



Published in final edited form as:

*Mol Nutr Food Res.* 2019 September ; 63(17): e1900140. doi:10.1002/mnfr.201900140.

## Plasma metabolites associated with frequent red wine consumption: a metabolomics approach within the PREDIMED study

Pablo Hernández-Alonso<sup>1,2,\*</sup>, Christopher Papandreou<sup>1,2,\*</sup>, Mònica Bulló<sup>1,2</sup>, Miguel Ruiz-Canela<sup>2,3,4</sup>, Courtney Dennis<sup>5</sup>, Amy Deik<sup>5</sup>, Dong D. Wang<sup>6</sup>, Marta Guasch-Ferré<sup>1,2,6</sup>, Edward Yu<sup>6</sup>, Estefanía Toledo<sup>2,3,4</sup>, Cristina Razquin<sup>2,3,4</sup>, Dolores Corella<sup>2,7</sup>, Ramon Estruch<sup>2,8</sup>, Emilio Ros<sup>2,9</sup>, Montserrat Fitó<sup>2,10</sup>, Fernando Arós<sup>2,11</sup>, Miquel Fiol<sup>2,12</sup>, Lluís Serra-Majem<sup>2,13</sup>, Liming Liang<sup>14</sup>, Clary B. Clish<sup>5</sup>, Miguel A Martínez-González<sup>2,3,4,6</sup>, Frank B Hu<sup>5,14,15</sup>, Jordi Salas-Salvadó<sup>1,2</sup>

<sup>1</sup>Human Nutrition Unit, Faculty of Medicine and Health Sciences, Sant Joan Hospital, Institut d'Investigació Sanitària Pere Virgili, Rovira i Virgili University, Reus, Spain

<sup>2</sup>CIBER Fisiopatología de la Obesidad y Nutrición (CIBEROBN), Instituto de Salud Carlos III, Madrid, Spain

<sup>3</sup>University of Navarra, Department of Preventive Medicine and Public Health, Pamplona, Spain

<sup>4</sup>Navarra Institute for Health Research (IdisNA), Pamplona, Navarra, Spain.

<sup>5</sup>Broad Institute of MIT and Harvard University, Cambridge, MA, USA

<sup>6</sup>Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA

<sup>7</sup>Department of Preventive Medicine, University of Valencia, Valencia, Spain

<sup>8</sup>Department of Internal Medicine, Department of Endocrinology and Nutrition Institut d'Investigacions Biomèdiques August Pi Sunyer (IDI-BAPS), Hospital Clinic, University of Barcelona, Barcelona, Spain

<sup>9</sup>Lipid Clinic, Department of Endocrinology and Nutrition Institut d'Investigacions Biomèdiques August Pi Sunyer (IDI-BAPS), Hospital Clinic, University of Barcelona, Barcelona, Spain

**Address correspondence and reprint requests to:** Mònica Bulló, BSc, PhD, and Jordi Salas-Salvadó, MD, PhD, Human Nutrition Unit, Faculty of Medicine and Health Sciences, Universitat Rovira i Virgili, St/Sant Llorenç 21, 43201 Reus, Spain (monica.bullo@urv.cat, jordi.salas@urv.cat).

\*These authors contributed equally to this work.

Author contributions:

FH, JS-S and MM-G designed research; PH-A, CP, MB, MR-C, CD, AD, DW, MG-F, EY, ET, CR, DC, RE, ER, MF, FA, MFiol, LS-M, LL, CC, MM-G, FH and JS-S conducted research; DC, RE, ER, MF, FA, MFiol, LS-M, MM-G and JS-S were the coordinators of subject recruitment at the outpatient clinics; PH-A, CP and MB analyzed the data; PH-A, CP, MB, FBH and JS-S interpreted statistical analysis and data; CC, CD and AD acquired and processed metabolomics data; PH-A and CP drafted the paper; FH, JS-S and MM-G supervised the study and MB and JS-S took the responsibility for the integrity of the data and the accuracy of the data analysis. All authors revised the manuscript for important intellectual content, read and approved the final manuscript.

Conflict of interest statement:

JS-S is a non-paid member of the Scientific Committee of the International Nut and Dried Fruit Foundation. He has received grants/research support from the American Pistachio Growers and International Nut and Dried Fruit Foundation through his Institution. He has received honoraria from Nuts for Life, Danone and Eroski. He reports personal fees from Danone. He is a member of the executive committee of the Instituto Danone Spain. I certify that any of my other co-authors have a conflict of interest that is relevant to the subject matter or materials included in this Work.

<sup>10</sup>Cardiovascular and Nutrition Research Group, Institut de Recerca Hospital del Mar, Barcelona, Spain

<sup>11</sup>Department of Cardiology, University Hospital of Alava, Vitoria, Spain

<sup>12</sup>Institute of Health Sciences IUNICS, University of Balearic Islands and Hospital Son Espases, Palma de Mallorca, Spain

<sup>13</sup>Research Institute of Biomedical and Health Sciences UIBS, University of Las Palmas de Gran Canaria, Las Palmas, Spain

<sup>14</sup>Departments of Epidemiology and Statistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA

<sup>15</sup>Channing Division for Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, MA, USA

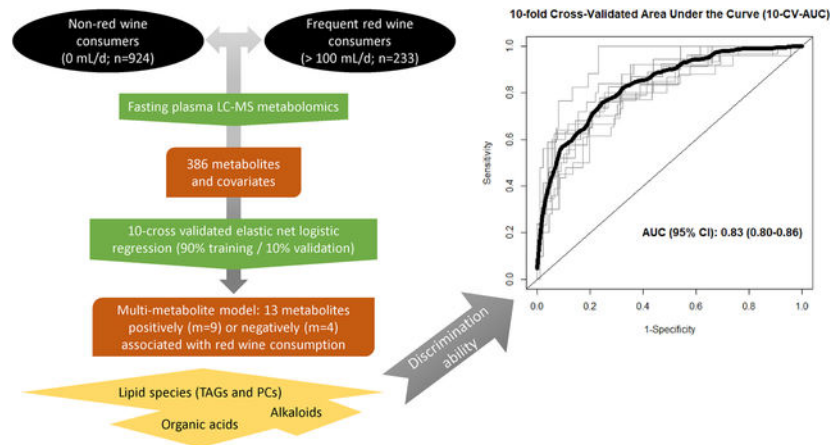
## Abstract

**Scope:** The relationship between red wine (RW) consumption and metabolism is poorly understood. We aimed to assess the systemic metabolomic profiles in relation to frequent RW consumption as well as the ability of a set of metabolites to discriminate RW consumers.

**Methods and Results:** Cross-sectional analysis of 1,157 participants. Subjects were divided as non-RW consumers versus RW consumers (> 1 glass/day RW (100 mL/day)). Plasma metabolomics analysis was performed using LC-MS. Associations between 386 identified metabolites and RW consumption were assessed using elastic net regression analysis taking into consideration baseline significant covariates. Ten-cross-validation (CV) was performed and receiver operating characteristic curves were constructed in each of the validation datasets based on weighted models. A subset of 13 metabolites was consistently selected and discriminated RW consumers versus non-consumers. Based on the multi-metabolite model weighted with the regression coefficients of metabolites, the area under the curve was 0.83 (95% CI: 0.80–0.86). These metabolites mainly consisted of lipid species (e.g. triglycerides and phosphatidylcholines), some organic acids and alkaloids.

**Conclusions:** A multi-metabolite model identified in a Mediterranean population appeared useful to discriminate between frequent RW consumers and non-consumers. Further studies are needed to assess the contribution of these metabolites in health and disease.

## Graphical Abstract



The relationship between red wine (RW) consumption and metabolism is poorly understood. The aim of this study was to assess the circulating metabolomic profiles in relation to frequent RW consumption as well as the ability of a set of metabolites to discriminate RW consumers from non-consumers. Our validated 13-multi-metabolite model accurately discriminates those who frequently consume RW from non-RW consumers.

## Keywords

metabolomics; red wine; lipidomics; metabolites; LC-MS

## 1. INTRODUCTION

Epidemiological studies suggest that light to moderate red wine (RW) consumption is inversely associated with different outcomes such as metabolic syndrome [1], type 2 diabetes (T2D) [2] and diverse cardiovascular outcomes [3]. However, even though data points to a putative modulatory role of moderate RW consumption on lipid and glucose metabolism, together with effects on inflammatory and endothelial parameters [4], the evidence coming from randomized clinical trials (RCT) is still controversial (reviewed in [5]).

An important issue arisen from the evaluation of RW intake in these studies is the lack of knowledge about the systemic metabolic modulation accompanying RW consumption. In order to accurately monitor dietary consumption in RCTs and epidemiological studies, there is a need for reliable biomarkers. In this regard, previous small clinical trials have mainly focused on polyphenols, measured either in urine or plasma, to predict RW consumption [6–9]. The main concern of these predicting models is that: i) they are based on wine components (e.g. resveratrol) which are also present in equal or higher amounts in other food matrices and their intake may differ among populations; ii) large molecular weight polyphenols from food, including certain polyphenols from RW, are poorly absorbed in humans [10]; and iii) their discovery is based on a high daily intake of wine. For example, Urpi-Sarda et al. first analyzed both urine and plasma samples in 36 male subjects consuming 272 mL of RW to measure a set of 70 host and microbial phenolic metabolites

[7]. In a more recent study[8], in which 41 healthy subjects consumed 250mL/d RW for 28 days, 94 compounds were linked to wine consumption comprising wine components, microbial-derived phenolic metabolites and endogenous compounds. However, the set of metabolites modulated by RW consumption might implicate other important and previously unexplored molecules and pathways.

To date, no comprehensive metabolomics analysis has been performed using different metabolomics platforms to analyse a diverse set of metabolites and including a relatively large sample size of participants. We hypothesized that the plasma metabolomic profile differs between frequent RW consumers and non-consumers. Using a multi-platform metabolomics analysis, we examined the associations between circulating metabolites and RW consumption. We further investigated the set of metabolites that best discriminate RW consumers from non-consumers.

## 2. MATERIAL AND METHODS

This study is a cross-sectional evaluation of baseline data from two published nested case-cohort studies aiming to explore the risk of CVD or T2D by means of a plasma metabolomics modulation [11, 12] within the PREDIMED trial (ISRCTN35739639). The PREDIMED study is a large clinical trial carried out in Spain, aiming to assess the effects of the traditional Mediterranean diet (MedDiet) on the primary prevention of cardiovascular disease (CVD) in a population at high risk of CVD [13]. Recruitment took place between June 2003 and June 2009. The trial protocol was in accordance with the Helsinki Declaration and was approved by the institutional review boards of all the centers involved. All participants provided written informed consent. Participants were men (55–80 years) and women (60–80 years) without CVD at baseline and fulfilling at least one of the two following criteria: presence of T2D or three or more major CV risk factors: current smoking, hypertension, high low-density lipoprotein (LDL)-cholesterol, low high-density lipoprotein (HDL)-cholesterol, overweight or obesity, and family history of premature CVD.

### 2.1 Assessment of population characteristics and dietary habits

Body mass index (BMI) was calculated as weight divided by height squared ( $\text{kg}/\text{m}^2$ ). Waist circumference (WC) was measured midway between the lowest rib and the iliac crest using an anthropometric tape. Dietary habits at baseline were evaluated using a validated 137-item semi quantitative food frequency questionnaire (FFQ) [14]. Daily food and nutrient intakes were estimated from the FFQ by multiplying the frequency of consumption by the average portion size. Participants also filled out a general questionnaire on lifestyle habits, medication use and concurrent diseases, and a validated Spanish version of the Minnesota Leisure Time Physical Activity Questionnaire [15].

### 2.2 Subjects' classification according to wine consumption

For the current study, participants were selected from the CVD [11] and T2D [12] case-cohort PREDIMED grants (NIH-NHLBI-5R01HL118264 and NIH-NIDDK-1R01DK102896). At baseline, a semi-quantitative FFQ was used to evaluate the foods consumed during the preceding year. The FFQ included questions concerning the

usual intake of different types of alcoholic beverages including RW. Importantly, the validity and reproducibility for total alcohol and total wine consumption has been previously reported [14, 16]. The intraclass correlation coefficient (ICC) between RW intake from the FFQ and repeated food records was 0.64 (unpublished results), whereas it was 0.82 for total alcohol intake [14]. The Pearson correlation coefficient between total alcohol intake (g/d) and RW (mL/d) in our analysis was 0.86 (95% CI: 0.85–0.88). Categories for RW consumption in the FFQ were as follows: i) never/almost never (i.e. < 1 serving/month); ii) 1–3 servings/month; iii) 1 serving/week; iv) 2–4 servings per week; v) 5–6 servings per week; vi) 1 serving/day; vii) 2–3 servings/day; viii) 4–6 servings/day and ix) >6 servings/day. Due to the semi-quantitative basis of the FFQ, categories of consumption were created instead of using the semi-continuous variable. A portion of 10g of pure alcohol was used to categorize RW intake, thus 100 ml of wine is referred to as one glass or drink. The participants were categorized into three groups based on their RW intake: 0 drinks/d (non-RW consumers (NRWC)); 0 < drinks/d 1; > 1 drink/d (RW consumer (RWC)). In order to increase the discrimination power of the multi-variable model, the middle category (n = 677) was excluded from the analysis.

### 2.3 Plasma metabolomics

Fasting (for 8 hours) plasma EDTA samples were collected from subjects and stored at  $-80^{\circ}\text{C}$ . Samples for each participant were randomly ordered and analyzed using two liquid chromatography tandem mass spectrometry (LC-MS) methods to measure polar metabolites and lipids as described previously [17–19]. Briefly, amino acids (AA) and other polar metabolites were profiled a Shimadzu Nexera X2 U-HPLC (Shimadzu Corp.) coupled to a Q-Exactive mass spectrometer (ThermoFisher Scientific). Metabolites were extracted from plasma (10  $\mu\text{L}$ ) using 90  $\mu\text{L}$  of 74.9:24.9:0.2 (vol/vol/vol) of acetonitrile/methanol/formic acid containing stable isotope-labeled internal standards [valine-d8 (Sigma-Aldrich) and phenylalanine-d8 (Cambridge Isotope Laboratories)]. The samples were centrifuged (10 min; 9000 x g;  $4^{\circ}\text{C}$ ), and the supernatants were injected directly on to a  $150 \times 2\text{-mm}$ , 3- $\mu\text{m}$  Atlantis HILIC column (Waters). The column was eluted isocratically at a flow rate of 250  $\mu\text{L}/\text{min}$  with 5% mobile phase A (10 mmol ammonium formate/L and 0.1% formic acid in water) for 0.5 min followed by a linear gradient to 40% mobile phase B (acetonitrile with 0.1% formic acid) over 10 min. MS analyses were carried out using electrospray ionization in the positive-ion and full-scan spectra were acquired over 70–800 m/z. Lipids were profiled using a Shimadzu Nexera X2 U-HPLC (Shimadzu Corp.; Marlborough, MA) coupled to an Exactive Plus orbitrap mass spectrometer (Thermo Fisher Scientific; Waltham, MA). Lipids were extracted from plasma (10  $\mu\text{L}$ ) using 190  $\mu\text{L}$  of isopropanol containing 1,2-didodecanoyl-sn-glycero-3-phosphocholine (Avanti Polar Lipids; Alabaster, AL) as an internal standard. Lipid extracts (2  $\mu\text{L}$ ) were injected onto a  $100 \times 2.1\text{ mm}$ , 1.7  $\mu\text{m}$  ACQUITY BEH C8 column (Waters; Milford, MA). The column was eluted isocratically with 80% mobile phase A (95:5:0.1 vol/vol/vol 10mM ammonium acetate/methanol/formic acid) for 1 minute followed by a linear gradient to 80% mobile-phase B (99.9:0.1 vol/vol methanol/formic acid) over 2 minutes, a linear gradient to 100% mobile phase B over 7 minutes, then 3 minutes at 100% mobile-phase B. MS analyses were carried out using electrospray ionization in the positive ion mode using full scan analysis over 200–1100 m/z. Raw data were processed using Trace Finder version 3.1 and 3.3 (Thermo Fisher Scientific)

and Progenesis QI (Nonlinear Dynamics; Newcastle upon Tyne, UK). Polar metabolite identities were confirmed using authentic reference standards and lipids were identified by head group and total acyl carbon number and total acyl double bond content. To enable assessment of data quality and to facilitate data standardization across the analytical queue and sample batches, pairs of pooled plasma reference samples were analyzed at intervals of 20 study samples. One sample from each pair of pooled references served as a passive QC sample to evaluate the analytical reproducibility for measurement of each metabolite while the other pooled sample was used to standardized at using a “nearest neighbour” approach. Standardized values were calculated using the ratio of the value in each sample over the nearest pooled plasma reference multiplied by the median value measured across the pooled references.

## 2.4 Statistical analysis

Baseline characteristics of study participants were described as means and standard deviations (SD) for quantitative variables, and percentages for categorical variables. Missing values of individual metabolites were imputed (in those metabolites with less than 20% of missing values) using the random forest imputation approach (“missForest” function from the “randomForest” R package) as it has been previously recommended in metabolomics studies [20, 21]. Importantly, different alternatives to this approach were found to generate consistent results (Supplementary Table 4). Missing values correspond to those determinations that were below the limit of detection. To conduct the multivariate analysis, metabolomic data was first centered and scaled using the standard deviation as the scaling factor (i.e. autoscaling) [22]. Due to the high dimensionality and collinear nature of the data, logistic regression with elastic net penalty (implemented in the “glmnet” R package) was used to build a predictive model for frequent RWC. We conducted the analysis including both metabolites and covariates that were relevant for our analysis. We performed a 10 cross-validation (CV) approach, splitting the sample into training (90% of the sample) and validation set 10 independent times, and then within the training set we performed a further 10-fold CV to find the optimal value of the tuning parameter [ $\lambda$  (lambda)] that yielded the minimum mean-squared error (MSE). The values minMSE and minMSE + 1SE were calculated using argument  $s = \text{“lambda.min”}$  or  $s = \text{“lambda.1se”}$  in the `cv.glmnet` function (“glmnet” R package). In order to report the coefficients from each CV iteration we evaluated the lambda selection in the elastic net logistic regression. Even though  $s = \text{“lambda.min”}$  gives the minimum mean CV error, we used  $s = \text{“lambda.1se”}$  in order to select a simpler model (reduced retained predictors). In fact, this model cannot be distinguished from the best model in terms of error given the uncertainty in the k-fold CV estimate of the error of the best model. Weighted models were constructed for each training-validation dataset pair (90% training and 10% validation) using solely the coefficients for the metabolites and other covariates obtained from each elastic net regression in the training set, and a receiver operating characteristic (ROC) curve was then created using the validation dataset pair (including only metabolites or metabolites and selected covariates). Because few participants consumed more than one glass of wine on a daily basis, a skewed consumption frequency pattern emerged [23], resulting in a different number of samples in the sub-cohort study tested for each group. In this context, the ROC curve is a non-parametric measure of biomarker utility and there is no need for the two distributions to have an equal number of

individuals and equal variance [24]. For reproducibility purposes, regression coefficients are reported using 10 iterations of the 10-CV elastic regression approach in the whole dataset. Different sensitivity analysis were additionally performed: a) Main model excluding those subjects with RW consumption exceeding the general recommendations (250 mL/d) b) Main model including relevant dietary factors (i.e. macronutrients and certain food groups); c) Main model excluding smokers. Importantly, after conducting a sensitivity analysis splitting the population into non-alcohol consumers (0 mL/d) versus non-RW alcohol consumers (>100 mL/d) the model did not identify any positively or negatively associated metabolite. All the analyses were performed using R v. 3.4.2 statistical software. A two-sided *P*-value less than 0.05 was considered significant.

### 3. RESULTS

A total of 1,157 PREDIMED study participants (436 men and 721 women) were included in the present study. Supplementary Figure 1 shows the flow chart of study participants. Baseline characteristics of the participants according to categories of RW intake at baseline are summarized in Table 1. After categorizing the RW consumption, 924 participants were included in the NRWC (0 mL/d) and 233 in the RWC (> 1 glass/day). Only 15% of the subjects included in the RWC group consumed more than 250 mL/d. RWC group compared to the NRWC had a higher proportion of men, a lower mean age and BMI, but a higher WC, lower cholesterol and triglycerides (TAG) levels, and had a higher prevalence of current smoking habit. Compared to NRWC, RWC had a higher total energy, carbohydrates, fat and protein intake. Considering main food groups, RWC also reported to have a higher red meat consumption, whereas a lower intake of fruits and vegetables versus NRWC.

#### 3.1 Multi-metabolite main model and discrimination of RW consumption

From the 399 metabolites analyzed in the present study, 13 metabolites were removed due to high number of missing values (i.e. >20%), thus 386 metabolites were finally included in all the analysis. Based on the baseline characteristics of the participants, we included in the main model for the elastic net logistic regression the following variables: age, sex (male/female), smoking status (yes/no), sub-cohort (CVD/T2D/both), case/control, BMI, WC, circulating baseline cholesterol and TAGs. Only sex, age and smoking status were retained at least once out of the 10 iterations. Values for metabolites' and covariates' mean, SD and the times being selected in each iteration are shown in Supplementary Table 1. Moreover, Figure 1 shows the mean coefficient value (and SD) for those 13 metabolites consistently selected in the 10 CV. The vast majority of the metabolites persistently included in the models belong to lipid species (i.e. TAGs, cholesteryl esters (CE), phosphatidylcholines (PC) and sphingomyelins (SM)), some organic acids and alkaloids. Those metabolites with the highest negative coefficient value were C51:3 TAG and C18:0 CE, whereas those with the highest positive coefficient value were C52:1 TAG and C34:1 PC. Figure 2 shows ROC curves for each of the 10 iterations together with the 10-fold cross-validated ROC curve. The AUC value for the 10-fold cross-validated ROC curve discriminating RW consumption was 0.83 (95% CI: 0.80–0.86; Figure 2) in case of the multi-metabolite model excluding covariates. The non-significantly higher 10-CV-AUC value after including covariates (i.e. sex, age and smoking status) was 0.88 (95% CI: 0.86–0.91, *P* = 0.0729).

### Sensitivity analysis

**a) Setting maximum RW intake at 250 mL/d:** The sensitivity analysis excluding those subjects with RW consumption >250mL/d showed a similar pattern of metabolites positively and negatively associated with RW consumption, even though more covariates (i.e. WC, baseline TAGs, BMI and baseline cholesterol) were included in the CV model (Supplementary Figure 2). The AUC value for discriminating RWC versus NRWC was of 0.85 (95% CI: 0.83–0.88), non-significantly higher compared to the model considering all the ranges of RW intake.

**b) Adjusting for dietary factors:** The following dietary factors were additionally included into the main model: intake of total fat, saturated fatty acids, monounsaturated fatty acids, polyunsaturated fatty acids, carbohydrates, protein, total energy, non-RW alcohol, red meat and food groups with reported high amount of polyphenols [25] (coffee (caffeinated and decaffeinated), fruits, vegetables, olives and olive oil and nuts)). We found that none of these factors were selected as predictors (Supplementary Table 2). Moreover, the consistently selected metabolites and other covariates remained invariant, together with the overall magnitude and variability of their coefficients. The predictive capability of the model was comparable when considering these covariates (0.89 (0.87–0.91)), but non-significantly higher compared to the main model when removing covariates (0.85 (0.82–0.87), Supplementary Table 3).

**c) Excluding smokers:** Exclusion of smokers reduced the sample size to 824 NRWC and 151 RWC (11% and 35% of reduction, respectively). Only sex was selected as a covariate (Supplementary Table 2). Cotinine was not included into the model and some other metabolites were less consistently selected or were not selected at all. Importantly, for those that were selected, the coefficients did not differ in magnitude to the main model and the AUCs were non-significantly lower compared to the main model (Supplementary Table 3).

## 4. DISCUSSION

In the present analysis, we have identified a set of 36 plasma metabolites mainly related to lipid species (PCs, TAGs, SMs), positively or negatively associated with frequent RW consumption. Out of them, 13 metabolites were consistently selected in the CV procedure (i.e. 10 times). The identified multi-metabolite model appeared useful to discriminate between frequent RW consumers and non-consumers.

Biomarkers of total wine and/or RW consumption, such as polyphenols, have been mostly identified in urine [6–9], and scarce data comes from plasma and/or serum samples [7]. In 2009, Zamora-Ros and collaborators defined resveratrol metabolites in urine as biomarkers of total wine intake after applying LC-MS/MS to urine samples from 1,000 participants [9]. They reported an AUC of 0.98 (95% CI: 0.97–0.99) for discrimination of total wine intake (mainly based on red wine and based on a cut-off point for urinary resveratrol metabolites of 411.38 nmol/g). However, resveratrol is not only present in RW. Other types of wine, red grapes, peanuts and cocoa powder also contain appreciable amounts of resveratrol [26]. Therefore, the application of a resveratrol-based predictive model, with a single metabolite as predictor, into populations with a frequent consumption of the aforementioned alternative



sources of resveratrol may reduce their predictive value [27]. In addition, the lack of internal validation in that study may have led to the overfitting of the model.

Interestingly, Urpi-Sardà and collaborators designed a crossover RCT for 4 weeks with 36 male volunteers at high risk of CVD to assess biomarkers of chronic RW consumption (272 mL/d). An analysis of metabolites in urine and plasma was performed by these authors using an UPLC-MS/MS, and most of the identified metabolites were related to resveratrol, but also gallic acid and ethylgallate behaved as predictors. The AUCs ranged from 0.91–0.98 and 0.74–0.91 in urine and plasma, respectively [7]. More recently, Esteban-Fernández and collaborators also evaluated the 24h-urinary metabolome using untargeted UHPLC-TOF MS of 41 healthy subjects consuming moderate RW consumption (250 mL/d for 28 days). They found 94 compounds related to high intake of RW consumption, including wine components (tartrate), microbial-derived phenolic metabolites (4-hydroxyl-5-(phenyl)-valeric acid) and endogenous compounds [8]. In our study, we have found an inverse association between the levels of  $\alpha$ -keto- $\delta$ -(N(G),N(G)-dimethylguanidino)valeric acid (DMGV) and RW consumption. Interestingly, baseline levels of this type of valeric acid were described as a biomarker of hepatic steatosis and independently predicted incident T2D up to 12 years before disease onset in 3 different cohorts [18]. As no previous findings have related RW consumption with DMGV in plasma, further research is needed to identify this relationship. Importantly, these studies considered higher and narrower threshold for wine consumption compared to our analysis. However, our sensitivity analysis excluding those subjects consuming >250mL/d of RW (i.e. more than 2.5 glasses of RW per day) also exhibited a good discrimination accuracy that did not differ from the analysis considering the whole population.

A recent review of clinical trials assessing the effects of wine consumption revealed a significant impact on lipid metabolism mostly attributed to their ethanol content [5]. Although the beneficial effect of moderate alcohol consumption on HDL metabolism is well described, the effect of wine consumption on TAGs, LDL, VLDL, and lipoprotein(a) is still under debate [5]. In fact, a J-shaped relationship between alcohol intake and plasma TAGs has been consistently reported [28], but no studies have ever reported any association between plasma TAG species with RW consumption. In our study, the highest negative coefficient identified was found for C51:3 TAG. Even though we may hypothesize that C51:3 TAG represents a glycerol backbone esterified to three fatty acids (FA) with 17 carbons and 1 double bond each, other FA combinations cannot be excluded. Interestingly, a positive association with RW consumption was reported for TAGs with even number of carbons (i.e. C52:1, C58:7 and C56:5).

The negative association observed in our study between C18:0 CE levels and frequent RW consumption may be due to the effect of alcohol consumption on CE transfer protein (CETP) activity [29]. Plasma CE, synthesized within HDL, may be transferred from HDL particles to the apolipoprotein B by plasma CETP. However, in our study we also reported C20:5 and C16:1 CEs to be positively associated with RW consumption, thus a complex regulation may be taking place. Alcohol consumption is associated with increased HDL cholesterol concentration and reduced plasma CETP activity [29]. Remarkably, in our study, we found no differences in HDL cholesterol levels according to RW consumption categories,

maybe due to the low reported RW consumption. Previous findings have shown that chronic alcohol consumption is positively associated with most lysophosphatidylcholines (LPC) [30], which have a direct role in toxic inflammatory responses in a variety of organ systems [31]. Therefore, the lack of any positive association between LPCs and RW consumption requires further investigation.

The positive association between L-carnitine and RW consumption could indicate the role of ethanol on  $\beta$ -oxidation [32]. Furthermore, RW consumption appears to affect the carnitine system in a more complex way as we reported carnitines to be positively (C9 carnitine) and negatively (C18:2 carnitine) associated with RW consumption. Noticeably, even we further adjusted the elastic net logistic regression model for red meat consumption, carnitine was consistently selected and included in the multi-metabolite model. Ethanol also inhibits the Krebs cycle by decreasing pyruvate levels. A reduction in pyruvate levels leads to a reduction of Krebs cycle intermediates such as isocitrate and succinate [33] that could explain the negative association between isocitrate levels and RW consumption that we have reported [33].

Zheng *et al.* found different serum metabolomic profiles related to alcohol consumption in 1,977 African Americans from the Atherosclerosis Risk in Communities (ARIC) study. Even with different types of specific metabolites, their results mirror ours pointing to a greater modulation of lipid and some AA species and their derivatives [30]. Contrary to their results, we found no associations with LPCs, whereas we reported different positive/negative associations with some species of TAGs. These differences could be related to the modulatory effect of the non-alcoholic fraction of red wine (i.e. polyphenols, glycerol and polysaccharides) [34].

Cotinine, caffeine, inositol and piperine were found to be positively associated with RWC. Cotinine is an alkaloid found in tobacco and is also the predominant metabolite of nicotine [35]. Our results pointing to a positive association between cotinine and frequent RW consumption – even after adjusting for smoking status – may be explained by a residual effect, as those frequently consuming RW also have a higher prevalence of smoking. In fact, in our sensitivity model excluding smokers, the metabolite pattern was consistent with the main model but cotinine was not included in the model. Importantly, recent studies have reported that those subjects with higher alcohol consumption also had a higher caffeine consumption from coffee beverages, partially explaining the positive association between circulating caffeine levels and RW consumption [36]. Whether the positive association between RW consumption and plasma inositol is due to inositol usage as preservative of wine is a hypothesis that deserves further investigation [37]. Interestingly, a previous study assessing metabolites significantly associated with total alcohol consumption also found piperine (i.e. an alkaloid derived from black pepper) to be positively associated with alcohol consumption [30].

This approach has some drawbacks that deserve comment. First, as it has been performed in older adults at high CV risk from a Mediterranean area, the generalizability of the findings to other populations may be limited. Moreover, due to its cross-sectional design, causation cannot be inferred. Due to a low number of female subjects into the RWC category, we were

not able to assess a sex-specific metabolomics pattern. Finally, even though we included in the analysis a relatively large sample size that was analyzed using a validated FFQ, we cannot exclude misclassification bias. On the other hand, we have used a multi-metabolomics approach in order to analyze a wide range of biochemical compounds; we have cross-internally validated our results; we have used a relatively low threshold for RW consumption making our model applicable to the general population; and we have performed different sensitivity analysis to assess the role of other putative confounders such as smoking status, dietary factors and non-RW alcohol intake into the selected metabolites.

In conclusion, our findings show that frequent RW consumption is associated with a set of plasma metabolites mainly related to lipid species and other organic compounds, which may have an impact in metabolic pathways related to disease development or prevention. Moreover, we have demonstrated that multi-metabolite models might accurately discriminate frequent RW consumption. In the current study, using a combination of different metabolomics platforms to cover a wide range of metabolites and examine their association with RW consumption, we provided a deeper understanding of the metabolic response to RW providing new functional insight to its role in health.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments:

We thank the participants for their enthusiastic collaboration, the PREDIMED personnel for excellent assistance, and the personnel of all affiliated primary care centers.

Funding sources:

This study was funded by the National Institutes of Health (R01DK102896,F31DK114938), the Spanish Ministry of Health (Instituto de Salud Carlos III) and the Ministerio de Economía y Competitividad-Fondo Europeo de Desarrollo Regional(Projects CNIC-06/2007, RTIC G03/140, CIBER 06/03, PI06-1326, PI07-0954,PI11/02505, SAF2009-12304 and AGL2010-22319-C03-03) and by the Generalitat Valenciana (ACOMP2010-181, AP111/10, AP-042/11, ACOM2011/145,ACOMP/2012/190, ACOMP/2013/159 and ACOMP/213/165). Dr. CP was supported by a postdoctoral fellowship granted by the Autonomous Government of Catalonia (PERIS 2016-2020 Incorporació de Científics i Tecnòlegs, SLT002/0016/00428). Dr. MG-F was supported by EFSO (European Foundation for the Study of Diabetes)/Lilly through the Institut d'Investigacions Sanitàries Pere I Virgili (IISPV).

## Abbreviations:

<b>AA</b>	amino acid
<b>ARIC</b>	Atherosclerosis Risk in Communities
<b>AUC</b>	area under the curve
<b>CE</b>	cholesteryl esters
<b>CETP</b>	cholesteryl esters transfer protein
<b>CV</b>	cross-validation
<b>CVD</b>	cardiovascular disease

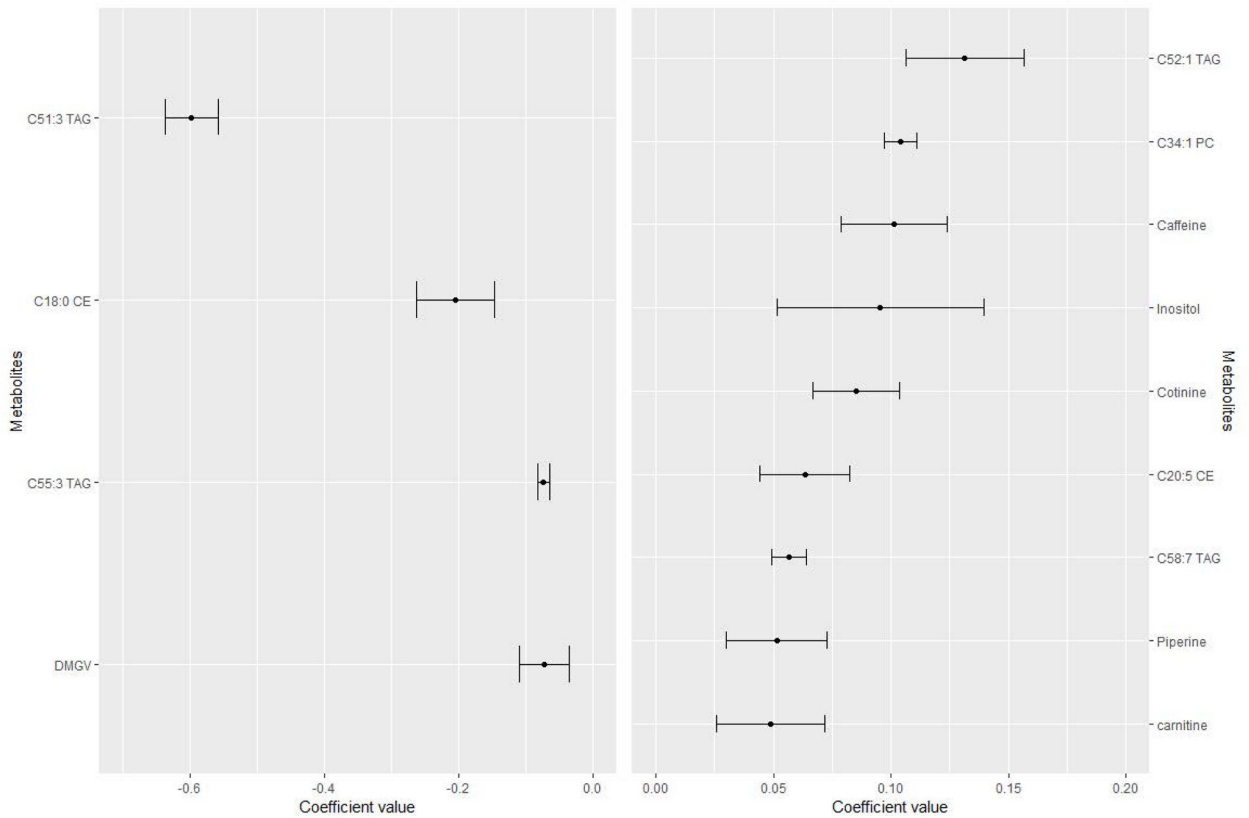
<b>DMGV</b>	$\alpha$ -keto- $\delta$ -(N(G),N(G)-dimethylguanidino)valeric acid
<b>FA</b>	fatty acid
<b>FFQ</b>	food frequency questionnaire
<b>ICC</b>	intraclass correlation coefficient
<b>LPC</b>	lysophosphatidylcholines
<b>MSE</b>	mean-squared error
<b>NRWC</b>	non-red wine consumers
<b>PC</b>	phosphatidylcholines
<b>RCT</b>	randomized clinical trial
<b>ROC</b>	receiver operating characteristic
<b>RW</b>	red wine
<b>RWC</b>	red wine consumers
<b>SM</b>	sphingomyelins
<b>T2D</b>	type 2 diabetes
<b>TAG</b>	triglyceride
<b>WC</b>	waist circumference

## 5. REFERENCES

- [1]. Tresserra-Rimbau A, Medina-Remón A, Lamuela-Raventós RM, Bulló M, Salas-Salvadó J, Corella D, Fitó M, Gea A, Gómez-Gracia E, Lapetra J, Arós F, Fiol M, Ros E, Serra-Majem L, Pintó X, Muñoz MA, Estruch R, Moderate red wine consumption is associated with a lower prevalence of the metabolic syndrome in the PREDIMED population. *Br. J. Nutr* 2015, 113 Suppl, S121–30. [PubMed: 26148915]
- [2]. Koppes LL, Dekker JM, Hendriks HF, Bouter LM, Heine RJ, Statements ADA, Moderate alcohol consumption lowers the risk of type 2 diabetes: a meta-analysis of prospective observational studies. *Diabetes Care* 2005, 28, 719–725. [PubMed: 15735217]
- [3]. Ronksley PE, Brien SE, Turner BJ, Mukamal KJ, Ghali WA, Association of alcohol consumption with selected cardiovascular disease outcomes: a systematic review and meta-analysis. *BMJ* 2011, 342, d671–d671. [PubMed: 21343207]
- [4]. Mohamed Saleem TS, Darbar Basha S, Red wine: A drink to your heart. *J. Cardiovasc. Dis. Res* 2010, 1, 171–176. [PubMed: 21264180]
- [5]. Fragopoulou E, Choleva M, Antonopoulou S, Demopoulos CA, Wine and its metabolic effects. A comprehensive review of clinical trials. *Metabolism* 2018, 83, 102–119. [PubMed: 29408458]
- [6]. Vázquez-Fresno R, Llorach R, Alcaro F, Rodríguez MÁ, Vinaixa M, Chiva-Blanch G, Estruch R, Correig X, Andrés-Lacueva C, 1H-NMR-based metabolomic analysis of the effect of moderate wine consumption on subjects with cardiovascular risk factors. *Electrophoresis* 2012, 33, 2345–2354. [PubMed: 22887155]
- [7]. Urpi-Sarda M, Boto-Ordóñez M, Queipo-Ortuño MI, Tulipani S, Corella D, Estruch R, Tinahones FJ, Andres-Lacueva C, Phenolic and microbial-targeted metabolomics to discovering and

- evaluating wine intake biomarkers in human urine and plasma. *Electrophoresis* 2015, 36, 2259–2268. [PubMed: 25929678]
- [8]. Esteban-Fernandez A, Ibañez C, Simó C, Bartolome B, Moreno-Arribas MV, An UHPLC-TOF MS metabolomic approach to study the impact of moderate red wine consumption on urinary metabolome. *J. Proteome Res* 2018, acs.jpoteome.7b00904.
- [9]. Zamora-Ros R, Urpí-Sardà M, Lamuela-Raventós RM, Estruch R, Martínez-González MÁ, Bulló M, Arós F, Cherubini A, Andres-Lacueva C, Resveratrol metabolites in urine as a biomarker of wine intake in free-living subjects: The PREDIMED Study. *Free Radic. Biol. Med* 2009, 46, 1562–1566. [PubMed: 19167481]
- [10]. Modun D, Music I, Vukovic J, Brizic I, Katalinic V, Obad A, Palada I, Dujic Z, Boban M, The increase in human plasma antioxidant capacity after red wine consumption is due to both plasma urate and wine polyphenols. *Atherosclerosis* 2008.
- [11]. Wang DD, Toledo E, Hruby A, Rosner BA, Willett WC, Sun Q, Razquin C, Zheng Y, Ruiz-Canela M, Guasch-Ferré M, Corella D, Gómez-Gracia E, Fiol M, Estruch R, Ros E, Lapetra J, Fito M, Aros F, Serra-Majem L, Lee CH, Clish CB, Liang L, Salas-Salvadó J, Martínez-González MA, Hu FB, Plasma ceramides, mediterranean diet, and incident cardiovascular disease in the PREDIMED trial (prevención con dieta mediterránea). *Circulation* 2017, 135, 2028–2040. [PubMed: 28280233]
- [12]. Ruiz-Canela M, Guasch-Ferré M, Toledo E, Clish CB, Razquin C, Liang L, Wang DD, Corella D, Estruch R, Hernández Á, Yu E, Gómez-Gracia E, Zheng Y, Arós F, Romaguera D, Dennis C, Ros E, Lapetra J, Serra-Majem L, Papandreou C, Portoles O, Fitó M, Salas-Salvadó J, Hu FB, Martínez-González MA, Plasma branched chain/aromatic amino acids, enriched Mediterranean diet and risk of type 2 diabetes: case-cohort study within the PREDIMED Trial. *Diabetologia* 2018.
- [13]. Estruch R, Ros E, Salas-Salvadó J, Covas M-I, Corella D, Arós F, Gómez-Gracia E, Ruiz-Gutiérrez V, Fiol M, Lapetra J, Lamuela-Raventós RM, Serra-Majem L, Pintó X, Basora J, Muñoz MA, Sorlí JV, Martínez JA, Fitó M, Gea A, Hernán MA, Martínez-González MA, PREDIMED Study Investigators, Primary Prevention of Cardiovascular Disease with a Mediterranean Diet Supplemented with Extra-Virgin Olive Oil or Nuts. *N. Engl. J. Med* 2018, 378, e34. [PubMed: 29897866]
- [14]. Fernández-Ballart JD, Piñol JL, Zazpe I, Corella D, Carrasco P, Toledo E, Perez-Bauer M, Martínez-González MA, Salas-Salvadó J, Martín-Moreno JM, Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain. *Br. J. Nutr* 2010, 103, 1808–16. [PubMed: 20102675]
- [15]. Elosua R, Marrugat J, Molina L, Pons S, Pujol E, Validation of the Minnesota Leisure Time Physical Activity Questionnaire in Spanish men. The MARATHOM Investigators. *Am. J. Epidemiol* 1994, 139, 1197–1209. [PubMed: 8209878]
- [16]. De La Fuente-Arrillaga C, Vzquez Ruiz Z, Bes-Rastrollo M, Sampson L, Martinez-González MA, Reproducibility of an FFQ validated in Spain, in: *Public Health Nutrition*, 2010.
- [17]. Mascanfroni ID, Takenaka MC, Yeste A, Patel B, Wu Y, Kenison JE, Siddiqui S, Basso AS, Otterbein LE, Pardoll DM, Pan F, Priel A, Clish CB, Robson SC, Quintana FJ, Metabolic control of type 1 regulatory T cell differentiation by AHR and HIF1- $\alpha$ . *Nat. Med* 2015, 21, 638–646. [PubMed: 26005855]
- [18]. O'Sullivan JF, Morningstar JE, Yang Q, Zheng B, Gao Y, Jeanfavre S, Scott J, Fernandez C, Zheng H, O'Connor S, Cohen P, Vasan RS, Long MT, Wilson JG, Melander O, Wang TJ, Fox C, Peterson RT, Clish CB, Corey KE, Gerszten RE, Dimethylguanidino valeric acid is a marker of liver fat and predicts diabetes. *J. Clin. Invest* 2017, 127, 4394–4402. [PubMed: 29083323]
- [19]. Rowan S, Jiang S, Korem T, Szymanski J, Chang M-L, Szelog J, Cassalman C, Dasuri K, McGuire C, Nagai R, Du X-L, Brownlee M, Rabbani N, Thornalley PJ, Baleja JD, Deik AA, Pierce KA, Scott JM, Clish CB, Smith DE, Weinberger A, Avnit-Sagi T, Lotan-Pompan M, Segal E, Taylor A, Involvement of a gut–retina axis in protection against dietary glycemia-induced age-related macular degeneration. *Proc. Natl. Acad. Sci* 2017, 114, E4472–E4481. [PubMed: 28507131]
- [20]. Gromski P, Xu Y, Kotze H, Correa E, Ellis D, Armitage E, Turner M, Goodacre R, Influence of Missing Values Substitutes on Multivariate Analysis of Metabolomics Data. *Metabolites* 2014.

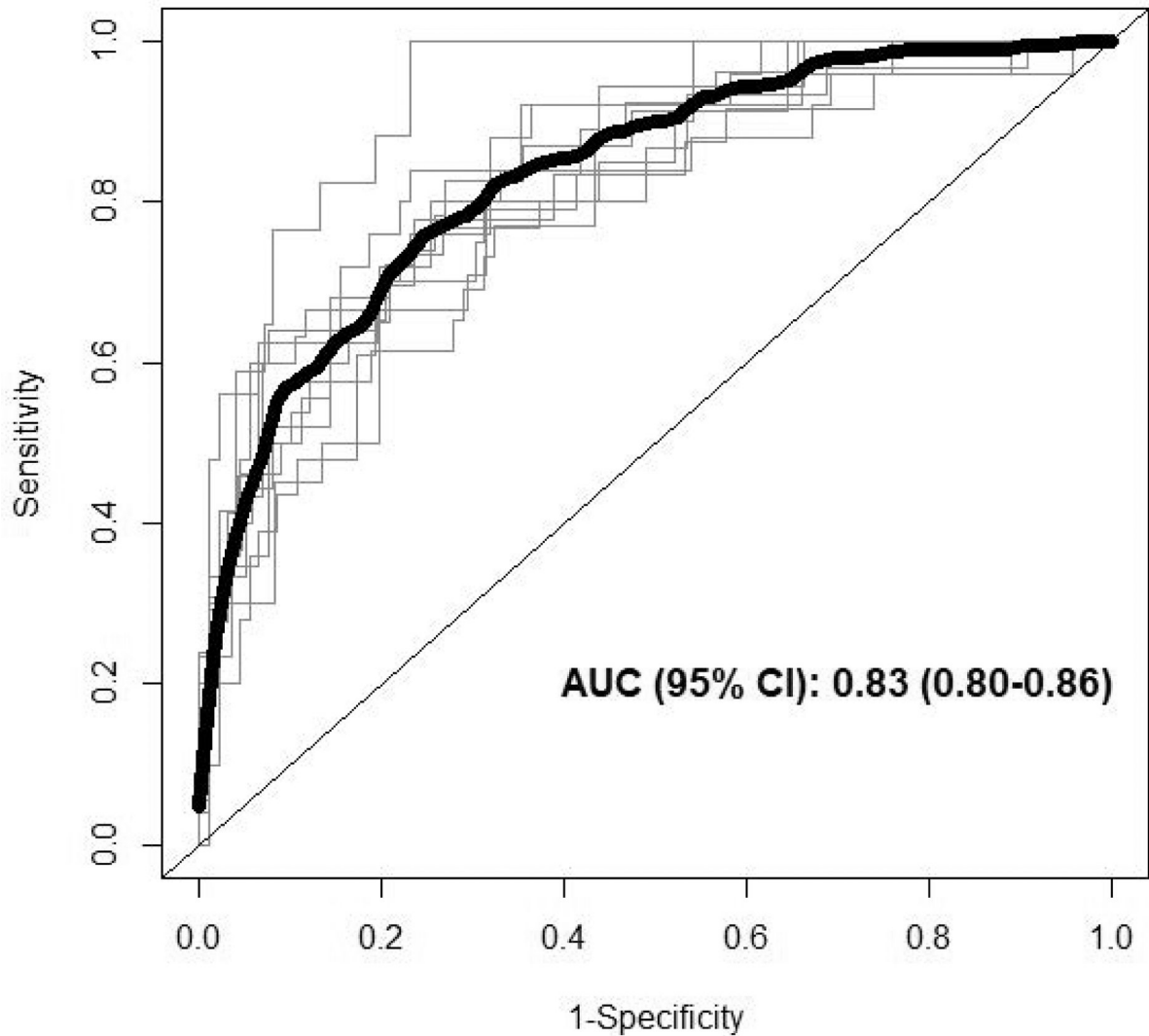
- [21]. Wei R, Wang J, Su M, Jia E, Chen S, Chen T, Ni Y, Missing Value Imputation Approach for Mass Spectrometry-based Metabolomics Data. *Sci. Rep* 2018.
- [22]. van den Berg RA, Hoefsloot HCJ, Westerhuis JA, Smilde AK, van der Werf MJ, Centering, scaling, and transformations: Improving the biological information content of metabolomics data. *BMC Genomics* 2006, 7.
- [23]. Lloyd AJ, Beckmann M, Haldar S, Seal C, Brandt K, Draper J, Data-driven strategy for the discovery of potential urinary biomarkers of habitual dietary exposure. *Am. J. Clin. Nutr* 2013, 97, 377–389. [PubMed: 23269817]
- [24]. Xia J, Broadhurst DI, Wilson M, Wishart DS, Translational biomarker discovery in clinical metabolomics: An introductory tutorial. *Metabolomics* 2013, 9, 280–299. [PubMed: 23543913]
- [25]. Tresserra-Rimbau A, Medina-Remón A, Pérez-Jiménez J, Martínez-González MA, Covas MI, Corella D, Salas-Salvador J, Gómez-Gracia E, Lapetra J, Arós F, Fiol M, Ros E, Serra-Majem L, Pintó X, Muñoz MA, Saez GT, Ruiz-Gutiérrez V, Warnberg J, Estruch R, Lamuela-Raventós RM, Dietary intake and major food sources of polyphenols in a Spanish population at high cardiovascular risk: The PREDIMED study. *Nutr. Metab. Cardiovasc. Dis* 2013.
- [26]. USDA, National Nutrient Database. US Dep. Agric. Agric. Res. Serv. Nutr. Data Lab. USDA Natl. Nutr. Database Stand. Ref. Release 28. Version Curr. Sept. 2015, Slightly Revis. May 2016 2015.
- [27]. Zamora-Ros R, Rothwell JA, Achaintre D, Ferrari P, Boutron-Ruault MC, Mancini FR, Affret A, Kühn T, Katzke V, Boeing H, Küppel S, Trichopoulou A, Lagiou P, La Vecchia C, Palli D, Contiero P, Panico S, Tumino R, Ricceri F, Noh H, Freisling H, Romieu I, Scalbert A, Evaluation of urinary resveratrol as a biomarker of dietary resveratrol intake in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Br. J. Nutr* 2017, 117, 1596–1602. [PubMed: 28637522]
- [28]. Klop B, Rego AT Do, Cabezas, M.C., Alcohol and plasma triglycerides. *Curr. Opin. Lipidol* 2013, 24, 321–326. [PubMed: 23511381]
- [29]. Hannuksela M, Marcel YL, Kesäniemi YA, Savolainen MJ, Reduction in the concentration and activity of plasma cholesteryl ester transfer protein by alcohol. *J Lipid Res* 1992, 33, 737–44. [PubMed: 1619365]
- [30]. Zheng Y, Yu B, Alexander D, Steffen LM, Nettleton JA, Boerwinkle E, Metabolomic patterns and alcohol consumption in African Americans in the Atherosclerosis Risk in Communities Study. *Am. J. Clin. Nutr* 2014, 99, 1470–1478. [PubMed: 24760976]
- [31]. Cunningham TJ, Yao L, Lucena A, Product inhibition of secreted phospholipase A2 may explain lysophosphatidylcholines' unexpected therapeutic properties. *J. Inflamm* 2008, 5.
- [32]. Crabb DW, Liangpunsakul S, Alcohol and lipid metabolism, in: *Journal of Gastroenterology and Hepatology (Australia)*, 2006.
- [33]. BADAWY AA-B, A Review of the Effects of Alcohol on Carbohydrate Metabolism. *Alcohol Alcohol* 1977, 12, 120–136.
- [34]. Markoski MM, Garavaglia J, Oliveira A, Olivares J, Marcadenti A, Molecular properties of red wine compounds and cardiometabolic benefits. *Nutr. Metab. Insights* 2016.
- [35]. Benowitz NL, Cotinine as a biomarker of environmental tobacco smoke exposure. *Epidemiol. Rev* 1996, 18, 188–204. [PubMed: 9021312]
- [36]. Torres-Collado L, García-De la Hera M, Navarrete-Muñoz EM, Compañ-Gabucio LM, Gonzalez-Palacios S, Vioque J, Coffee drinking and associated factors in an elderly population in Spain. *Int. J. Environ. Res. Public Health* 2018.
- [37]. *Inositol & its Phosphates: Basic Science to Practical Applications*, 2015.



**Figure 1.**

Coefficients (mean  $\pm$  SD) for the metabolites selected 10 times in the 10-CV logistic elastic regression. Mean and SD of the set of 13 metabolites selected the 10 times in the 10-CV elastic logistic regression procedure (using  $\lambda_{1se}$ ) employing the whole dataset of subjects. Metabolites with negative coefficients ( $m = 4$ ) are plotted in the left part, whereas those with positive coefficients ( $m = 9$ ) are shown in the right part. **Abbreviations:** CE, cholesteryl ester; CV, cross-validation; DMGV,  $\alpha$ -keto- $\delta$ -(N(G),N(G)-dimethylguanidino)valeric acid; PC, phosphatidylcholine; TAG, triglyceride.

## 10-fold Cross-Validated Area Under the Curve (10-CV-AUC)



**Figure 2.**

Ten-fold cross-validated ROC curves. Black bold line represents the 10-fold cross-validated ROC curve, whereas the thin grey lines show ROC curves for each of the 10 iterations using the training-validation (90%–10%) pair datasets excluding consistently selected covariates (i.e. sex, age and smoking) from the model. **Abbreviations:** CI, confidence interval; ROC, receiver operating characteristic.



**Table 1.**

Characteristics of the study subjects

	Non-RW consumers ( $\leq 1$ glass/day) (n = 924)	RW consumers ( $> 1$ glass/day) (n = 233)	Total subjects (N = 1,157)
<b>Characteristic</b>			
Median red wine consumption (mL/d)	0	250	0
Male sex, N (%)	245 (26.5)	191 (82.0) <sup>c</sup>	436 (37.7)
Age (years)	68.13 $\pm$ 5.97	65.95 $\pm$ 6.04 <sup>a</sup>	67.69 $\pm$ 6.04
Body mass index (kg/m <sup>2</sup> )	30.15 $\pm$ 3.73	29.29 $\pm$ 3.14 <sup>a</sup>	29.98 $\pm$ 3.63
Waist circumference (cm)	99.75 $\pm$ 10.51	102.34 $\pm$ 9.77 <sup>a</sup>	100.27 $\pm$ 10.41
Cholesterol (mg/dL)	213.6 $\pm$ 36.21	206.5 $\pm$ 36.50 <sup>a</sup>	212.15 $\pm$ 36.36
HDL-C (mg/dL)	52.1 $\pm$ 11.52	52.18 $\pm$ 10.80	52.11 $\pm$ 11.38
Triglycerides (mg/dL) <sup>‡</sup>	120.9 [92.3, 161.99]	111.24 [86.3, 153.69] <sup>b</sup>	118.89 [91.19, 159.95]
Smoking, N (%)			
Yes	100 (10.8)	82 (35.2) <sup>c</sup>	182 (15.7)
Type 2 Diabetes, N (%)	278 (30.1)	56 (24.0)	334 (28.9)
Hypercholesterolemia, N (%)	697 (75.4)	171 (73.4)	868 (75.0)
Hypertension, N (%)	818 (88.5)	204 (87.6)	1022 (88.3)
Family history of CVD, N (%)	247 (26.7)	46 (19.7)	293 (25.3)
Cardiac medication, N (%)	90 (10.0)	23 (9.9)	113 (10.1)
Hypotensive medication, N (%)	716 (77.7)	168 (72.1)	884 (76.6)
Cholesterol lowering medication, N (%)	432 (46.8)	114 (48.9)	546 (47.3)
Insulin medication, N (%)	49 (5.3)	5 (2.1)	54 (4.7)
Oral antidiabetics, N (%)	192 (20.8)	46 (19.7)	238 (20.6)
Cases of CVD incidence (%) <sup>‡‡</sup>	122 (13.2)	27 (11.6)	149 (12.9)
Cases of T2D incidence (%) <sup>§</sup>	128 (13.9)	44 (18.9)	172 (14.9)
Energy intake (Kcal/d) <sup>‡</sup>	2105.8 [1793.7, 2469.2]	2578.3 [2256.4, 2867.6] <sup>b</sup>	2192.4 [1853.7, 2593.9]
Carbohydrates (g/d) <sup>‡</sup>	226.0 [182.3, 274.1]	248.4 [202.3, 299.4] <sup>b</sup>	228.7 [184.9, 277.6]
Protein (g/d) <sup>‡</sup>	88.49 [75.85, 101.8]	94.9 [83.5, 110.7] <sup>b</sup>	89.4 [77.2, 104.4]
Fat (g/d) <sup>‡</sup>	91.9 [73.5, 110.5]	103.9 [88.8, 122.7] <sup>b</sup>	95.4 [75.7, 113.2]
SFA (g/d) <sup>‡</sup>	23.1 [18.3, 28.7]	26.5 [22.0, 32.7] <sup>b</sup>	23.9 [18.8, 29.5]
MUFA (g/d) <sup>‡</sup>	46.8 [34.1, 56.5]	52.6 [43.3, 61.0] <sup>b</sup>	47.9 [35.7, 57.5]
PUFA (g/d) <sup>‡</sup>	13.5 [10.4, 17.9]	16.4 [13.2, 21.2] <sup>b</sup>	14.0 [10.9, 18.6]
Total alcohol (g/d) <sup>‡</sup>	0 [0, 0.7]	30.9 [26.7, 47.8] <sup>b</sup>	0 [0, 10.2]
Total alcohol (mL/d) <sup>‡</sup>	0 [0, 21.6]	300 [255.2, 521.4] <sup>b</sup>	0 [0, 141.4]
Non-RW alcohol (mL/d) <sup>‡</sup>	0 [0, 21.6]	35.7 [6.7, 144.8] <sup>b</sup>	0 [0, 29.1]

	Non-RW consumers ( $\leq 1$ glass/day) (n = 924)	RW consumers ( $> 1$ glass/day) (n = 233)	Total subjects (N = 1,157)
Fruit intake (g/d) <sup>‡</sup>	340.7 [225.4, 481.6]	301.0 [190.7, 449.7] <sup>b</sup>	328.8 [217.6, 474.3]
Vegetable intake (g/d) <sup>‡</sup>	306.58 [225, 396.3]	312.17 [234.3, 412]	307 [226.8, 398.7]
Olive oil intake (mL/d) <sup>‡</sup>	35 [25, 50]	50 [25, 50]	35 [25, 50]
Nut intake (g/d) <sup>‡</sup>	8.57 [2, 17.1]	4.29 [0, 12.9] <sup>b</sup>	6.29 [2, 17.1]
Caffeinated coffee intake (mL/d) <sup>‡</sup>	0 [0, 50]	0 [0, 50]	0 [0, 50]
Decaffeinated coffee intake (mL/d) <sup>‡</sup>	21.43 [0, 50]	0 [0, 50]	7.14 [0, 50]
Red meat intake (g/d) <sup>‡</sup>	51.43 [31.4, 74.3]	74.29 [42.9, 85.7] <sup>b</sup>	52.9 [31.4, 84.3]

<sup>‡</sup>Data shows mean  $\pm$  SD or number (%).

<sup>a</sup>median [IQR];

<sup>b</sup>*P*-value  $< 0.05$  (Student's *t*-test between red wine categories);

<sup>c</sup>*P*-value  $< 0.05$  (Mann-Whitney *U*-test between red wine categories);

*P*-value  $< 0.05$  ( $\chi^2$  between red wine categories); CVD, cardiovascular disease; MUFA, monounsaturated fatty acids; PUFA, polyunsaturated fatty acids; RW, red wine; SFA, saturated fatty acids; T2D, type 2 diabetes. One glass is defined as 100 mL of red wine (10g of pure alcohol). As these baseline cross-sectional analyses include data from two nested case-cohort studies, we have added the following information:

<sup>‡</sup>incident CVD cases in the case-cohort study of CVD;

<sup>§</sup>incident T2D cases in the case-cohort study of T2D.