" molecules in the given experimental *in vitro* conditions. Although α-tocopherol is well known to inhibit LDL oxidation *in vitro* (more precisely the lag time), recent studies show that at lysosomal pH (in macrophages), α-tocopherol was unable to inhibit LDL oxidation induced by ferritin (Ojo and Leake, 2018). Moreover, it was suggested that dehydroascorbate (the oxidation product of ascorbate) could have in the presence of ferritin, at lysosomal pH, a pro-oxidant effect on partially oxidized LDLs (Ojo and Leake, 2021). Furthermore, lipophilic tocopherols generally interact with the hydrophilic ascorbic acid, which is missing in our experimental set-up. As it was debated in a recent analysis resulted from the MARK-AGE data, all the so called "low molecular weight antioxidants" - α- and γ-tocopherol, lycopene, retinol, and ascorbic acid - reflect nutritional patterns and do not automatically possess the capacity to counteract the oxidative stress exerted on specific lipoprotein particles, such as LDLs (Pinchuk et al., 2021). In the present study, the age- and sex-associations of these studied dietary antioxidants, as well as their relationships with other oxidative stress biomarkers (LDLox, MDA) display the same trend with the findings reported in other MARK-AGE studies, in which larger cohorts were analyzed (> 2200 participants) (Weber et al., 2017; Pinchuk et al., 2019).

In our studied sample population based on RASIG subjects aged 40–75 years old, LDLox appeared to be significantly positively correlated with chronological age in women and negatively correlated with ageing in men. Such sex-specific particularities regarding the ageing-related changes of LDL oxidizability could result from the differences existing among the male and female groups, as recorded for some clinical (SBP, DBP), anthropometric (BMI, WTHR) characteristics and additional measured MARK-AGE biomarkers (Table 1). Beyond the association of LDLox with cholesterol and triglycerides content in lipoprotein fractions and subfractions, the multiple correlation analysis identified numerous other sex-specific metabolic determinants for LDLox, which could also explain such differences in LDL oxidation profiles between men and women. And lastly, it was interesting to analyze, within the studied cohort, the age-related changes of the MARK-AGE parameters which were identified to be significantly correlated with LDLox.

Overall, the "ageing metabolic landscape" displays an impaired glucose homeostasis by pointing out significantly in women, an increase in TyG index and HOMA-IR with chronological age. Also, only in the female group was pointed out a significant increase with age in triglycerides and cholesterol levels in all lipoprotein fractions and subfractions (LDL, VLDL, LDL2, VLDL2 and HDL2) excepting the protective, anti-atherogenic HDL-cholesterol fraction. In other words, the LDL, VLDL and HDL2 lipoprotein particles content of specific substrates for lipid peroxidation increased with chronological age only in women. These age-related changes in the advanced lipoprotein profile could explain the ageing-related increases in the LDL oxidizability pointed out in women. Numerous studies demonstrated that lipoprotein-associated cardiovascular risk, besides that dependent on LDL-cholesterol, is determined by a cluster of metabolic abnormalities, referred to as the "atherogenic lipid phenotype", characterized by elevated triglycerides (TG)-rich lipoproteins levels, increased small dense LDL and decreased HDL (Grundle et al., 2021; Kollar et al., 2021). It is acknowledged that serum levels of VLDLs enriched in TG are metabolized predominantly to small, dense LDL; VLDL1 once secreted from the liver enters a delipidation cascade leading to the formation of smaller VLDL2, IDL, and LDL. LDL is derived from the delipidation of VLDL1 but the extent of conversion is lower than from VLDL2 (Packard et al., 2020). In this respect, small LDL (LDL2) and small HDL (HDL2) particles were associated with impaired glucose tolerance and reduced insulin sensitivity, and overall

Table 8
Average feature importance (%) results for each validation type and outcome.

	Internal cross-validation (RASIG population)			External validation (GO population)			External validation (SGO population)		
Variable	**Case a**	**Case b**	**Case c**	**Case a**	**Case b**	**Case c**	**Case a**	**Case b**	**Case c**
LDLox	0.14 %	0.23 %	1.06 %	0.18 %	0.26 %	0.01 %	0.22 %	0.00 %	0.00 %
Risk_HDL-C+LDLox_12	0.48 %	0.44 %	2.56 %	0.22 %	0.02 %	0.68 %	0.00 %	0.00 %	0.00 %
Risk_HDL-C+LDLox_16	0.09 %	0.10 %	0.75 %	0.00 %	0.03 %	2.83 %	0.35 %	2.14 %	3.28 %
Risk_HDL-C	0.08 %	0.07 %	0.58 %	0.00 %	0.00 %	0.26 %	0.15 %	0.35 %	2.21 %
Risk_LDLox/HDL-C	0.16 %	0.56 %	3.18 %	0.48 %	0.19 %	4.61 %	2.27 %	0.75 %	9.90 %
FHL2_13,14,15+LDLox	0.62 %	1.06 %	0.58 %	1.05 %	1.81 %	0.01 %	7.37 %	5.22 %	1.03 %
FHL2_16,17+LDLox	0.30 %	0.60 %	0.81 %	0.51 %	0.61 %	0.15 %	0.43 %	0.00 %	0.00 %
Sex	5.50 %	0.31 %	0.94 %	2.76 %	0.50 %	1.31 %	2.08 %	0.00 %	2.84 %
Chronological age	45.96 %	22.37 %	0.64 %	7.86 %	0.78 %	0.06 %	7.61 %	0.79 %	0.20 %
HOMA-IR	7.15 %	19.28 %	52.34 %	9.33 %	25.36 %	54.14 %	12.85 %	24.67 %	48.28 %
TyG	8.26 %	9.54 %	2.75 %	19.71 %	9.50 %	2.02 %	20.41 %	20.48 %	1.97 %
AIP	0.43 %	1.29 %	1.77 %	4.55 %	4.40 %	0.91 %	1.74 %	0.98 %	1.06 %
Adiponectin	1.03 %	1.78 %	0.51 %	3.84 %	4.54 %	0.80 %	0.15 %	0.16 %	0.91 %
Uric acid	21.79 %	34.27 %	18.21 %	26.20 %	40.06 %	13.65 %	19.05 %	23.51 %	3.98 %
Smoking years	0.14 %	0.08 %	0.17 %	0.00 %	0.64 %	0.18 %	0.00 %	0.07 %	1.10 %

*Case a = HPB+WTHR, Case b = HBP+$BMI_{overweight}$, Case c = HBP+$BMI_{obese;}$ HDL-C, HDL-cholesterol

associated with an atherogenic risk profile (Hämäläinen et al., 2018). Recently, Fernández-Cidón et al. demonstrated that subjects with residual risk of premature cardiovascular disease had higher concentrations of small dense LDL-cholesterol (sdLDL) and triglycerides (TG) in LDL and HDL particles (Fernández-Cidón et al., 2021). In observational analyses, a higher triglyceride composition within HDL subclasses was associated with higher risk of CHD, independently of total cholesterol and triglycerides (Kettunen et al., 2019). Also, it has been proposed that small HDL subclass (HDL2) exhibits atherogenic properties (dysfunctional HDL) whereas large and intermediate HDL subclasses are atheroprotective (Serban et al., 2014).

In women, other metabolic and systemic inflammation parameters that are positively associated with the increase in LDL oxidizability such as ferritin, fibrinogen and hsCRP, were also found to display a significant age-related increase. Serum ferritin express the body iron stores and is an acute phase reactant known to coordinate cellular defense against oxidative stress and inflammation (Galaris et al., 2019). Ferritin is assessed routinely in the diagnosis of anemia because low levels are a criterion for iron deficiency anemia. Its concentration increases in women after the menopause, with a concomitant increase in the risk for several diseases (Engelfriet et al., 2013). Recent experimental and clinical studies demonstrated the direct involvement of ferritin in LDL oxidation through reactive oxygen species generation and oxidative stress associated with hyperglycemia and hyperlipidemia (Martínez-Soto et al., 2021). In the MARK-AGE sample population, ferritin levels were higher in women than in men and increased significantly with age only in women. The ageing-related increase in all the pro-inflammatory, pro-thrombotic, pro-oxidative and pro-atherogenic biomarkers/factors such as: fibrinogen, ferritin, glucose, uremic environment and small-dense lipoprotein particles LDL2, seem to be counteracted by a single biomarker with anti-inflammatory and anti-atherosclerotic properties, namely the adiponectin. Notably, in women, the significantly age-related increased levels of pro-inflammatory parameters and LDL oxidizability were pointed out together with increased levels of plasma malondialdehyde (MDA), a less specific biomarker of lipid peroxidation. As regards the nitric oxide metabolic pathway products (NOx), a recent MARK-AGE study in the RASIG sample population evidenced that blood bacterial DNA levels were positively, significantly associated with plasma NOx levels (Giacconi et al., 2023).

By contrast, in male subjects' group there was a significant decrease in LDL oxidation biomarkers (LDLox and LDLox/apoB) with chronological age. Only in men LDLox was significantly negatively correlated with the protective, anti-atherogenic fraction HDL-cholesterol and HDL2-cholesterol. Although LDLox is positively correlated with fibrinogen and glucose levels/insulin resistance it should be noted that both fibrinogen and insulin resistance (HOMA-IR and TyG index) increase with chronological age are less pronounced in men *versus* women. LDLox was not significantly associated with ferritin levels but was negatively correlated with adiponectin. At a glance, it seems that all these more or less significant ageing-related metabolic differences between male and female subjects have made male subjects' LDLs particles less susceptible (vulnerable) to the oxidative stress exerted *in vitro*, in our experimental setting.

The present study evidenced in both sexes an age-related increase in adiponectin levels, and its protective effect against LDL oxidation. Data regarding the correlation between adiponectin levels and age are contradictory (Baker et al., 2019), but clinical studies have suggested a positive association of circulating adiponectin with healthspan and lifespan: for instance, hypoadiponectinemia was closely associated with type 2 diabetes, metabolic syndrome (MS), atherosclerotic cardiovascular diseases (CVD) and oxidative stress (Baker et al., 2019; Gradinaru et al., 2017). A decrease in adiponectin clearance in the kidney may be the cause of high levels of adiponectin in the elderly. Adiponectin level seems to be influenced strongly by blood urea nitrogen (BUN) and to be increased by a decline in renal function with ageing (Isobe et al., 2005). The MARK-AGE study also confirmed, in men and women, the age-related increase in serum urea concentration which might reflect an age-related decline in renal function (Bürkle et al., 2015).

LDL oxidation and lipid peroxidation are the result of metabolic stress to which the female body is subjected (probably following menopause) at systemic level and in the endothelial microenvironment: ferritin - and possible subsequent iron release, in certain conditions such as hyperglycemia, and hyperuremia. Our data are in line with recent findings from a cross-sectional and longitudinal study showing that transition from pre-menopausal period to post-menopause has effects on multiple circulating metabolic biomarkers, over and above the underlying age trajectory. The adverse changes in multiple apolipoprotein-B-containing lipoprotein subclasses, increased inflammation and oxidative stress may underlie women's increased cardiometabolic risk in their post-menopausal years (Wang et al., 2018; Pinchuk et al., 2019).

The MARK-AGE study identified the CpG islands of *FHL2* gene, whose methylation levels strongly correlates with chronological age in both sexes (Garagnani et al., 2012; Bürkle et al., 2015). In recent studies *FHL2* age-related hypermethylation is validated amongst the strongest biomarkers of biological age, in men and women (Han et al., 2020). In our sample population a significant negative association between LDL oxidizability and this epigenetic biomarker of ageing (*FHL2_CpG_16,17*) was pointed out only in female subjects. The four and a half LIM domains 2 (*FHL2*) gene encode for a transcriptional cofactor that acts as a scaffolding protein, and is expressed most abundantly in the heart and blood vessels (Habibe et al., 2021). *FHL2* hypermethylation and expression are

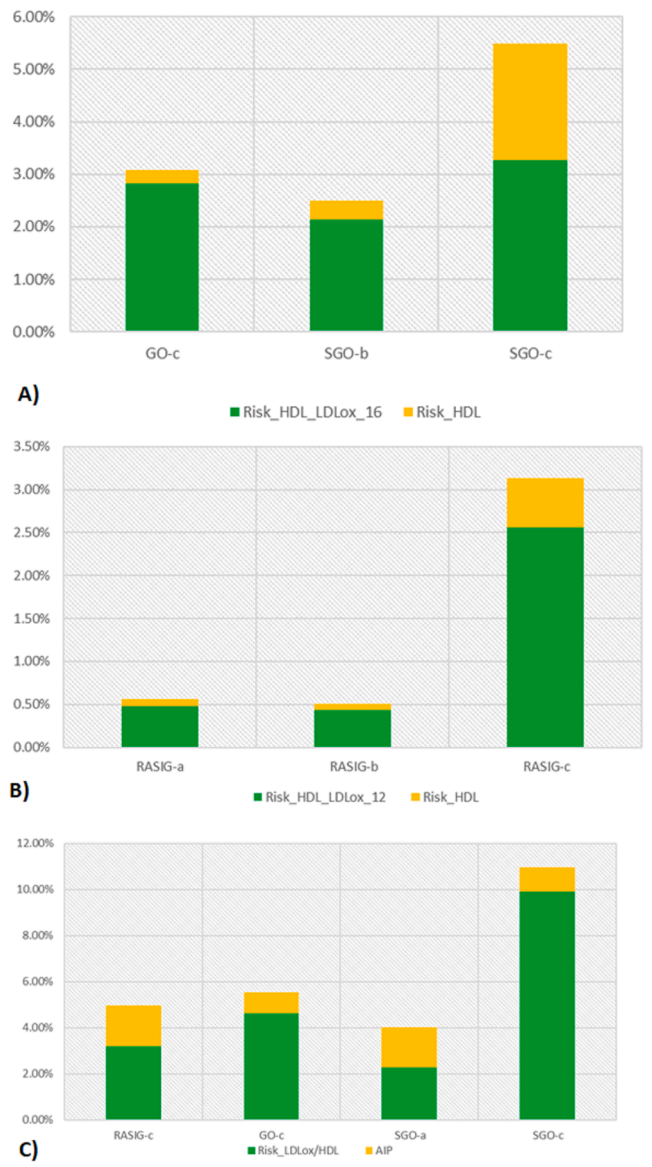


Fig. 4. Comparative feature importance of the combined LDLox and HDL cholesterol variables: *Risk_HDL-C+LDLox_16* **(A)**, *Risk_HDL-C+LDLox_12* **(B)** and *Risk_LDLox/HDL-C versus* the classical risk of HDL cholesterol **(A,B)** and Plasma Atherogenic Index (AIP) **(C)**, for the three combined prediction models with hypertension (HBP) and obesity outcomes (Cases a, b and c) in RASIG, GO and SGO population. Legend: a = HPB+WTHR, b = HBP+$BMI_{overweight}$, c = HBP+BMI_{obese}, RASIG = internal cross-validation, GO = external validation on GO population, SGO = external validation on SGO population.

increased in ageing, and in different pathologies such as cardiovascular dysfunction, vascular disease and obesity (Garagnani et al., 2012; Chen et al., 2020). In mice, *FHL2* in both myeloid and vascular cells may play an important role in atherosclerosis by promoting proinflammatory chemokine production, adhesion molecule expression, and proinflammatory monocyte recruitment (Ebrahimian et al., 2015). Functionally, in vascular smooth muscle cells the absence of *FHL2* (in knockout mice) resulted in attenuated cholesterol efflux to ApoA-1 and high-density lipoprotein (HDL), in agreement with an altered cholesterol synthesis and liver X receptor (LXR) signaling pathways (Kurakula et al., 2015).

In the present study, the different association between the hypermethylation of CpG sites (*FHL2_CpG_16,17* and *FHL2_CpG_13,14,15*) of the promoter region of *FHL2* and LDL oxidation could indicate the vascular endothelium shift during ageing to a pro-inflammatory, pro-oxidative and pro-atherosclerotic status/phenotype. As the LIM domains allow *FHL2* to interact with a variety of targets proteins (such as different transcription factors), it modulates a large number of functions of various tissues and their pathologies, therefore it could influence LDL oxidizability.

The correlation analysis which explored the relationships between ageing, LDLox and relevant MARK-AGE biomarkers proved that LDL susceptibility to oxidation is influenced - *in vitro* and *in vivo* - by a series of metabolic, prooxidant, proinflammatory and proatherogenic factors. Therefore, LDLox was further used in three prediction models - machine learning algorithms - together with other traditional clinical parameters such as BMI, WTHR, SBP, and DBP, with the aim of estimating the cardiometabolic risk, quantified through the combined status of high blood pressure (HBP) and obesity for the MARK-AGE subjects.

With regards to the results obtained in terms of machine learning algorithm validation (Tables 6 and 7), one of the most notable aspects which can be noticed is the enhanced predictive ability obtained during the internal repeated (10 times) 5-fold cross-validation on RASIG population (1089 subjects), as opposed to the external validation, performed on GO (346 subjects) and SGO (192 subjects) populations. For example, for Case a (HBA+WTHR), the computed internal cross-validation ROC AUC Score (which quantified how well calibrated are the probabilities estimated by the model – the ROC AUC Score can vary between 0 and 1) varied between 0.783 and 0.785 (with a maximum of 0.785 obtained for SVM algorithm), while the Matthews Correlation Coefficient (a less biased classification metric when compared to F1 Score; it should be noted that the Matthews Correlation Coefficient can take values from −1–1, with −1 meaning a completely opposite classification and 1 perfect classification) yielded values of 0.401–0.43 (with a maximum of 0.43 obtained for the SVM algorithm). On the other hand, the combined hypertension and BMI prediction models (Cases b and c) yielded higher ROC AUC Scores (for Case b – HPB+$BMI_{overweight}$: 0.792–0.808; for Case c – HBP+BMI_{obese}: 0.803–0.839; for both cases the XGB algorithm obtained the best performance), but variable Matthews Correlation Coefficients (Case b – HPB+$BMI_{overweight}$: 0.392–0.455; for Case c – HBP+BMI_{obese}: 0.174–0.401; for both cases the XGB algorithm obtained the best performance). Hence, overall, the internal cross-validation results performed on more than 1000 subjects from the MARK-AGE general population yielded ROC AUC Scores ranging from a minimum of 0.783 to a maximum of 0.839; it is reasonable to note that the computed values are satisfactory, especially considering the small dataset and the complex nature of the three outcomes: combined status of high blood pressure (at least stage 1) and overweight or obesity, quantified through waist-to-hip ratio and BMI (Leggio et al., 2017; Muhammad et al., 2022). Therefore, the main aim of the obtained algorithms could be to identify and stratify individuals with a high cardiometabolic risk, based on the specific probabilities estimated by the machine learning models (Pedregosa et al., 2011; Chen and Guestrin, 2016).

On the other hand, unlike for the internal cross-validation, for the GO and SGO external validation lower ROC AUC Scores were obtained. Indeed, for Case a (HBP+WTHR), Table 7 illustrates a ROC AUC Score in the range 0.673–0.705 for the validation on GO population (the maximum value – 0.705, was obtained with the SVM algorithm), while the validation on SGO yielded a 0.648–0.659 range (the maximum value – 0.659, was obtained with the MLP algorithm). Enhanced results were obtained when validating the estimated probabilities for the combined high blood pressure and BMI-quantified overweight (Case b) and obesity status (Case c). For Case b, the ROC AUC Score ranged from 0.715 to 0.738 (maximum value: SVM algorithm) for the GO external validation and from 0.682 to 0.688 (maximum value: MLP algorithm) for the SGO external validation respectively. Even though the dataset was more imbalanced when the combined status of HBP and BMI-quantified obesity was used as outcome (see Table 4 for details regarding the distribution of the outcomes for each of the three populations – RASIG, GO and SGO), optimized ROC AUC Scores were computed for Case c

(0.73–0.787 for GO external validation and 0.711–0.759 for SGO external validation; in both situations, the SVM algorithm led to the best results). As with the ROC AUC Score, the computed Matthews Correlation Coefficients were lower than for the RASIG internal validation, with overall values ranging from 0.225 to 0.338 for the GO external validation and from 0.185 to 0.319 for the SGO external validation.

Hence, taking into consideration the most important validation metrics which were computed, our results support the use of the SVM algorithm for estimating the risk of developing combined HBP and WTHR-quantified obesity in the general population, while the XGB algorithm might be better suited for the combined HBP and BMI-quantified overweight and obesity status risk estimation. The significantly lower validation results which were obtained for the external validation make it more difficult to support the use of such algorithms for estimating the cardiometabolic risk in GO and SGO populations, perhaps with the notable exception of the combined HBP and BMI-quantified obesity status, where the ROC AUC Scores consistently reached values above 0.75 for the SVM algorithm.

With regards to other published studies which implemented artificial intelligence for predicting hypertension and/or obesity in adults, it is worth mentioning that the vast majority of the research focused on the individual prediction of the two outcomes, hypertension (Fitriyani et al., 2019; Montagna et al., 2022; Islam et al., 2023; Nematollahi et al., 2023) or obesity (Thamrin et al., 2021; Jeon et al., 2023; Lin et al., 2023), while the obesity prediction models were based on BMI and not on WTHR. Even though most of the studies relied on bigger datasets (at least 10000 subjects) than the one from our study (approximately 1000 subjects), the obtained metrics showed a high variability and the majority of the published research did not reach superior results when compared to our study; for example, Jeon et al. reported 0.65–0.7 AUC for predicting obesity by using data from 21100 Korean subjects, while Thamrin et al. reported a 0.74–0.79 AUC for predicting obesity on 618898 Indonesian subjects (Thamrin et al., 2021; Jeon et al., 2023). In terms of relevant studies performed on European populations, Montagna et al. predicted hypertension by using data obtained from 20206 subjects and reported an AUC of 0.816 (Montagna et al., 2022). In addition, it should be mentioned that all the mentioned studies purely relied on a raw classification approach and did not implement probability prediction; therefore, it should be mentioned that our developed algorithms might be better suited for stratifying individuals with high cardiometabolic risk (Pedregosa et al., 2011; Leggio et al., 2017). Therefore, our study is the first to implement a machine learning prediction of combined HBP and obesity status risk, as well as to include the WTHR as relevant outcome in a machine learning algorithm; it has been suggested by several studies that WTHR is a better cardiovascular risk predictor than BMI (Cao et al., 2018; Harris, 2023).

By implementing a standardized methodology during the machine learning validation process, it was possible to extract the importance of each variable for the prediction of each of the three outcomes. The feature importance was assessed on unseen data by using the ROC AUC Score as evaluation metric, which lead to a less biased measure of the contribution of each predictive variable in terms of outcome probability estimation. Regarding the obtained results (Table 8, Fig. 4), several clinically relevant findings are worth being mentioned. First of all, for all validation types and cardiometabolic outcome, at least one of the combined thresholds based on LDLox ($\geq$12 and $\geq$16) and HDL cholesterol ($<$40 mg/dl for males or $<$50 mg/dl for females) (Risk_HDL-C_LDLox_12, Risk_HDL-C_LDLox_16) had a higher average predictive importance than the HDL-cholesterol threshold ($<$40 mg/dl for males or $<$50 mg/dl for females) (Risk_HDL-C). These findings suggest that the clinical context (such as the susceptibility to oxidation of LDL particles - LDLox) of the subjects with HDL cholesterol lower than the recommended thresholds could add valuable information for estimating their cardiometabolic risk, especially considering the fact that the relationship between HDL cholesterol and cardiovascular risk is considered non-linear (März et al., 2017). Secondly, in several situations, the ratio between LDLox and HDL cholesterol (Risk_LDLox/HDL-C) outperformed the classical AIP (the ratio between triglycerides and HDL cholesterol) in terms of predictive importance. It is worth mentioning that the ratio between LDLox and HDL reached the highest predictive relevance for Case c (HBP+BMI_{obese}) for all validation types (internal RASIG validation: 3.18 %; external GO validation: 4.61 %; external SGO validation: 9.90 %): these aspects might be explained by the fact that for subjects with obesity, oxidative stress is one of the most important factors, while obesity, high blood pressure and oxidative stress are all important risk factors for atherosclerosis and cardiovascular risk (Aoki et al., 2012; Leggio et al., 2017; Muhammad et al., 2022). In addition, other combined indices centered on the susceptibility of LDL particles to oxidation, such as the product between LDLox and the DNA methylation status of FHL2 promoter (on 13, 14 and 15 CpG sites), also showed important contributions to the overall prediction, especially for the external validation performed on SGO population (Case a – 7.37 %, Case b – 5.22 %). Since the FHL2 methylation status is an important epigenetic marker, with an incompletely understood role in lipoprotein oxidation and cardiovascular risk, our results deem further research on a larger number of subjects. Hence, the current study is the first to analyze the combined predictive value of epigenetic markers and lipoprotein parameters on cardiovascular risk (Habibe et al., 2021), highlighting their complex and patient specific role for the cardiometabolic risk assessment.

However, even though LDLox based variables reached in a number of specific situations higher prediction importance when compared to the classical risk biomarkers (such as HDL threshold and AIP), there were more cases when LDLox based variables were outperformed by several clinically relevant features. For example, the internal validation (RASIG population) yielded importance scores of under 1 % for all LDLox based variables for Case a (HBP + WTHR), while sex, age, HOMA-IR, TyG and uric acid all reached scores of over 5 % (hence at least 5 times higher) in that situation. In this clinical context, it is worth mentioning that Case a, due to defining obesity through WTHR, might be more relevant than Cases b and c (which defined overweight and obesity status through BMI) from a cardiovascular risk perspective (Cao et al., 2018; Harris, 2023). In addition, LDLox showed a high importance in specific situations mainly when combined with other biomarkers, such as HDL cholesterol or the DNA methylation status of FHL2 promoter, which emphasizes that the susceptibility to oxidation of LDL particles might have an impact on cardiometabolic risk only through some complex interactions involving specific gene and lipid metabolism dysregulations.

Nevertheless, a limitation of the developed and validated machine learning models would be the fact that we didn't include as predictive variables any information regarding pharmacological treatment and subjects' comorbidities. While such information would be relevant from a clinical viewpoint, the main aim of the current stage of this analysis was to assess the cardiometabolic risk based solely on candidate geromarkers and lifestyle factors.

Another important limitation of the current study is the relatively small number of predictive biomarkers. While it is reasonable to state that the inclusion of more relevant biomarkers would have lead to a better predictive ability of the machine learning models (Kaneko, 2024), it would have made more difficult the assessment of the importance of LDLox in relation to the other variables (Pedregosa et al., 2011). In addition, since future studies would aim at enlarging the cohort size and building an online platform for clinical use of the predictive algorithms, the inclusion of a too high number of biomarkers would be invaluable from a practical and economical point of view (Bajwa et al., 2021).

Thirdly, the feature importance assessment was undertaken based on one single relevant method, the permutation technique, which randomly shuffled the specific variable which was analyzed, in order to compute the average loss in the ROC AUC Score. While the average results which were obtained on all three algorithm types (MLP, SVM and XGB) for each outcome and validation cohort (RASIG, GO and SGO) can be hence considered as a standardized manner of comparing the relative

importances of the predictive variables, the main caveat in applying this method in the current study would be its inability of handling interactions between variables (Cava et al., 2020). Hence, it should be mentioned that, given the complex environment which regulates oxidative stress and cardiometabolic risk, it is reasonable to state that specific interactions between LDLox and other predictive biomarkers and/or lifestyle factors might occur (Gradinaru et al., 2015; Sverdlov et al., 2016).

The current study offers novel insights regarding the combined effects of LDL oxidation and other ageing markers on cardiometabolic risk assessment. The MARK-AGE project and the present study pointed out for several biomarkers strong sex- and age-related correlations, suggesting that biological differences between women and men extend beyond the hormonal differences. As a consequence, the vascular ageing could express sex-specific biomarkers and should address sex-specific challenges and health conditions. Therefore, future studies must aim to validate the obtained results on larger patient cohorts, in order to obtain reproducible clinical assessment models.

5. Conclusions

In the MARK-AGE study population, the LDL oxidizability (LDLox) might be regarded as a candidate geromarker that illustrates the cardiometabolic risk as a consequence of proatherogenic and prooxidant conditions existing at systemic level, resulting from an increased ageing-related metabolic stress (dyslipidemia, chronic hyperglycemia, hyperinsulinemia, hyperuricemia), inflammation and pro-thrombotic environments, which are also acknowledged to be commonly associated with hypertension and/or overweight/obesity. Despite the disadvantages of the implemented machine learning methods (small number of subjects and of predictive variables, as well as the lack of external validation on a group of subjects from the general population), the current research is the first to develop predictive models for estimating the personalized cardiometabolic risk defined as the overlapping hypertension and obesity status. Furthermore, the non-inferior predictive value of the combined LDLox and HDL-cholesterol indices when compared to the classical HDL-cholesterol based markers (sex-based thresholds and Atherogenic index) emphasizes the complex relationships between the various lipoprotein fractions within the vascular environment, as well as the possible justification of considering LDLox as a clinically relevant marker in cardiometabolic risk estimation, after confirmation by more large-scale studies. Overall, LDLox offers an additional advantage in the clinical assessment of cardiometabolic risk associated with ageing.

Funding

This work was supported by the European Commission (Project Full Name: European Study to Establish Biomarkers of Human Ageing; Project Acronym: MARK-AGE; Project No. 200880). The current research and publication were undertaken with the support of "Carol Davila" University of Medicine and Pharmacy, Bucharest.

CRediT authorship contribution statement

Efstathios S Gonos: Writing – review & editing, Investigation, Formal analysis. **Eugène HJM Jansen:** Writing – review & editing, Investigation, Formal analysis. **Daniela Gradinaru:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Martijn E T Dollé:** Writing – review & editing, Investigation, Formal analysis. **Tilman Grune:** Writing – review & editing, Investigation, Formal analysis. **María Moreno-Villanueva:** Writing – review & editing, Investigation, Formal analysis. **Claudio Franceschi:** Writing – review & editing, Investigation, Formal analysis. **Alexander Bürkle:** Writing – review & editing, Investigation, Formal analysis. **Jan Frank:** Writing – review & editing, Investigation, Formal analysis. **Grazyna Mosieniak:** Writing – review & editing, Investigation, Formal analysis. **Ewa Sikora:** Writing – review & editing, Investigation, Formal analysis. **Simone Fiegl:** Writing – review & editing, Investigation, Formal analysis. **Birgit Weinberger:** Writing – review & editing, Investigation, Formal analysis. **Beatrix Grubeck-Loebenstein:** Writing – review & editing, Investigation, Formal analysis. **Jürgen Bernhardt:** Writing – review & editing, Investigation, Formal analysis. **Chiara Pirazzini:** Writing – review & editing, Investigation, Formal analysis. **Paolo Garagnani:** Writing – review & editing, Investigation, Formal analysis. **Miriam Capri:** Writing – review & editing, Investigation, Formal analysis. **Florence Debacq-Chainiaux:** Writing – review & editing, Investigation, Formal analysis. **Olivier Toussaint:** Investigation, Formal analysis. **Wolfgang Stuetz:** Writing – review & editing, Investigation, Formal analysis. **Daniela Weber:** Writing – review & editing, Investigation, Formal analysis. **Denisa Margina:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis. **Nicolle Breusing:** Writing – review & editing, Investigation, Formal analysis. **Andrei Valeanu:** Writing – review & editing, Writing – original draft, Validation, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Duncan Talbot:** Writing – review & editing, Investigation, Formal analysis. **Eline P Slagboom:** Writing – review & editing, Investigation, Formal analysis. **Antti Hervonen:** Writing – review & editing, Investigation, Formal analysis. **Maria Giulia Bacalini:** Writing – review & editing, Investigation, Formal analysis.

Data Availability

Data will be made available on request.

Acknowledgements

The authors extend their thanks to all participants from the different study centres in MARK-AGE project. Furthermore, thanks are extended to all the working groups involved, particulary to the research group from „Ana Aslan" National Institute of Gerontology and Geriatrics, Bucharest, Romania: Claudia Borsa, Cristina Ionescu and Gianina Constantin, for their experimental support.

Studies in humans

All procedures were performed in compliance with relevant laws and institutional guidelines and have been approved by the appropriate institutional committee(s). The privacy rights of human subjects have been observed and informed consent was obtained for experimentation with human subjects.

Informed consent

Ethics committee approval and informed consent have been obtained.

Author contributions

All authors contributed to the design, acquisition of data, drafting the article and revising it critically for important intellectual content. All authors reviewed and approved the final version of the manuscript.

References

Afonso, C.B., Spickett, C.M., 2019. Lipoproteins as targets and markers of lipoxidation. Redox Biol. *23*, 101066. https://doi.org/10.1016/j.redox.2018.101066.

Ahotupa, M., Ruutu, M., Mäntylä, E., 1996. Simple methods of quantifying oxidation products and antioxidant potential of low density lipoproteins. Clin. Biochem. *29* (2), 139–144. https://doi.org/10.1016/0009-9120(95)02043-8.

Aoki, T., Abe, T., Yamada, E., Matsuto, T., Okada, M., 2012. Increased LDL susceptibility to oxidation accelerates future carotid artery atherosclerosis. Lipids Health Dis. *11*, 4. https://doi.org/10.1186/1476-511X-11-4.

Bajwa, J., Munir, U., Nori, A., Williams, B., 2021. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc. J. *8* (2), e188–e194. https://doi.org/10.7861/fhj.2021-0095.

Baker, J.F., Newman, A.B., Kanaya, A., Leonard, M.B., Zemel, B., Miljkovic, I., Long, J., Weber, D., Harris, T.B., 2019. The Adiponectin Paradox in the Elderly: Associations With Body Composition, Physical Functioning, and Mortality. J. Gerontol. Ser. A Biol. Sci. Med. Sci. 74 (2), 247–253. https://doi.org/10.1093/gerona/gly017.

Baur, J., Kötter, T., Moreno-Villanueva, M., Sindlinger, T., Berthold, M.R., Bürkle, A., Junk, M., 2015a. The MARK-AGE extended database: data integration and pre-processing. Mech. Ageing Dev. 151, 31–37. https://doi.org/10.1016/j.mad.2015.05.006.

Baur, J., Moreno-Villanueva, M., Kötter, T., Sindlinger, T., Bürkle, A., Berthold, M.R., Junk, M., 2015b. MARK-AGE data management: cleaning, exploration and visualization of data. Mech. Ageing Dev. *151*, 38–44. https://doi.org/10.1016/j.mad.2015.05.007.

Bürkle, A., Moreno-Villanueva, M., Bernhard, J., Blasco, M., Zondag, G., Hoeijmakers, J. H., Toussaint, O., Grubeck-Loebenstein, B., Mocchegiani, E., Collino, S., Gonos, E.S., Sikora, E., Gradinaru, D., Dollé, M., Salmon, M., Kristensen, P., Griffiths, H.R., Libert, C., Grune, T., Breusing, N., Aspinall, R., 2015. MARK-AGE biomarkers of ageing. Mech. Ageing Dev. 151, 2–12. https://doi.org/10.1016/j.mad.2015.03.006.

Cao, Q., Yu, S., Xiong, W., Li, Y., Li, H., Li, J., Li, F., 2018. Waist-hip ratio as a predictor of myocardial infarction risk: A systematic review and meta-analysis. Medicine *97* (30), e11639. https://doi.org/10.1097/MD.0000000000011639.

Capri, M., Moreno-Villanueva, M., Cevenini, E., Pini, E., Scurti, M., Borelli, V., Palmas, M.G., Zoli, M., Schön, C., Siepelmeyer, A., Bernhardt, J., Fiegl, S., Zondag, G., de Craen, A.J., Hervonen, A., Hurme, M., Sikora, E., Gonos, E.S., Voutetakis, K., Toussaint, O., Franceschi, C., 2015. MARK-AGE population: From the human model to new insights. Mech. Ageing Dev. *151*, 13–17. https://doi.org/10.1016/j.mad.2015.03.010.

Cava, W., Bauer, C., Moore, J.H., Pendergrass, S.A., 2020. Interpretation of machine learning predictions for patient outcomes in electronic health records. AMIA Annu. Symp. . Proc. AMIA Symp. *2019*, 572–581.

Chen, C.Y., Tsai, H.Y., Tsai, S.H., Chu, P.H., Huang, P.H., Chen, J.W., Lin, S.J., 2020. Deletion of the FHL2 gene attenuates intima-media thickening in a partially ligated carotid artery ligated mouse model. J. Cell. Mol. Med. *24* (1), 160–173. https://doi.org/10.1111/jcmm.14687.

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). New York, NY, USA: ACM. https://doi.org/10.1145/2939672.2939785.

Dobiásová, M., Frohlich, J., 2001. The plasma parameter log (TG/HDL-C) as an atherogenic index: correlation with lipoprotein particle size and esterification rate in apoB-lipoprotein-depleted plasma (FER(HDL). Clin. Biochem. *34* (7), 583–588. https://doi.org/10.1016/s0009-9120(01)00263-6.

Ebrahimian, T., Simon, D., Lemarié, C.A., Simeone, S., Heidari, M., Mann, K.K., Wassmann, S., Lehoux, S., 2015. Absence of Four-and-a-Half LIM Domain Protein 2 Decreases Atherosclerosis in ApoE-/- Mice. Arterioscler., Thromb., Vasc. Biol. *35* (5), 1190–1197. https://doi.org/10.1161/ATVBAHA.114.305071.

Engelfriet, P.M., Jansen, E.H., Picavet, H.S., Dollé, M.E., 2013. Biochemical markers of aging for longitudinal studies in humans. Epidemiol. Rev. *35* (1), 132–151. https://doi.org/10.1093/epirev/mxs011.

Esterbauer, H., Striegl, G., Puhl, H., Rotheneder, M., 1989. Continuous monitoring of in vitro oxidation of human low density lipoprotein. Free Radic. Res. Commun. *6* (1), 67–75. https://doi.org/10.3109/10715768909073429.

Esterbauer, H., Gebicki, J., Puhl, H., Jürgens, G., 1992. The role of lipid peroxidation and antioxidants in oxidative modification of LDL. Free Radic. Biol. Med. *13* (4), 341–390. https://doi.org/10.1016/0891-5849(92)90181-f.

Fernández-Cidón, B., Candás-Estébanez, B., Gil-Serret, M., Amigó, N., Corbella, E., Rodríguez-Sánchez, M.Á., Padró-Miquel, A., Brotons, C., Hernández-Mijares, A., Calmarza, P., Jarauta, E., Brea, A.J., Mauri, M., Guijarro, C., Vila, À., Valdivielso, P., Corbella, X., Pintó, X., 2021. Physicochemical Properties of Lipoproteins Assessed by Nuclear Magnetic Resonance as a Predictor of Premature Cardiovascular Disease. PRESARV-SEA Study. J. Clin. Med. *10* (7), 1379. https://doi.org/10.3390/jcm10071379.

Fitriyani, N.L., Syafrudin, M., Alfian, G., Rhee, J., 2019. Development of disease prediction model based on ensemble learning approach for diabetes and hypertension (doi:). IEEE Access 7, 144777–144789. https://doi.org/10.1109/ACCESS.2019.2945129.

Galaris, D., Barbouti, A., Pantopoulos, K., 2019. Iron homeostasis and oxidative stress: an intimate relationship. Biochim. Et. Biophys. Acta Mol. Cell Res. *1866* (12), 118535. https://doi.org/10.1016/j.bbamcr.2019.118535.

Garagnani, P., Bacalini, M.G., Pirazzini, C., Gori, D., Giuliani, C., Mari, D., Di Blasio, A. M., Gentilini, D., Vitale, G., Collino, S., Rezzi, S., Castellani, G., Capri, M., Salvioli, S., Franceschi, C., 2012. Methylation of ELOVL2 gene as a new epigenetic marker of age. Aging Cell *11* (6), 1132–1134. https://doi.org/10.1111/acel.12005.

Ghezzi, P., Jaquet, V., Marcucci, F., Schmidt, H.H.H.W., 2017. The oxidative stress theory of disease: levels of evidence and epistemological aspects. Br. J. Pharmacol. *174* (12), 1784–1796. https://doi.org/10.1111/bph.13544.

Giacconi, R., Malavolta, M., Bürkle, A., Moreno-Villanueva, M., Franceschi, C., Capri, M., Slagboom, P.E., Jansen, E.H.J.M., Dollé, M.E.T., Grune, T., Weber, D., Hervonen, A., Stuetz, W., Breusing, N., Ciccarone, F., Zampieri, M., Aversano, V., Caiafa, P., Formentini, L., Piacenza, F., Cardelli, M., 2019. Nutritional factors modulating alu methylation in an Italian sample from the mark-age study including offspring of healthy nonagenarians. Nutrients *11* (12), 2986. https://doi.org/10.3390/nu11122986.

Giacconi, R., D'Aquila, P., Malavolta, M., Piacenza, F., Bürkle, A., Villanueva, M.M., Dollé, M.E.T., Jansen, E., Grune, T., Gonos, E.S., Franceschi, C., Capri, M., Gradinaru, D., Grubeck-Loebenstein, B., Sikora, E., Stuetz, W., Weber, D., Toussaint, O., Debacq-Chainiaux, F., Hervonen, A., Provinciali, M., 2023. Bacterial DNAemia in older participants and nonagenarian offspring and association with redox biomarkers: results from MARK-AGE study. J. Gerontol. Ser. A, Biol. Sci. Med. Sci. 78 (1), 42–50. https://doi.org/10.1093/gerona/glac154.

Giampieri, E., Remondini, D., Bacalini, M.G., Garagnani, P., Pirazzini, C., Yani, S.L., Giuliani, C., Menichetti, G., Zironi, I., Sala, C., Capri, M., Franceschi, C., Bürkle, A., Castellani, G., 2015. Statistical strategies and stochastic predictive models for the MARK-AGE data. Mech. Ageing Dev. 151, 45–53. https://doi.org/10.1016/j.mad.2015.07.001.

Gradinaru, D., Margina, D., Borsa, C., 2009. In vitro studies regarding the antioxidant effects of procaine, Gerovital H3 and Aslavital. Rev. Roum. De. Chim. 54 (9), 761–766.

Gradinaru, D., Borsa, C., Ionescu, C., Prada, G.I., 2015. Oxidized LDL and NO synthesis–Biomarkers of endothelial dysfunction and ageing. Mech. Ageing Dev. *151*, 101–113. https://doi.org/10.1016/j.mad.2015.03.003.

Gradinaru, D., Margina, D., Borsa, C., Ionescu, C., Ilie, M., Costache, M., Dinischiotu, A., Prada, G.I., 2017. Adiponectin: possible link between metabolic stress and oxidative stress in the elderly. Aging Clin. Exp. Res. *29* (4), 621–629. https://doi.org/10.1007/s40520-016-0629-z.

Habibe, J.J., Clemente-Olivo, M.P., de Vries, C.J., 2021. How (Epi)genetic regulation of the LIM-domain protein FHL2 impacts multifactorial disease. Cells *10* (10), 2611. https://doi.org/10.3390/cells10102611.

Han, Y., Franzen, J., Stiehl, T., Gobs, M., Kuo, C.C., Nikolić, M., Hapala, J., Koop, B.E., Strathmann, K., Ritz-Timme, S., Wagner, W., 2020. New targeted approaches for epigenetic age predictions. BMC Biol. *18* (1), 71. https://doi.org/10.1186/s12915-020-00807-2.

Harris, E., 2023. Study: Waist-to-Hip Ratio Might Predict Mortality Better Than BMI. JAMA *330* (16), 1515–1516. https://doi.org/10.1001/jama.2023.19205.

Hämäläinen, P., Saltevo, J., Kautiainen, H., Mäntyselkä, P., Vanhala, M., 2018. Hemoglobin level and lipoprotein particle size. Lipids Health Dis. 17 (1), 10. https://doi.org/10.1186/s12944-018-0655-2.

Heijmans, B.T., Beekman, M., Houwing-Duistermaat, J.J., Cobain, M.R., Powell, J., Blauw, G.J., van der Ouderaa, F., Westendorp, R.G., Slagboom, P.E., 2006. Lipoprotein particle profiles mark familial and sporadic human longevity. PLoS Med. 3 (12), e495. https://doi.org/10.1371/journal.pmed.0030495.

Islam, M.M., Alam, M.J., Maniruzzaman, M., Ahmed, N.A.M.F., Ali, M.S., Rahman, M.J., Roy, D.C., 2023. Predicting the risk of hypertension using machine learning algorithms: a cross sectional study in Ethiopia. PloS One *18* (8), e0289613. https://doi.org/10.1371/journal.pone.0289613.

Isobe, T., Saitoh, S., Takagi, S., Takeuchi, H., Chiba, Y., Katoh, N., Shimamoto, K., 2005. Influence of gender, age and renal function on plasma adiponectin level: the Tanno and Sobetsu study. Eur. J. Endocrinol. *153* (1), 91–98. https://doi.org/10.1530/eje.1.01930.

Jansen, E., Beekhof, P., Cremers, J., Weinberger, B., Fiegl, S., Toussaint, O., Bernhard, J., Gonos, E., Capri, M., Franceschi, C., Sikora, E., Moreno-Villanueva, M., Breusing, N., Grune, T., Bürkle, A., Dollé, M.E., 2015. Quality control data of physiological and immunological biomarkers measured in serum and plasma. Mech. Ageing Dev. *151*, 54–59. https://doi.org/10.1016/j.mad.2015.06.004.

Jeon, J., Lee, S., Oh, C., 2023. Age-specific risk factors for the prediction of obesity using a machine learning approach. Front. Public Health *10*, 998782. https://doi.org/10.3389/fpubh.2022.998782.

Kananen, L., Enroth, L., Raitanen, J., Jylhävä, J., Bürkle, A., Moreno-Villanueva, M., Bernhardt, J., Toussaint, O., Grubeck-Loebenstein, B., Malavolta, M., Basso, A., Piacenza, F., Collino, S., Gonos, E.S., Sikora, E., Gradinaru, D., Jansen, E.H.J.M., Dollé, M.E.T., Salmon, M., Stuetz, W., Jylhä, M., 2021. Self-rated health in individuals with and without disease is associated with multiple biomarkers representing multiple biological domains. Sci. Rep. *11* (1), 6139. https://doi.org/10.1038/s41598-021-85668-7.

Kananen, L., Hurme, M., Bürkle, A., Moreno-Villanueva, M., Bernhardt, J., Debacq-Chainiaux, F., Grubeck-Loebenstein, B., Malavolta, M., Basso, A., Piacenza, F., Collino, S., Gonos, E.S., Sikora, E., Gradinaru, D., Jansen, E.H.J.M., Dollé, M.E.T., Salmon, M., Stuetz, W., Weber, D., Grune, T., Jylhävä, J., 2023. Circulating cell-free DNA in health and disease - the relationship to health behaviours, ageing phenotypes and metabolomics. GeroScience *45* (1), 85–103. https://doi.org/10.1007/s11357-022-00590-8.

Kander, M.C., Cui, Y., Liu, Z., 2017. Gender difference in oxidative stress: a new look at the mechanisms for cardiovascular diseases. J. Cell. Mol. Med. *21* (5), 1024–1032. https://doi.org/10.1111/jcmm.13038.

Kaneko, H., 2024. Clustering method for the construction of machine learning model with high predictive ability. Chemom. Intell. Lab. Syst. *246*, 105084. https://doi.org/10.1016/j.chemolab.2024.105084.

Kettunen, J., Holmes, M.V., Allara, E., Anufrieva, O., Ohukainen, P., Oliver-Williams, C., Wang, Q., Tillin, T., Hughes, A.D., Kähönen, M., Lehtimäki, T., Viikari, J., Raitakari, O.T., Salomaa, V., Järvelin, M.-R., Perola, M., Davey Smith, G., Chaturvedi, N., Danesh, J., Ala-Korpela, M., 2019. Lipoprotein signatures of cholesteryl ester transfer protein and HMG-CoA reductase inhibition. e3000572. Article e3000572 PLoS Biol. 17 (12). https://doi.org/10.1371/journal.pbio.3000572.

Kollar, B., Siarnik, P., Hluchanova, A., Klobucnikova, K., Mucska, I., Turcani, P., Paduchova, Z., Katrencikova, B., Janubova, M., Konarikova, K., Argalasova, L., Oravec, S., Zitnanova, I., 2021. The impact of sleep apnea syndrome on the altered lipid metabolism and the redox balance. Lipids Health Dis. *20* (1), 175. https://doi.org/10.1186/s12944-021-01604-8.

Kresanov, P., Mykkänen, J., Ahotupa, M., Ala-Korpela, M., Juonala, M., Kaikkonen, J., Kähönen, M., Lehtimäki, T., Vasankari, T., Viikari, J., Raitakari, O.T., 2021. The associations of oxidized lipoprotein lipids with lipoprotein subclass particle concentrations and their lipid compositions. The Cardiovascular Risk in Young Finns Study. Free Radic. Biol. Med. *162*, 225–232. https://doi.org/10.1016/j.freeradbiomed.2020.10.020.

Kurakula, K., Sommer, D., Sokolovic, M., Moerland, P.D., Scheij, S., van Loenen, P.B., Koenis, D.S., Zelcer, N., van Tiel, C.M., de Vries, C.J., 2015. LIM-only protein FHL2 is a positive regulator of liver X receptors in smooth muscle cells involved in lipid homeostasis. Mol. Cell. Biol. *35* (1), 52–62. https://doi.org/10.1128/MCB.00525-14.

Leggio, M., Lombardi, M., Caldarone, E., Severi, P., D'Emidio, S., Armeni, M., Bravi, V., Bendini, M.G., Mazza, A., 2017. The relationship between obesity and hypertension: an updated comprehensive overview on vicious twins. Hypertens. Res.: Off. J. Jpn. Soc. Hypertens. *40* (12), 947–963. https://doi.org/10.1038/hr.2017.75.

Lin, W., Shi, S., Huang, H., Wen, J., Chen, G., 2023. Predicting risk of obesity in overweight adults using interpretable machine learning algorithms. Front. Endocrinol. *14*, 1292167. https://doi.org/10.3389/fendo.2023.1292167.

Liu, M.L., Ylitalo, K., Vakkilainen, J., Nuotio, I., Valkonen, M., Lahdenperä, S., Viikari, J., Taskinen, M.R., 2002. Susceptibility of LDL to oxidation in vitro and antioxidant capacity in familial combined hyperlipidemia: comparison of subjects with different lipid phenotypes. Ann. Med. *34* (1), 48–54. https://doi.org/10.1080/078538902317338643.

Mach, F., Baigent, C., Catapano, A.L., Koskinas, K.C., Casula, M., Badimon, L., Chapman, M.J., De Backer, G.G., Delgado, V., Ference, B.A., Graham, I.M., Halliday, A., Landmesser, U., Mihaylova, B., Pedersen, T.R., Riccardi, G., Richter, D. J., Sabatine, M.S., Taskinen, M.R., Tokgozoglu, L., ESC Scientific Document Group, 2020. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. Eur. Heart J. *41* (1), 111–188. https://doi.org/10.1093/eurheartj/ehz455.

Martínez-Soto, J.M., Candia-Plata, M.D.C., López-Soto, L.F., Soto-Guzmán, J.A., Camacho-Villa, A.Y., Álvarez-Hernández, G., Mata-Pineda, A.L., Galván-Moroyoqui, J.M., 2021. Increased serum ferritin is associated with oxidized low-density lipoprotein in prediabetes patients: a pilot study. Heliyon *7* (4), e06720. https://doi.org/10.1016/j.heliyon.2021.e06720.

Matthews, D.R., Hosker, J.P., Rudenski, A.S., Naylor, B.A., Treacher, D.F., Turner, R.C., 1985. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia *28* (7), 412–419. https://doi.org/10.1007/BF00280883.

März, W., Kleber, M.E., Scharnagl, H., Speer, T., Zewinger, S., Ritsch, A., Parhofer, K.G., von Eckardstein, A., Landmesser, U., Laufs, U., 2017. HDL cholesterol: reappraisal of its clinical relevance. Clin. Res. Cardiol.: Off. J. Ger. Card. Soc. *106* (9), 663–675. https://doi.org/10.1007/s00392-017-1106-1.

de Mello Barros Pimentel, M.V., Bertolami, A., Fernandes, L.P., Barroso, L.P., Castro, I.A., 2023. Could a lipid oxidative biomarker be applied to improve risk stratification in the prevention of cardiovascular disease? Biomed. Pharmacother. = Biomedecine Pharmacother. *160*, 114345. https://doi.org/10.1016/j.biopha.2023.114345.

Montagna, S., Pengo, M.F., Ferretti, S., Borghi, C., Ferri, C., Grassi, G., Muiesan, M.L., Parati, G., 2022. Machine learning in hypertension detection: a study on world hypertension day data. J. Med. Syst. *47* (1), 1. https://doi.org/10.1007/s10916-022-01900-5.

Moreno-Villanueva, M., Kötter, T., Sindlinger, T., Baur, J., Oehlke, S., Bürkle, A., Berthold, M.R., 2015b. The MARK-AGE phenotypic database: structure and strategy. Mech. Ageing Dev. *151*, 26–30. https://doi.org/10.1016/j.mad.2015.03.005.

Moreno-Villanueva, M., Capri, M., Breusing, N., Siepelmeyer, A., Sevini, F., Ghezzo, A., de, Craen, A.J., Hervonen, A., Hurme, M., Schön, C., Grune, T., Franceschi, C., Bürkle, A., 2015a. MARK-AGE standard operating procedures (SOPs): a successful effort. Mech. Ageing Dev. *151*, 18–25. https://doi.org/10.1016/j.mad.2015.03.007.

Moreno-Villanueva, M., Bürkle, A., 2019. Epigenetic and redox biomarkers: novel insights from the MARK-AGE study. Mech. Ageing Dev. *177*, 128–134. https://doi.org/10.1016/j.mad.2018.06.006.

Moriyama, K., 2020. The association between the triglyceride to high-density lipoprotein cholesterol ratio and low-density lipoprotein subcLasses. Intern. Med. (Tokyo, Jpn.) *59* (21), 2661–2669. https://doi.org/10.2169/internalmedicine.4954-20.

Muhammad, T., Paul, R., Rashmi, R., Srivastava, S., 2022. Examining sex disparity in the association of waist circumference, waist-hip ratio and BMI with hypertension among older adults in India. Sci. Rep. *12* (1), 13117. https://doi.org/10.1038/s41598-022-17518-z'.

Nematollahi, M.A., Jahangiri, S., Asadollahi, A., Salimi, M., Dehghan, A., Mashayekh, M., Roshanzamir, M., Gholamabbas, G., Alizadehsani, R., Bazrafshan, M., Bazrafshan, H., Bazrafshan Drissi, H., Shariful Islam, S.M., 2023. Body composition predicts hypertension using machine learning methods: a cohort study. Sci. Rep. *13* (1), 6885. https://doi.org/10.1038/s41598-023-34127-6.

Ojo, O.O., Leake, D.S., 2018. Low density lipoprotein oxidation by ferritin at lysosomal pH. Chem. Phys. Lipids *217*, 51–57. https://doi.org/10.1016/j.chemphyslip.2018.09.016.

Ojo, O.O., Leake, D.S., 2021. Vitamins E and C do not effectively inhibit low density lipoprotein oxidation by ferritin at lysosomal pH. Free Radic. Res. *55* (5), 525–534. https://doi.org/10.1080/10715762.2021.1964494.

Packard, C.J., Boren, J., Taskinen, M.R., 2020. Causes and consequences of hypertriglyceridemia. Front. Endocrinol. *11*, 252. https://doi.org/10.3389/fendo.2020.00252.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn Res *12*, 2825–2830.

Pinchuk, I., Kohen, R., Stuetz, W., Weber, D., Franceschi, C., Capri, M., Hurme, M., Grubeck-Loebenstein, B., Schön, C., Bernhardt, J., Debacq-Chainiaux, F., Dollé, M.E. T., Jansen, E.H.J.M., Gonos, E.S., Sikora, E., Breusing, N., Gradinaru, D., Moreno-Villanueva, M., Bürkle, A., Grune, T., Lichtenberg, D., 2021. Do low molecular weight antioxidants contribute to the Protection against oxidative damage? The interrelation between oxidative stress and low molecular weight antioxidants based on data from the MARK-AGE study. Arch. Biochem. Biophys. *713*, 109061. https://doi.org/10.1016/j.abb.2021.109061.

Pinchuk, I., Weber, D., Kochlik, B., Stuetz, W., Toussaint, O., Debacq-Chainiaux, F., Dollé, M.E.T., Jansen, E.H.J.M., Gonos, E.S., Sikora, E., Breusing, N., Gradinaru, D., Sindlinger, T., Moreno-Villanueva, M., Bürkle, A., Grune, T., Lichtenberg, D., 2019. Gender- and age-dependencies of oxidative stress, as detected based on the steady state concentrations of different biomarkers in the MARK-AGE study. Redox Biol. *24*, 101204. https://doi.org/10.1016/j.redox.2019.101204.

Pinchuk, I., Lichtenberg, D., 2014. Analysis of the kinetics of lipid peroxidation in terms of characteristic time-points. Chem. Phys. Lipids *178*, 63–76. https://doi.org/10.1016/j.chemphyslip.2013.12.001.

Python Software Foundation (2021). Python language reference, version 3.9.2. Available at: ⟨http://www.python.org⟩ (accessed 24 July 2024).

Ramdas Nayak, V.K., Satheesh, P., Shenoy, M.T., Kalra, S., 2022. Triglyceride Glucose (TyG) Index: a surrogate biomarker of insulin resistance. Jpma. J. Pak. Med. Assoc. *72* (5), 986–988. https://doi.org/10.47391/JPMA.22-63.

Scoccia, A.E., Molinuevo, M.S., McCarthy, A.D., Cortizo, A.M., 2001. A simple method to assess tHE OXIDative susceptibility of low density lipoproteins. BMC Clin. Pathol. *1* (1), 1. https://doi.org/10.1186/1472-6890-1-1.

Serban, C., Muntean, D., Mikhailids, D.P., Toth, P.P., Banach, M., 2014. Dysfunctional HDL: the journey from savior to slayer. Clin. Lipidol. *9* (*1*), 49–59. https://doi.org/10.2217/clp.13.83.

Stuetz, W., Weber, D., Dollé, M.E., Jansen, E., Grubeck-Loebenstein, B., Fiegl, S., Toussaint, O., Bernhardt, J., Gonos, E.S., Franceschi, C., Sikora, E., Moreno-Villanueva, M., Breusing, N., Grune, T., Bürkle, A., 2016. Plasma carotenoids, tocopherols, and retinol in the age-stratified (35-74 Years) general population: a cross-sectional study in Six European countries. Nutrients *8* (10), 614. https://doi.org/10.3390/nu8100614.

Sverdlov, A.L., Figtree, G.A., Horowitz, J.D., Ngo, D.T., 2016. Interplay between oxidative stress and inflammation in cardiometabolic syndrome. Mediat. Inflamm. *2016*, 8254590. https://doi.org/10.1155/2016/8254590.

Suzuki-Sugihara, N., Kishimoto, Y., Saita, E., Taguchi, C., Kobayashi, M., Ichitani, M., Ukawa, Y., Sagesaka, Y.M., Suzuki, E., Kondo, K., 2016. Green tea catechins prevent low-density lipoprotein oxidation via their accumulation in low-density lipoprotein particles in humans. Nutr. Res. (N. Y., N. Y.) *36* (1), 16–23. https://doi.org/10.1016/j.nutres.2015.10.012.

Thamrin, S.A., Arsyad, D.S., Kuswanto, H., Lawi, A., Nasir, S., 2021. Predicting Obesity in Adults Using Machine Learning Techniques: An Analysis of Indonesian Basic Health Research 2018. Front. Nutr. *8*, 669155. https://doi.org/10.3389/fnut.2021.669155.

Vaarhorst, A.A., Beekman, M., Suchiman, E.H., van Heemst, D., Houwing-Duistermaat, J. J., Westendorp, R.G., Slagboom, P.E., Heijmans, B.T., Leiden Longevity Study (LLS) Group, 2011. Lipid metabolism in long-lived families: the Leiden Longevity Study. Age (Dordr., Neth.) *33* (2), 219–227. https://doi.org/10.1007/s11357-010-9172-6.

Vlachopoulos, C., Xaplanteris, P., Aboyans, V., Brodmann, M., Cífková, R., Cosentino, F., De Carlo, M., Gallino, A., Landmesser, U., Laurent, S., Lekakis, J., Mikhailidis, D.P., Naka, K.K., Protogerou, A.D., Rizzoni, D., Schmidt-Trucksäss, A., Van Bortel, L., Weber, T., Yamashina, A., Zimlichman, R., Townsend, R.R., 2015. The role of vascular biomarkers for primary and secondary prevention. A position paper from the European Society of Cardiology Working Group on peripheral circulation: Endorsed by the Association for Research into Arterial Structure and Physiology (ARTERY) Society. Atherosclerosis *241* (2), 507–532. https://doi.org/10.1016/j.atherosclerosis.2015.05.007.

Wang, Q., Ferreira, D.L.S., Nelson, S.M., Sattar, N., Ala-Korpela, M., Lawlor, D.A., 2018. Metabolic characterization of menopause: cross-sectional and longitudinal evidence. BMC Med. *16* (1), 17. https://doi.org/10.1186/s12916-018-1008-8.

Weber, D., Stuetz, W., Bernhard, W., Franz, A., Raith, M., Grune, T., Breusing, N., 2014. Oxidative stress markers and micronutrients in maternal and cord blood in relation to neonatal outcome. Eur. J. Clin. Nutr. *68* (2), 215–222. https://doi.org/10.1038/ejcn.2013.263.

Weber, D., Stuetz, W., Toussaint, O., Debacq-Chainiaux, F., Dollé, M.E.T., Jansen, E., Gonos, E.S., Franceschi, C., Sikora, E., Hervonen, A., Breusing, N., Sindlinger, T., Moreno-Villanueva, M., Bürkle, A., Grune, T., 2017. Associations between specific redox biomarkers and age in a large European cohort: the MARK-AGE project. Oxid. Med. Cell. Longev. *2017*, 1401452. https://doi.org/10.1155/2017/1401452.

Williams, B., Mancia, G., Spiering, W., Agabiti Rosei, E., Azizi, M., Burnier, M., Clement, D.L., Coca, A., de Simone, G., Dominiczak, A., Kahan, T., Mahfoud, F., Redon, J., Ruilope, L., Zanchetti, A., Kerins, M., Kjeldsen, S.E., Kreutz, R., Laurent, S., Lip, G.Y.H., ESC Scientific Document Group, 2018. 2018 ESC/ESH Guidelines for the management of arterial hypertension. Eur. Heart J. *39* (33), 3021–3104. https://doi.org/10.1093/eurheartj/ehy339.

World Health Organization (2008). Waist circumference and waist–hip ratio: report of a WHO expert consultation, Geneva, 8–11 December 2008. Available at: ⟨https://iris.who.int/bitstream/handle/10665/44583/9789241501491_eng.pdf?sequence=1⟩ (accessed 24 July 2024).

Yates, T., Khunti, K., Wilmot, E.G., Brady, E., Webb, D., Srinivasan, B., Henson, J., Talbot, D., Davies, M.J., 2012. Self-reported sitting time and markers of inflammation, insulin resistance, and adiposity. Am. J. Prev. Med. *42* (1), 1–7. https://doi.org/10.1016/j.amepre.2011.09.022.