

Non-targeted LC-MS/MS metabolomic profiling of human plasma uncovers a novel Mediterranean diet biomarker panel

Shirin Macias

Queen's University Belfast

Ali Yilmaz

Beaumont Health

Joseph Kirma

Beaumont Health

Sarah E. Moore

Queen's University Belfast

Jayne V. Woodside

Queen's University Belfast

Stewart F. Graham

Beaumont Health

Brian D. Green (✉ b.green@qub.ac.uk)

Queen's University Belfast

Research Article

Keywords: Biomarkers, Mediterranean diet, Metabolomics, Dietary patterns, Mass spectrometry

Posted Date: June 7th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-3019157/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Metabolomics on December 8th, 2023. See the published version at <https://doi.org/10.1007/s11306-023-02058-y>.

Abstract

- Introduction: Consumption of a Mediterranean diet (MD) has established health benefits, and the identification of novel biomarkers could enable objective monitoring of dietary pattern adherence.
- Objectives: The present investigation performed untargeted metabolomics on blood plasma from a controlled study of MD adherence, to identify novel blood-based metabolite biomarkers associated with the MD pattern, and to build a logistic regression model that could be used to characterise MD adherence.
- Methods: A hundred and thirty-five plasma samples from $n = 58$ patients collected at different time points were available. Using a 14-point scale MD Score (MDS) subjects were divided into 'high' or 'low' MDS adherence groups and liquid chromatography-mass spectrometry (LC-MS/MS) was applied for analysis.
- Results: The strongest association with MDS was pectenotoxin 2 seco acid ($r = 0.53$; ROC = 0.78), a non-toxic marine xenobiotic metabolite. Several lipids were useful biomarkers including eicosapentaenoic acid, the structurally related lysophospholipid (20:5(5Z,8Z,11Z,14Z,17Z)/0:0), a phosphatidylcholine (P-18:1(9Z)/16:0) and also xi-8-hydroxyhexadecanedioic acid. Two metabolites negatively correlated with MDS, these were the monoacylglycerides (0:0/16:1(9Z)/0:0) and (0:0/20:3(5Z,8Z,11Z)/0:0). By stepwise elimination we selected a panel of 3 highly discriminatory metabolites and developed a linear regression model which identified 'high MDS' individuals with high sensitivity and specificity (AUC (95% CI) = 0.83 (0.76–0.97)).
- Conclusion: Our study highlights the utility of metabolomics as an approach for developing novel panels of dietary biomarkers. Quantitative profiling of these metabolites is required to validate their utility for evaluating dietary adherence.

1. Introduction

In recent years, there has been a surge of interest in the discovery of dietary biomarkers, particularly markers of healthy and sustainable dietary patterns. This area has seen an expansion in the application of metabolite-profiling (or metabolomics) technologies such as liquid chromatography-mass spectrometry (LC-MS) and nuclear magnetic resonance (NMR) spectroscopy. Studies have profiled biofluids such as urine, blood or saliva, and used this data to distinguish individual participant's adherence to patterns such as the Average Danish Diet (ADD) and the New Nordic Diet (NND) (1), the Baltic Sea diet (2), and, most commonly, the MD (3).

The MD is characterised by high consumption of fruit and vegetables, fish, legumes, nuts and olive oil; a moderate intake of wine, red meat and dairy, and low consumption of sweets and processed foods (4, 5). Measurement of adherence to the MD involves the use of various Mediterranean diet scores (MDS) which are based on food groups considered important within a MD. Two examples of these are the 9-item MedDiet score (6) and the *Prevención con Dieta Mediterránea* (PREDIMED) 14-item MedDiet score. A previous investigation in our group using $^1\text{H-NMR}$ metabolomic profiling was employed to identify a number of blood plasma biomarkers for the MD. This uncovered several modestly discriminatory metabolites including citric acid, pyruvic acid, betaine, mannose, and myo-inositol (7). Three similar studies have also performed $^1\text{H-NMR}$ profiling, although each of these studies examined urine rather than blood. One study found that MD adherence was associated with changes in the levels of 3-hydroxybutyrate, citric acid, and *cis*-aconitate, oleic acid, suberic acid, various amino acids and some microbial co-metabolites (8). Another study identified 34 metabolites associated with low or high adherence to MD (9). A third NMR study found higher urinary hippurate in individuals consuming a MD (supplemented with Coenzyme Q10) compared with individuals consuming a Western diet rich in saturated fat (10). There are some disadvantages with $^1\text{H-NMR}$ metabolomic profiling, firstly that relatively few metabolites are measured and, secondly, that it lacks the sensitivity of other analytical platforms such as LC-MS (11).

We are only aware of two studies that have applied untargeted LC-MS-based metabolomics in the search for blood biomarkers of a MD dietary pattern. The first compared four different diet quality indices in male Finish smokers and correlated them with serum metabolites: the Healthy Eating Index (HEI), the Alternate Mediterranean Diet Score (aMED), the WHO Healthy Diet Indicator (HDI), and the Baltic Sea Diet (BSD). A total of 46 metabolites were associated with a MD pattern measured by aMED, 21 of which were identifiable, including 4 amino acids, 1 carbohydrate, 2 vitamins/cofactors, 11 lipids and 3 xenobiotics. Moderate and low correlations were observed between these metabolites and aMED, with the highest correlation corresponding to the lipid 1-myristoleoylglycerophosphocholine (14:1) ($r = -0.3$) (2).

The second reported study involved 72 participants with high body mass index (BMI) and at least two features of Metabolic Syndrome (12). Participants followed either a Mediterranean Diet or a control diet in accordance with American Heart Association guidelines for a period of 6 months (2 months active intervention and 4 months self-administered). Results showed that levels of several blood lipids (mainly lipids and lysophospholipids) were altered by a MD-based intervention after 2 months, however these changes decreased from the second to the sixth month of the study, possibly due to lack of adherence to a MD during the self-adherence months (12).

The overall aim of the current study was to discover novel blood-based metabolite biomarkers associated with the MD pattern. We performed high LC-MS metabolomic analysis on blood plasma from a controlled study of MD adherence in a Northern European population. We then attempted to identify those ions with strong biomarker performance and to select an optimised biomarker panel. Lastly, using the identified metabolites, we built a logistic regression model that could best reflect MD adherence.

2. Materials and Methods

The Mediterranean Diet in Northern Ireland (MEDDINI) study was a pilot randomised controlled parallel group trial where 61 patients previously diagnosed with coronary heart disease (CHD) were recruited (135 plasma samples from $n = 58$ patients collected at different time points were available for the present study) from the Cardiology Directorate, Royal Victoria Hospital, Belfast. Patients provided informed written consent and were aged between 39 and 78 years. Seven-

day food diaries were used to collect food consumption data. Patients were asked to record the foods consumed over seven consecutive days, including an estimation of quantity consumed and information on preparation methods used. From seven-day food diaries, a database was created registering all food amounts eaten by all patients during the course of the intervention (baseline, 6 months, and 12 months). Foods portions were described in detail in the food diaries and all amounts were registered in grams/day in the database. Further details from the intervention study have been previously described and reported elsewhere (7, 13).

Food diaries from patients were scored using the validated 14-point MDS questionnaire based on the PREDIMED score. Scores are based on the answers to 14 questions along with food diaries. The questionnaire considered both the type of food and its frequency of intake (Supplementary information Table S1). A score of 0 indicated lowest adherence to an MD and a score of 14 indicated highest adherence. Recent MD advice was taken into consideration and types and quantities of foods within the PREDIMED score were adapted to reflect the typical diet and dietary recommendations in Northern Ireland. After scoring all patients food diets, they were divided into two groups split by the median (Low and High MDS). Details have been previously described in sections 2.2.2 and 2.2.4 (7).

2.1. Sample preparation

A total of 135 plasma samples from 58 participants from the MEDDINI study were analysed with the Dionex Ultimate 3000 UHPLC system coupled to an LTQ Orbitrap Elite mass spectrometer. This method was chosen for its high sensitivity. The extraction method was as follows: Plasma samples were stored at -80°C . Subsequently, all samples were thawed slowly on ice for 30 minutes prior to extraction, then 300 μL of ice cold methanol was added 100 μL of plasma, mixed for 10 minutes at 700 rpm, and subsequently centrifuged at $13,000\times g$ for 15 min under vacuum, and reconstituted in 100 μL of ultra-pure water. Samples were then filtered by centrifugation using a 0.22 μm Costar spin-X centrifuge tube filter ($8000\times g$ at 4°C for 5 min; Corning Incorporated, Corning, NY 14831, USA) and transferred to maximum recovery vials for analysis.

2.2. UPLC-MS analysis

All solvents were purchased from Fisher Scientific (Pittsburg, USA) and were LC-MS grade or equivalent. Chromatography was performed on a Dionex Ultimate 3000 UHPLC system (Dionex, Softron GmbH, Germany) coupled to an LTQ Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). 5 μL of extracted plasma was injected ($n = 3$ injections per sample) onto an Acquity UPLC CSH C18 column (2.1 x 100 mm, 1.7 μm , Waters, Wexford, Ireland) operating at 50°C and applying a binary mobile phase. The sample manager temperature was maintained at 4°C and the order in which the samples were injected was randomised throughout the experiment. The gradient elution buffers were A (water with 0.1% formic acid (vol/vol)) and B (methanol with 0.1% formic acid (vol/vol)). Solvent B was varied as follows: 0 min 1%, 2.5 min 1%, 16 min 99%, 18 min 99%, 18.1 min 1% and 20 min 1% with a flow rate of 0.4 mL.min⁻¹. Positive ionisation mode was employed with these conditions; source heater temperature at 400°C , sheath gas at 60 arbitrary units (AU), aux gas at 45 (AU) and sweep gas at 1 (AU), capillary temp was maintained at 325°C and source voltage at 3.5 kV. Mass spectra data were acquired in profile mode over the 50-1200 m/z range with a mass resolution of 60,000 at mass 400 (FWHM) and a scan time of 0.5s. In further experiments, the samples were subjected to mass fragmentation analysis (FT HCD (10, 30 and 70 NCE), MS2) with an isolation width of 1 Da and 60,000 FWHM at 400 m/z. Prior to sample analysis 10 pooled conditioning samples were injected. To determine chromatographic reproducibility of retention times and peak intensities, pooled samples were injected after every 10 sample injections throughout the experiment (14, 15). Pooled samples were comprised of all plasma samples from the study and were subjected to the same extraction procedure applied to the individual plasma samples.

2.3. Data analysis

UPLC-MS acquired data were analysed using Progenesis QI software (Waters Corporation, Milford, MA) for peak alignment, data normalisation and peak picking. Peak picking thresholds were set between 0.5 and 20 min. A peak threshold filter of 2.5 AU was applied (15). Data was normalised to all compounds by correcting for multiple features to determine a global scaling factor. From 3548 features found on Progenesis, filtering was applied by selecting those ions with $p\text{-value} < 0.01$, fold change > 1.5 and the coefficient of variation (CV) was $< 100\%$. Features with $> 20\%$ missing values were excluded. Filtered features were uploaded to Metaboanalyst where univariate (T-test) and multivariate (PCA, PLS-DA and O-PLS-DA) were performed and the model validated. Putative identifications were further examined to increase confidence in identification by mass fragmentation analysis whereby ms/ms spectra were used to search spectral libraries via additional online databases HMDB (16) and FooDB (17).

2.4. Metabolite identification

For each of the features related to the MDS obtained from Progenesis, identification was accomplished based on accurate mass (mass tolerance ≤ 4 ppm) and m/z values as matched against online databases. Adducts suggested by Progenesis were also matched with putative identifications from online libraries confirming that both neutral mass and m/z matched online databases. The confidence level of annotation was categorized according to the Metabolomics Standard Initiative (MSI) (18). Fragment masses observed from MS/MS experiments were also searched for using Xcalibur™ Software - Thermo Fisher Scientific. For a further level of confidence, ion intensities were correlated to previously measured targeted biomarkers using the software IBM® SPSS® Statistics. The identified metabolites were further used for performing a logistic regression analysis that could maximise the AUC (ROC) of a panel of biomarkers using the software Metaboanalyst 4.0.

2.5. Correlation with MDS and food groups

The association between putatively identified markers and MDS was analysed by correlating shortlisted ions of interest with the MDS. LC-MS data were non-normally distributed and hence metabolite correlations were examined using non-parametric correlations and Spearman's rank correlation coefficient using SPSS.

From the criteria used to measure adherence to MD, 14 food groups were selected, and their values were correlated to the shortlisted ions of interest. These food groups were: fruit, fruit juice, vegetables, combined fruit with fruit juice and vegetables, fish, nuts, legumes, red meat, processed meat, chicken and turkey, whole grain cereals, alcohol beverages and sweet foods. Non-parametric correlations (Spearman r) were selected, and correlations were carried out using SPSS. All p -values underwent Benjamini-Hochberg correction (q -values) and were deemed significant if $p \leq 0.05$ and $q \leq 0.05$.

Previous targeted biomarker analysis was carried out with MEDDINI serum and plasma samples using HPLC and GC-MS (19). These were Vitamin C, EPA and TG. To obtain further confirmation, metabolites putatively identified were also correlated to targeted data using SPSS.

3 Results

3.1. Univariate Analysis and identification of metabolites

A total of 135 plasma samples from 58 participants whose adherence to MD was assessed over time and scored, were divided into either 'low' MDS ($n = 63$) or 'high' MDS ($n = 72$). A total of 3548 features were detected using Progenesis. After data filtering and performing T-tests a total of 73 statistically significant features differed between low and high MDS groups. These features were ranked according to their: p -value, FDR and their variable importance in projection (VIP) scores from partial least squared discriminant analysis (PLS-DA) modelling. It was possible to assign metabolite identities to 7 features (ions of interest) from this list, which were: two monoacylglyceride: MG (0:0/16:1(9Z)/0:0 and MG (0:0/20:3(5Z,8Z,11Z)/0:0)), four fatty acid metabolites (EPA, lysoPC (20:5(5Z,8Z,11Z,14Z,17Z)/0:0), PC(P-18:1(9Z)/16:0), and xi-8-hydroxyhexa- decanedioic acid), and one xenobiotic (Pectenotoxin-2 secoacid (PTX2SA)). The characteristics and performance of each of these is outlined in detail in Table 1. The two monoacylglycerides were significantly higher in the 'low' MDS group. The other 5 metabolites were significantly higher in the 'high' MDS group.

3.2. Multivariate analysis

Figure 1 shows principal component analysis (PCA) of the 73 ions of the LC-MS data. Principal component 1 (PC1) explained 24.5% of the variance and component 2 (PC2) explained 17.4% of the variance. Two supervised multivariate methods were then applied; PLS-DA, and orthogonal partial least squares discriminant analysis (O-PLS-DA) both improved the separation between the groups. The PLS-DA model was subsequently cross-validated on Metaboanalyst using the 10-fold cross validation method (Supplementary Figure S1). The validation of the model showed R^2 of 0.66 and Q^2 of 0.48 after using a maximum of 4 components. Variable importance in projection (VIP) scores were used to rank the top 15 ions, and of these 5 were assigned identities.

3.3. Correlations with MDS and food groups

Correlations with MDS were also determined for each of 7 metabolites as well as food group-metabolite correlations (Table 2). The 2 monoacylglycerides (MG): MG (0:0/16:1(9Z)/0:0), MG (0:0/20:3(5Z,8Z11Z)/0:0) were significantly ($p < 0.05$) negatively correlated with MDS. The other 5 metabolites: EPA (ion), LysoPC (20:5(5Z,8Z,11Z,14Z,17Z)/0:0), PTX2SA, PC (P-18:1(9Z)/16:0), and xi-8-Hydroxyhexa decanedioic acid were significantly positively correlated with MDS (Table 2).

Table 1

Characteristics of seven putatively identified metabolites with biomarker potential for the Mediterranean dietary pattern. Table shows the univariate statistical (0–4) and High (5–10) MDS groups. Average intensities of ranked metabolites in both Low and High MDS groups and their respective standard deviations (SD) along with p -values, false discovery rate (FDR) corrected p -values, receiver operating characteristic (ROC) curve area under the curve (AUC), charge, adducts, neutral mass, putative identifications and proposed chemical formula.

	m/z	RT (min)	Low MDS (mean intensity)	Std. Dev	High MDS (mean intensity)	Std. Dev	p-value	FDR	ROC (AUC)	Charge	Adducts	Polarity	Putative identifi
1	541.3148	14.83	2876.45	1629.01	4922.09	3471.75	1.38E-8	4.76E-7	0.76	1	M + H	Positive	LysoPC(20:5(5Z
2	894.5209	17.40	990.39	654.44	2142.43	1436.66	1.81E-8	4.76E-7	0.78	1	M + NH ₄	Positive	Pectenotoxin 2 :
3	302.2235	15.74	8142.76	3775.01	13893.73	11077.42	5.31E-8	1.04E-6	0.77	1	M + Na	Positive	Eicosapentaenc
4	766.5722	17.86	6762.24	4066.45	16134.27	15842.27	1.86E-7	2.10E-6	0.79	1	M + Na	Positive	PC(P-18:1(9Z)/'
5	328.2602	15.57	27708.71	36777.19	6146.53	3961,89	2.37E-7	2.12E-6	0.74	1	M + Na	Positive	MG(0:0/16:1(9Z
6	403.2801	15.99	3326.77	3522.58	1296.85	563,90	5.11E-6	1.83E-5	0.71	1	M + Na	Positive	MG(0:0/20:3(5Z
7	302.2315	10.22	1870.19	1580.56	3383.72	2983,85	6.27E-5	1.83E-4	0.73	1	M + NH ₄ -H ₂ O	Positive	xi-8-Hydroxyhex

Table 2

Correlation of metabolites with MDA and food intake. Table shows the Spearman rank correlations of shortlisted metabolites with food groups used in calculations ($p < 0.05$) are marked in bold. Benjamini-Hochberg multiple comparison correction was significant if ($q < 0.05$). Number of correlations

Consumption of food grouping (g/day)		MG (0:0/20:3(5Z,8Z,11Z)/0:0)	EPA	LysoPC (20:5(5Z,8Z,11Z,14Z,17Z)/0:0)	PC(P-18:1(9Z)/16:0)	MG (0:0/16:1(9Z)/0:0)	Pectenotoxin 2 secoacid	xi-8-Hydr acid
MDS	r	-0.405	0.352	0.432	0.511	-0.464	0.534	0.44
	p	1.0E-6	6.9E-5	1.6E-7	1.04E-9	1.40E-8	2.6E-11	6.83E-11
	q	1.0E-5	3.0E-4	1.9E-6	3.9E-8	2.8E-7	3.1E-9	1.0E-10
Fruit & Fruit Juice	r	-0.330	0.197	0.238	0.362	-0.381	0.360	0.38
	p	9.3E-5	0.023	0.005	1.6E-5	5.0E-6	1.8E-5	5.0E-6
	q	3.8E-4	0.042	0.010	1.0E-4	3.7E-5	1.0E-4	3.7E-5
Fruit	r	-0.280	0.188	0.181	0.310	-0.354	0.328	0.35
	p	0.001	0.030	0.035	2.5E-4	2.6E-5	1.1E-4	2.4E-5
	q	0.002	0.052	0.058	8.1E-4	1.3E-4	4.2E-4	1.2E-4
Vegetables	r	-0.147	0.183	0.197	0.133	-0.080	0.186	0.19
	p	0.092	0.036	0.024	0.129	0.356	0.033	0.02
	q	0.134	0.059	0.043	0.173	0.423	0.055	0.04
Fruit, Fruit Juice & Vegetables	r	-0.319	0.230	0.293	0.350	-0.356	0.378	0.40
	p	1.6E-4	0.008	0.001	3.2E-5	2.2E-5	6.0E-6	2.0E-6
	q	6.0E-4	0.016	0.025	1.5E-4	1.2E-4	4.2E-5	1.84E-5
Legumes	r	-0.074	-0.051	-0.046	0.070	-0.084	-0.025	-0.11
	p	0.399	0.558	0.599	0.428	0.336	0.776	0.19
	q	0.464	0.625	0.659	0.493	0.403	0.809	0.25
Fish	r	-0.204	0.426	0.524	0.492	-0.315	0.459	0.31
	p	0.019	2.5E-7	7.2E-11	1.3E-9	1.9E-4	3.1E-8	2.0E-8
	q	0.036	2.72E-6	4.32E-9	3.9E-8	6.7E-4	5.31E-7	6.8E-8
Red Meat	r	0.078	-0.177	-0.133	-0.206	0.151	-0.258	-0.23
	p	0.373	0.042	0.129	0.018	0.084	0.003	0.00
	q	0.438	0.067	0.173	0.034	0.124	6.6E-3	0.01
Processed Meat	r	0.437	-0.307	-0.316	-0.383	0.294	-0.364	-0.34
	p	1.1E-7	2.9E-4	1.8E-4	5.0E-6	5.3E-4	1.8E-5	4.1E-5
	q	1.4E-6	8.9E-4	6.5E-4	3.7E-5	1.5E-3	1.0E-4	1.8E-4
White meat	r	-0.165	0.014	0.015	-0.021	-0.094	0.137	0.16
	p	0.059	0.875	0.867	0.814	0.283	0.118	0.06
	q	0.093	0.882	0.881	0.842	0.343	0.166	0.09
Cereals	r	-0.312	0.140	0.228	0.271	-0.289	0.290	0.47
	p	2.3E-4	0.109	0.009	.002	6.7E-4	0.001	8.5E-5
	q	7.6E-4	0.155	0.018	0.005	0.002	0.003	2.0E-5
Sweets and carbonated drinks	r	0.198	-0.041	-0.097	-0.062	0.113	-0.192	-0.15
	p	0.023	0.642	0.267	0.477	0.197	0.027	0.08
	q	0.042	0.688	0.330	0.545	0.252	0.047	0.12
Sweets	r	0.161	0.010	-0.043	-0.051	0.112	-0.117	-0.09
	p	0.066	0.908	0.628	0.565	0.203	0.183	0.27

Consumption of food grouping (g/day)		MG (0:0/20:3(5Z,8Z,11Z)/0:0)	EPA	LysoPC (20:5(5Z,8Z,11Z,14Z,17Z)/0:0)	PC(P-18:1(9Z)/16:0)	MG (0:0/16:1(9Z)/0:0)	Pectenotoxin 2 secoacid	xi-8-Hydr acid
	q	0.100	0.908	0.685	0.627	0.252	0.238	0.330
Alcohol	r	-0.127	0.257	0.254	0.133	-0.041	0.186	-0.03
	p	0.145	0.003	0.003	0.129	0.643	0.033	0.688
	q	0.193	0.006	0.006	0.173	0.688	0.055	0.730
Nuts	r	-0.126	0.281	0.264	0.265	-0.133	0.307	0.160
	p	0.151	9.4E-4	0.002	0.002	0.128	3.3E-4	0.060
	q	0.199	0.002	0.005	0.005	0.173	9.6E-4	0.090

3.4 Logistic regression model using identified metabolites.

We then performed a range of logistic regression analyses based on combinations of the 7 putatively identified metabolites until optimised to obtain the highest possible ROC (AUC) value (Fig. 2). Using the intensities of MG(0:0/16:1(9Z)/0:0), PTX2SA, and PC(P-18:1(9Z)/16:0) and following cross validation (1000 permutations), we observed significant separation ($p < 0.001$) between Low and High MDS. We developed a logistic regression algorithm with an AUC (95% CI) = 0.830 (0.763–0.894) with corresponding sensitivity and specificity equal to 0.794 (0.794–0.894) and 0.722 (0.619–0.826) respectively, following 10-fold cross validation. Supplementary Table S2 lists the summary of each feature used to develop the following predictive algorithm:

$$\text{logit}(P) = \log \left(\frac{P}{1 - P} \right) = 0.313 - \text{MG}(0:0/16:1(9Z)/0:0) - 0.001 \text{ Pectenotoxin-2 seco acid} - \text{PC}(P-18:1(9Z)/16:0)$$

Where P is $\text{Pr}(y = 1|x)$. The best threshold (or Cutoff) for the predicted P is 0.44.

4 Discussion

The present study acquired food diary data and concomitant blood samples from a Mediterranean diet intervention study in a Northern European population. Non-targeted metabolomic profiling was performed on plasma and this was combined with the calculated MDS to shortlist 73 features significantly differing between 'low' and 'high' consumers of a Mediterranean diet. Ultimately, this led to the putative identification of 7 high performing metabolite biomarkers (ROCAUC ≤ 0.79), which were also highly influential in multivariate modelling, and strongly correlated with MDS. Correlation of these biomarkers against each of the food groups involved in the calculation of MDS provided potential information on the possible origin of some of these. Using logistic regression analysis, we developed a model which accurately distinguished between the two dietary groups. We developed a predictive algorithm using the acquired data which had an AUC (95% CI) = 0.83 (0.76–0.89) with corresponding sensitivity and specificity equal to 0.79 (0.79–0.89) and 0.72 (0.62–0.83). To our knowledge, no logistic regression model currently exists for human plasma that distinguishes low and high MDS with such a high degree of accuracy.

The results appear impressive, given that adherence in the MEDDINI intervention study (from which samples and data were obtained) would be described as sub-maximal for MD (Maximum MDS achieved was 10 from the 14-point scale. Mean MDS in the High MDS group was 6.68). Participants were not supplemented with any foods, and as such, did not eat identical foods. We should be mindful of the potential complexities here, particularly that the participants had a history of CVD and were taking prescribed medications. However, a number of the metabolites uncovered as biomarkers have obvious connections with food consumption (especially those with ROCAUC values > 0.7). Top performing putative biomarkers included PC (P-18:1(9Z)/16:0), EPA (an n-3 long chain polyunsaturated fatty acid), and a closely related metabolite LysoPC(20:5(5Z,8Z,11Z,14Z,17Z)/0:0). The strong correlation of each of these, not only with MDS, but also with fish consumption underscores dietary relevance. MEDDINI participants did not take any fish oil supplements, which rules this out as a possible source. EPA has previously been reported as a validated biomarker of fish intake (20), but the lysophospholipid metabolite performed substantially better than the free fatty acid (spearman $r = 0.524$ vs 0.426). This novel biomarker ought to be closely examined in future dietary biomarker studies, but it is encouraging that one other study is supportive of it as a biomarker of MD (12). Unfortunately, due to time, budget and instrument availability it was not possible to check metabolite identifications against analytical standards. However, the intensity values for EPA and LysoPC(20:5(5Z,8Z,11Z,14Z,17Z)/0:0) were correlated against quantitated values acquired from a previously published investigation (13) which found strong associations ($r = 0.595$; $p = 6.9E-14$ and $r = 0.589$; $p = 1.3E-13$, respectively) thus providing reasonable confidence in their identity. Fish oil supplementation studies in human volunteers indicate that LysoPC (20:5(5Z,8Z,11Z,14Z,17Z)/0:0) is associated with EPA intake (21). It is thought that LysoPC(20:5(5Z,8Z,11Z,14Z,17Z)/0:0) is more bioavailable than the free fatty acid, as gavage studies in mice increased the levels of EPA in the brain by > 100 -fold(22).

Despite the potential utility of eicosapentaenoic acid metabolites, ultimately, none of these were incorporated into the optimised logistic regression model predicting MDS group. The three metabolites incorporated were MG(0:0/16:1(9Z)/0:0), PTX2SA, and PC(P-18:1(9Z)/16:0). The combined correlation of the 3 metabolites included in the model and MDS was ($r = 0.64$; $p = 5.6E-17$). MG(0:0/16:1(9Z)/0:0) was one of two monoacylglyceride biomarkers identified (the other being MG(0:0/20:3(5Z,8Z,11Z)/0:0)). The respective fatty acid components of these MGs have potentially strong dietary relevance, and both markers significantly decreased with increasing MDS. Both of these metabolites were correlated with previous targeted measured blood triglycerides showing a highly significant association ($r = 0.511$; $p = 1.2E-9$) and ($r = 0.348$; $p = 6.8E-5$), respectively.

MG(0:0/20:3(5Z,8Z,11Z)/0:0) is the monoacylglyceride metabolite of mead acid (20:3(5Z,8Z,11Z)). The levels of mead acid in plasma are known to be a marker for overall essential fatty acid (EFA) status (23, 24). Mead acid is not itself considered an essential fatty acid, but in the absence of adequate essential fatty acids in human tissues, the fatty acid MG(0:0/20:3(5Z,8Z,11Z)/0:0) is metabolised to mead acid (24–27). Given that MG(0:0/20:3(5Z,8Z,11Z)/0:0) is significantly negatively correlated with MDS potentially indicates that some individuals in the 'low' MDS group, particularly those with very high MG(0:0/20:3(5Z,8Z,11Z)/0:0) levels, may exhibit deficiency in essential fatty acids (28, 29). Adherence to MD has proved an enhancement in essential fatty acid levels (30, 31). This is consistent with our findings which showed a significant difference in EFA concentrations between low and high MD adherence groups ($p = 0.001$) and an 11.10% increase between low and high MD adherence.

Alternatively, it is also possible that MG (0:0/20:3(5Z,8Z,11Z)/0:0) is derived from the diet. For instance, mead acid is present in very high levels in animal cartilage, and it is noteworthy that this metabolite is strongly correlated with processed meat intake. The other MG biomarker identified is a metabolite of trans-palmitoleic acid (16:1(9Z)) which has previously been established as a marker of full fat dairy intake (32). Unfortunately, the correlation between this ion and dairy intake was not statistically significant in our study ($p = 0.12$).

Another novel high performing MDS biomarker which correlated strongly with fish consumption is a xenobiotic compound called Pectenotoxin-2 seco acid (PTX2SA). PTX2SA correlated closely with both MDS and fish intake. Produced by toxic dinoflagellates, pectenotoxins accumulate in shellfish, and humans can potentially be exposed to these through shellfish consumption. PTX2 is one of the family of pectenotoxin compounds, which are polyether macrolide toxins responsible for diarrhetic shellfish poisoning (33). Reassuringly however, PTX2sa is a non-toxic metabolite of PTX2. Injection of mice with doses as high as 5mg/kg does not cause toxicity, and oral administration of PTX2sa is likely to be even less toxic (34). Intriguingly, PTX2sa and its epimer 7-epi-PTX2sa have previously been detected in Irish waters (35) however, we cannot find any evidence that PTX2sa has been detected in humans before. PTX2 is highly lipophilic and may not be released and absorbed during human digestion of shellfish. It also appears to be quite labile and, if it were to be liberated during digestion, acidity in the stomach would rapidly lead to metabolism to its non-toxic seco acid. PTX2sa has been detected in various marine samples, with mollusks and plankton being the most abundant sources (36). We are not aware that PTX2 or PTX2sa has been detected in fish specimens before, however, it seems likely given that trace amounts will occur, given that PTX2 has been measured at $\leq 8\text{ng/l}$ in seawater, $\leq 10\text{ ng}$ in suspended particulate matter and $\leq 2\text{ng/g}$ in marine sediment (37). It is entirely plausible that this compound could originate from fish intake. Mollusk/shellfish consumption among MEDDINI participants during the surveyed period was, however, extremely rare, and the calculation of MDS was based on fish intake, and not shellfish intake. Alternatively, given the fact that PTX2/PTX2sa are highly lipophilic they may persist long after absorption and may reflect shellfish intake outside of the food diary data collection period.

The third metabolite to be incorporated in the model was a phosphatidylcholine identified PC(16:0/18:1(11Z)), comprised of palmitic acid and vaccenic acid. This metabolite was the best performing individual biomarker overall (AUCROC = 0.79) and it strongly positively correlated with MDS ($r = 0.495$; $p = 1.04\text{E-}9$). It is difficult to pinpoint the dietary origin of PC(16:0/18:1(11Z)) as it correlated with a number of food types, but the strongest association was with fish intake ($r = 0.492$, $p\text{-value} = 1.3\text{E-}9$), which was almost equal in strength to its association with MDS.

One other noteworthy MDS biomarker detected was Xi-8-Hydroxyhexadecanedioic acid (Xi-8-HHDDA), also known as 8-hydroxyhexadecane dioic acid. Xi-8-HHDDA is a long-chain fatty acid, and here it was significantly associated with fruit, fruit juice and vegetable intake, and it also correlated with blood plasma levels of vitamin C ($r = 0.30$; $p = 8\text{E-}5$). It seems likely to originate from dietary plant intake, given that it is a cutin constituent of fruits and vegetables (38, 39). Its presence has been reported in fruit and tomatoes, which makes it a potential biomarker for the consumption of these food products (40, 41). It has also been identified as one of the major constituents of sweet cherries (42). Other dioic acids have been reported in other fruits, for example, 10,16-dihydroxyhexadecanoic was identified as major components of the of the cuticle of different apple varieties (43). To the best of our knowledge, this is the first time that the metabolite xi-8-Hydroxyhexadecanedioic has been identified as a potential biomarker of fruits and vegetable intake.

In conclusion, the present study is only the third to apply untargeted LC-MS metabolomics to a MD study and it provides a clear indication that this approach can be effective in a Northern European population with sub-maximal MD adherence. The findings further advance the ongoing search for a biomarker panel to determine adherence to a MD diet. Specifically, we propose a logistic regression model for accurately distinguishing low or high MDS, which will require careful validation using targeted and quantitative methods in other MD cohorts. There is clear evidence that the shortlisted metabolite biomarkers have statistically significant dietary associations (five for fish intake, one for fruit and vegetable intake, and two for processed meat), thus making the findings biologically plausible, and worthy of further investigation.

Declarations

Ethics approval.

This study was conducted in accordance with the guidelines laid down in the 1964 Declaration of Helsinki and its later amendments. All procedures involving human subjects were approved by the Queen's University Belfast Research Ethics Committee; Ethical approval references: RGHT000049 and 15.42 for the MEDDINI original study (2006) and for the present latest analysis respectively in view of the retrospective nature of the study.

Consent to participate and to publish.

Written informed consent was obtained from all MEDDINI participants.

Funding

This research was funded by the Research and Development Office, Northern Ireland. S.M. was funded by a PhD Scholarship from the Department for the Economy (DfE), Northern Ireland.

Competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

Data availability

Data uploaded to Metabolights: MTBLS7948: Non-targeted LC-MS/MS metabolomic profiling of human plasma uncovers a novel Mediterranean diet biomarker panel. <https://www.ebi.ac.uk/metabolights/editor/console>

Software availability

The software used for data analysis in this study (Progenesis) is available via: <https://www.nonlinear.com/progenesis/qi/>

And the online software Metaboanalyst: <https://www.metaboanalyst.ca/>

Author Contributions: S.M. collated data, performed statistical analysis, and wrote the manuscript. S.M. and J.K. carried out sample extraction. J.V.W. and B.D.G. supervised S.M. and wrote and edited the manuscript. S.F.G. designed and oversaw the LC-MS analysis and A.Y. assisted with the LC-MS-MS analysis. S.E.M. assisted with the MD scoring of participants. All authors provided critical feedback and helped shape the manuscript. All Authors read and approved the final manuscript.

Conflict of interest:

Shirin Macias declares she has no conflict of interest.

Ali Yilmaz declares he has no conflict of interest.

Joseph Kirma declares he has no conflict of interest.

Sarah E. Moore declares she has no conflict of interest.

Jayne V. Woodside declares she has no conflict of interest.

Stewart F. Graham declares he has no conflict of interest.

Brian D. Green declares he has no conflict of interest.

Ethical approval: Queen's University Belfast Research Ethics Committee (Ref: RGHT000049 and 15.42 for the MEDDINI original study and for the latest analysis respectively).

Acknowledgements: We thank the participants of the MEDDINI study and Karen Logan, who carried out the recruitment, delivered the interventions and completed all outcome assessment. A special thanks to Charlotte Neville for her assistance with governance procedures for the current analysis.

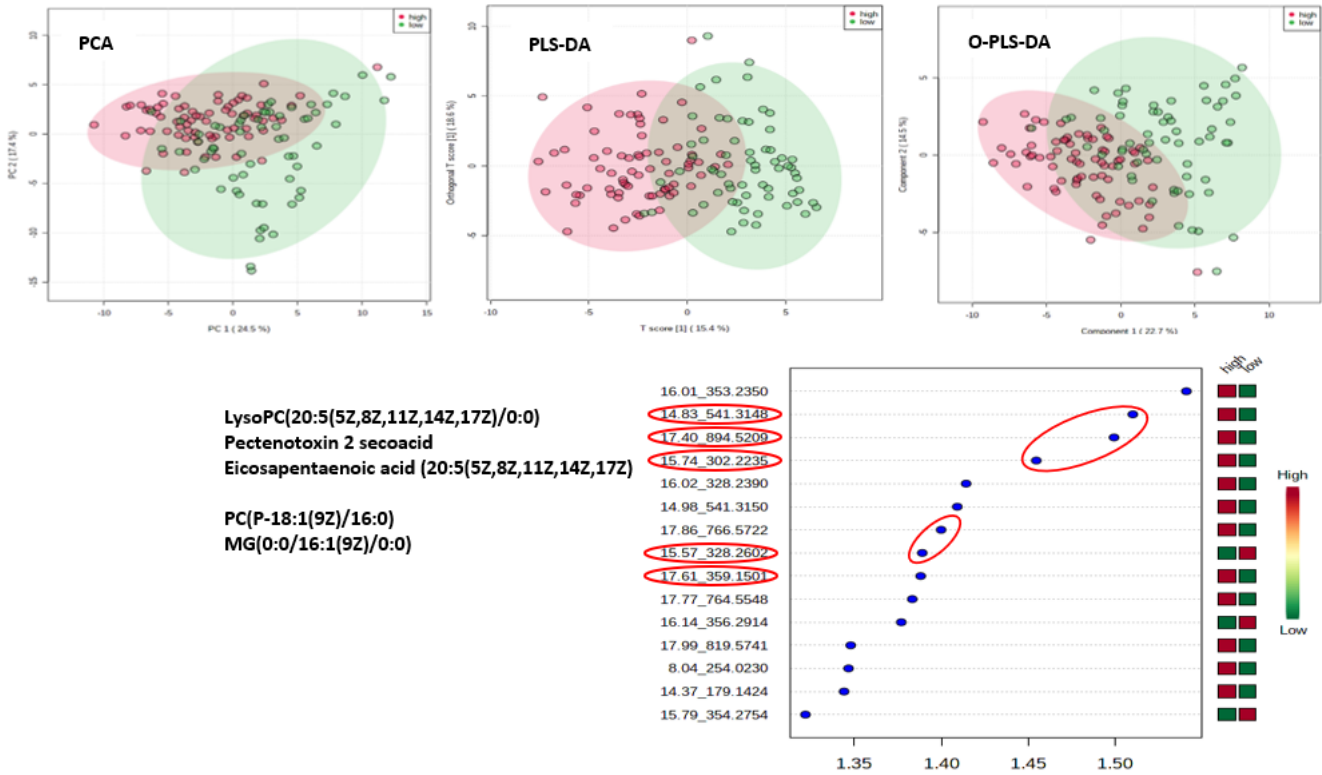
References

1. Khakimov, B., Kellebjerg Poulsen, S., Savorani, F., Acar, E., Gu zde, Larsen, T. M. (2016). New Nordic Diet versus Average Danish Diet: A Randomized Controlled Trial Revealed Healthy Long-Term Effects of the New Nordic Diet by GC – MS Blood Plasma Metabolomics. [cited 2021 Feb 17]; Available from: <https://pubs.acs.org/sharingguidelines>
2. Playdon, M. C., Moore, S. C., Derkach, A., Reedy, J., Subar, A. F., Sampson, J. N.. Identifying biomarkers of dietary patterns by using metabolomics. *Am J Clin Nutr* [Internet]. 2017 Feb 1 [cited 2021 Feb 17];105(2):450–65. Available from: [/pmc/articles/PMC5267308/](https://pubmed.ncbi.nlm.nih.gov/273087308/)
3. Fitó, M., Melander, O., Martínez, J., Toledo, E., Carpéné, C., & Corella, D.. Advances in Integrating Traditional and Omic Biomarkers When Analyzing the Effects of the Mediterranean Diet Intervention in Cardiovascular Prevention. *Int J Mol Sci* [Internet]. 2016 Sep 2 [cited 2021 Jan 19];17(9):1469. Available from: <http://www.mdpi.com/1422-0067/17/9/1469>
4. Martínez-González, M. A., García-Arellano, A., Toledo, E., Salas-Salvadó, J., Buil-Cosiales, P., Corella, D.. A 14-item mediterranean diet assessment tool and obesity indexes among high-risk subjects: The PREDIMED trial. *PLoS One* [Internet]. 2012 Aug 14 [cited 2021 Feb 17];7(8). Available from: [/pmc/articles/PMC3419206/](https://pubmed.ncbi.nlm.nih.gov/23419206/)
5. Estruch, R., Ros, E., Salas-Salvadó, J., Covas, M. I., Corella, D., Arós, F. (2018). Primary Prevention of Cardiovascular Disease with a Mediterranean Diet Supplemented with Extra-Virgin Olive Oil or Nuts. *N Engl J Med* [Internet]. Jun 21 [cited 2021 Feb 17];378(25):e34. Available from: <https://pubmed.ncbi.nlm.nih.gov/29897866/>
6. Martínez-González, M. A., Fernández-Jarne, E., Serrano-Martínez, M., Wright, M., & Gomez-Gracia, E.. Development of a short dietary intake questionnaire for the quantitative estimation of adherence to a cardioprotective Mediterranean diet. *Eur J Clin Nutr* [Internet]. 2004 Nov 26 [cited 2020 Sep 14];58(11):1550–2. Available from: www.nature.com/ejcn
7. Macias, S., Kirma, J., Yilmaz, A., Moore, S. E., McKinley, M. C., McKeown, P. P. (2019). Application of 1H-NMR Metabolomics for the Discovery of Blood Plasma Biomarkers of a Mediterranean Diet. *Metabolites* [Internet]. Oct 1 [cited 2020 Aug 27];9(10). Available from: <https://pubmed.ncbi.nlm.nih.gov/31569638/>

8. Vázquez-Fresno, R., Llorach, R., Urpi-Sarda, M., Lupianez-Barbero, A., Estruch, R., Corella, D. (2015). Metabolomic pattern analysis after mediterranean diet intervention in a nondiabetic population: A 1- and 3-year follow-up in the PREDIMED study. *J Proteome Res* [Internet]. Jan 2 [cited 2021 Jan 18];14(1):531–40. Available from: <https://pubs.acs.org/sharingguidelines>
9. Almanza-Aguilera, E., Urpi-Sarda, M., Llorach, R., Vázquez-Fresno, R., Garcia-Aloy, M., Carmona, F., et al. (2017 Oct). Microbial metabolites are associated with a high adherence to a Mediterranean dietary pattern using a 1H-NMR-based untargeted metabolomics approach. *Journal Of Nutritional Biochemistry*, 7, 48:36–43.
10. González-Guardia, L., Yubero-Serrano, E. M., Delgado-Lista, J., Perez-Martinez, P., Garcia-Rios, A., Marin, C.. Effects of the mediterranean diet supplemented with coenzyme Q10 on metabolomic profiles in elderly men and women. *Journals Gerontol - Ser A Biol Sci Med Sci* [Internet]. 2015 Jan 1 [cited 2021 Feb 18];70(1):78–84. Available from: <https://pubmed.ncbi.nlm.nih.gov/24986061/>
11. Brennan, L. (2014). *NMR-based metabolomics: From sample preparation to applications in nutrition research* (83 vol., pp. 42–49). Elsevier B.V.. Progress in Nuclear Magnetic Resonance Spectroscopy.
12. Bondía-Pons, I., Martínez, J. A., de la Iglesia, R., Lopez-Legarrea, P., Poutanen, K., Hanhineva, K.. Effects of short- and long-term Mediterranean-based dietary treatment on plasma LC-QTOF/MS metabolic profiling of subjects with metabolic syndrome features: The Metabolic Syndrome Reduction in Navarra (RESMENA) randomized controlled trial. *Mol Nutr Food Res* [Internet]. 2015 Apr 1 [cited 2021 Jan 19];59(4):711–28. Available from: <http://doi.wiley.com/10.1002/mnfr.201400309>
13. Logan, K. J., Woodside, J. V., Young, I. S., McKinley, M. C., Perkins-Porras, L., & McKeown, P. P. Adoption and maintenance of a Mediterranean diet in patients with coronary heart disease from a Northern European population: a pilot randomised trial of different methods of delivering Mediterranean diet advice. *J Hum Nutr Diet* [Internet]. 2010 Feb 1 [cited 2021 Jan 21];23(1):30–7. Available from: <http://doi.wiley.com/10.1111/j.1365-277X.2009.00989.x>
14. Graham, S. F., Chevallier, O. P., Roberts, D., Hölscher, C., Elliott, C. T., & Green, B. D. (2013). Investigation of the human brain metabolome to identify potential markers for early diagnosis and therapeutic targets of Alzheimer's disease. *Anal Chem* [Internet]. Feb 5 [cited 2021 Feb 22];85(3):1803–11. Available from: <https://pubs.acs.org/sharingguidelines>
15. Cowan, E., Kumar, P., Burch, K. J., Grieve, D. J., Green, B. D., & Graham, S. F.. Treatment of lean and diet-induced obesity (DIO) mice with a novel stable obestatin analogue alters plasma metabolite levels as detected by untargeted LC–MS metabolomics. *Metabolomics* [Internet]. 2016 Jul 1 [cited 2020 Sep 14];12(7). Available from: </pmc/articles/PMC4932145/?report=abstract>
16. Wishart, D. S., Knox, C., Guo, A. C., Eisner, R., Young, N., Gautam, B.. HMDB: A knowledgebase for the human metabolome. *Nucleic Acids Res* [Internet]. 2009 [cited 2020 Sep 14];37(SUPPL. 1). Available from: <https://pubmed.ncbi.nlm.nih.gov/18953024/>
17. FooDB [Internet]. [cited 2021 Aug 27]. Available from: <https://foodb.ca/>
18. Sumner, L. W., Alexander, A. E., Ae, A., Ae, D. B., Ae, M. H. B., Beger, R.. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* [Internet]. 2007 [cited 2021 Feb 22];3:211–21. Available from: <http://msi-workgroups.sourceforge.net/http://www.metabolomicsociety.org/Reference:http://msi-workgroups.sourceforge.net/bio-metadata/reporting/pbc/http://msi-workgroups.sourceforge.net/chemical-analysis/>
19. Logan, K. J., Woodside, J. V., Young, I. S., McKinley, M. C., Perkins-Porras, L., & McKeown, P. P. Adoption and maintenance of a Mediterranean diet in patients with coronary heart disease from a Northern European population: A pilot randomised trial of different methods of delivering Mediterranean diet advice. *J Hum Nutr Diet* [Internet]. 2010 Feb [cited 2020 Aug 27];23(1):30–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/19788708/>
20. Cuparencu, C., Praticó, G., Hemeryck, L. Y., Sri Harsha, P. S. C., Noerman, S., Rombouts, C. (2019). Biomarkers of meat and seafood intake: An extensive literature review [Internet]. Vol. 14, Genes and Nutrition. BioMed Central Ltd.; [cited 2021 Feb 18]. p. 1–30. Available from: <https://doi.org/10.1186/s12263-019-0656-4>
21. Block, R. C., Duff, R., Lawrence, P., Kakinami, L., Brenna, J. T., Shearer, G. C.. The Effects of EPA, DHA, and Aspirin Ingestion on Plasma Lysophospholipids and Autotaxin. *Prostaglandins Leukot Essent Fat Acids* [Internet]. 2010 Feb [cited 2021 Feb 22];82(2–3):87–95. Available from: </pmc/articles/PMC2926929/>
22. Yalagala, P. C. R., Sugasini, D., Dasarathi, S., Pahan, K., & Subbaiah, P. V. (2019). Dietary lysophosphatidylcholine-EPA enriches both EPA and DHA in the brain: Potential treatment for depression. *J Lipid Res* [Internet], 60(3), 566–578. Available from: </pmc/articles/PMC6399499/>.
23. Fokkema, M. R., Smit, E. N., Martini, I. A., Woltil, H. A., Boersma, E. R., & Muskiet, F. A. J.. Assessment of essential fatty acid and ω3-fatty acid status by measurement of erythrocyte 20:3ω9 (Mead acid), 22:5ω6/20:4ω6 and 22:5ω6/22:6ω3. *Prostaglandins Leukot Essent Fat Acids* [Internet]. 2002 Nov 1 [cited 2021 Feb 19];67(5):345–56. Available from: <http://www.idealibrary.com>
24. Mead Acid - an overview | ScienceDirect Topics [Internet]. [cited 2021 Feb 19]. Available from: <https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/mead-acid>
25. Kinoshita, Y., Yoshizawa, K., Hamazaki, K., Emoto, Y., Yuri, T., Yuki, M. (2016). Dietary effects of mead acid on n-methyl-n-nitrosourea-induced mammary cancers in female sprague-dawley rats. *Biomed Reports* [Internet]. Jan 1 [cited 2021 Feb 19];4(1):33–9. Available from: </pmc/articles/PMC4726903/>
26. Strandvik, B. (2006). Mediterranean diet and cystic fibrosis. *Br J Nutr* [Internet]. Aug [cited 2021 Feb 19];96(2):199–200. Available from: <https://doi.org/10.1079/BJN20061815>
27. Gramlich, L., Ireton-Jones, C., Miles, J. M., Morrison, M., & Pontes-Arruda, A. (2019). Essential Fatty Acid Requirements and Intravenous Lipid Emulsions. *J Parenter Enter Nutr*, 43, 697–707.
28. Kaur, N., Chugh, V., & Gupta, A. K. (2014). Essential fatty acids as functional components of foods- a review [Internet]. Vol. 51, Journal of Food Science and Technology. Springer; [cited 2021 Feb 21]. p. 2289–303. Available from: </pmc/articles/PMC4190204/>
29. Kris-Etherton, P., Eckel, R. H., Howard, B. V., St. Jeor, S., & Bazzarre, T. L. (2001). Lyon diet heart study: Benefits of a Mediterranean-style, National Cholesterol Education Program/American Heart Association Step I dietary pattern on cardiovascular disease. *Circulation* [Internet]. Apr 3 [cited 2021 Feb

- 21];103(13):1823–5. Available from: <http://www.circulationaha.org>
30. Ristic-Medic, D., Kovacic, M., Takic, M., Arsic, A., Petrovic, S., Paunovic, M. (2020). Calorie-Restricted Mediterranean and Low-Fat Diets Affect Fatty Acid Status in Individuals with Nonalcoholic Fatty Liver Disease. [cited 2021 Feb 19]; Available from: <https://dx.doi.org/10.3390/nu13010015>
31. Hagfors, L., Nilsson, I., Sköldstam, L., & Johansson, G. (2005). Fat intake and composition of fatty acids in serum phospholipids in a randomized, controlled, Mediterranean dietary intervention study on patients with rheumatoid arthritis. *Nutr Metab* [Internet]. Oct 10 [cited 2021 Feb 22];2(1):26. Available from: <http://nutritionandmetabolism.biomedcentral.com/articles/10.1186/1743-7075-2-26>
32. Pranger, I. G., Corpeleijn, E., Muskiet, F. A. J., Kema, I. P., Singh-Povel, C., & Bakker, S. J. L.. Circulating fatty acids as biomarkers of dairy fat intake: data from the lifelines biobank and cohort study. *Biomarkers* [Internet]. 2019 May 19 [cited 2021 Feb 22]; 24(4):360–72. Available from: <https://www.tandfonline.com/doi/full/10.1080/1354750X.2019.1583770>
33. Polyether toxins produced by dinoflagellates [Internet]. [cited 2021 Feb 22]. Available from: <https://agris.fao.org/agris-search/search.do?recordID=US201302683668>
34. Miles, C. O., Wilkins, A. L., Munday, J. S., Munday, R., Hawkes, A. D., Jensen, D. J. (2006). Production of 7-epi-pectenotoxin-2 seco acid and assessment of its acute toxicity to mice. *J Agric Food Chem* [Internet]. Feb 22 [cited 2021 Feb 22];54(4):1530–4. Available from: <https://pubs.acs.org/sharingguidelines>
35. Daiguji, M., Satake, M., James, K. J., Bishop, A., MacKenzie, L., Naoki, H. (1998). Structures of new pectenotoxin analogs, pectenotoxin-2 seco acid and 7-epi-pectenotoxin-2 seco acid, isolated from a dinoflagellate and greenshell mussels. *Chem Lett* [Internet]. Sep 24 [cited 2021 Feb 22];(7):653–4. Available from: <https://www.journal.csj.jp/doi/abs/10.1246/cl.1998.653>
36. Vale, P., & Sampayo, M. A. D. M. (2002). Pectenotoxin-2 seco acid, 7-epi-pectenotoxin-2 seco acid and pectenotoxin-2 in shellfish and plankton from Portugal. *Toxicon* [Internet]. [cited 2021 Feb 22];40(7):979–87. Available from: <https://pubmed.ncbi.nlm.nih.gov/12076652/>
37. Chen, J., Li, X., Wang, S., Chen, F., Cao, W., Sun, C.. Screening of lipophilic marine toxins in marine aquaculture environment using liquid chromatography–mass spectrometry. *Chemosphere* [Internet]. 2017 Feb 1 [cited 2021 Feb 22];168:32–40. Available from: <http://dx.doi.org/10.1016/j.chemosphere.2016.10.052>
38. Human Metabolome Database: Showing metabocard for xi-8-Hydroxyhexadecanedioic acid (HMDB0037831) [Internet]. [cited 2020 Sep 15]. Available from: <https://hmdb.ca/metabolites/HMDB0037831>
39. ContaminantDB: xi-8-Hydroxyhexadecanedioic acid [Internet]. [cited 2021 Feb 22]. Available from: <https://contaminantdb.ca/contaminants/CHEM031260>
40. Dictionary of Food Compounds with CD-ROM: Additives, Flavors, and Ingredients - Google Books [Internet]. [cited 2021 Feb 22]. Available from: https://books.google.co.uk/books/about/Dictionary_of_Food_Compounds_with_CD_ROM.html?id=eW20PCXlqv8C&redir_esc=y
41. Kosma, D. K., Parsons, E. P., Isaacson, T., Lü, S., Rose, J. K. C., & Jenks, M. A. (2010). Fruit cuticle lipid composition during development in tomato ripening mutants. *Physiol Plant* [Internet]. May 1 [cited 2021 Feb 22];139(1):107–17. Available from: <http://doi.wiley.com/10.1111/j.1399-3054.2009.01342.x>
42. Peschel, S., Franke, R., Schreiber, L., & Knoche, M.. Composition of the cuticle of developing sweet cherry fruit. *Phytochemistry*. 2007 Apr 1;68(7):1017–25.
43. Holloway, P. J.. Cutins of *Malus pumila* fruits and leaves. *Phytochemistry*. 1973 Dec 1;12(12):2913–20.

Figures



Multivariate statistical modelling of LC-MS data. Plots (top left, centre and right) shows group separations achieved by principal component analysis (PCA), orthogonal partial least squares discriminant analysis (OPLS-DA) and partial least squares discriminant analysis (PLS-DA), respectively. Red circles (●) represent patients with high MDS and green circles (●) represent individuals with low MDS. Plot at the bottom is the resulting variable importance in projection (VIP) plot indicating the 15 most influential metabolites responsible for the observed separation in the PLS-DA model.

Figure 1

See image above for figure legend

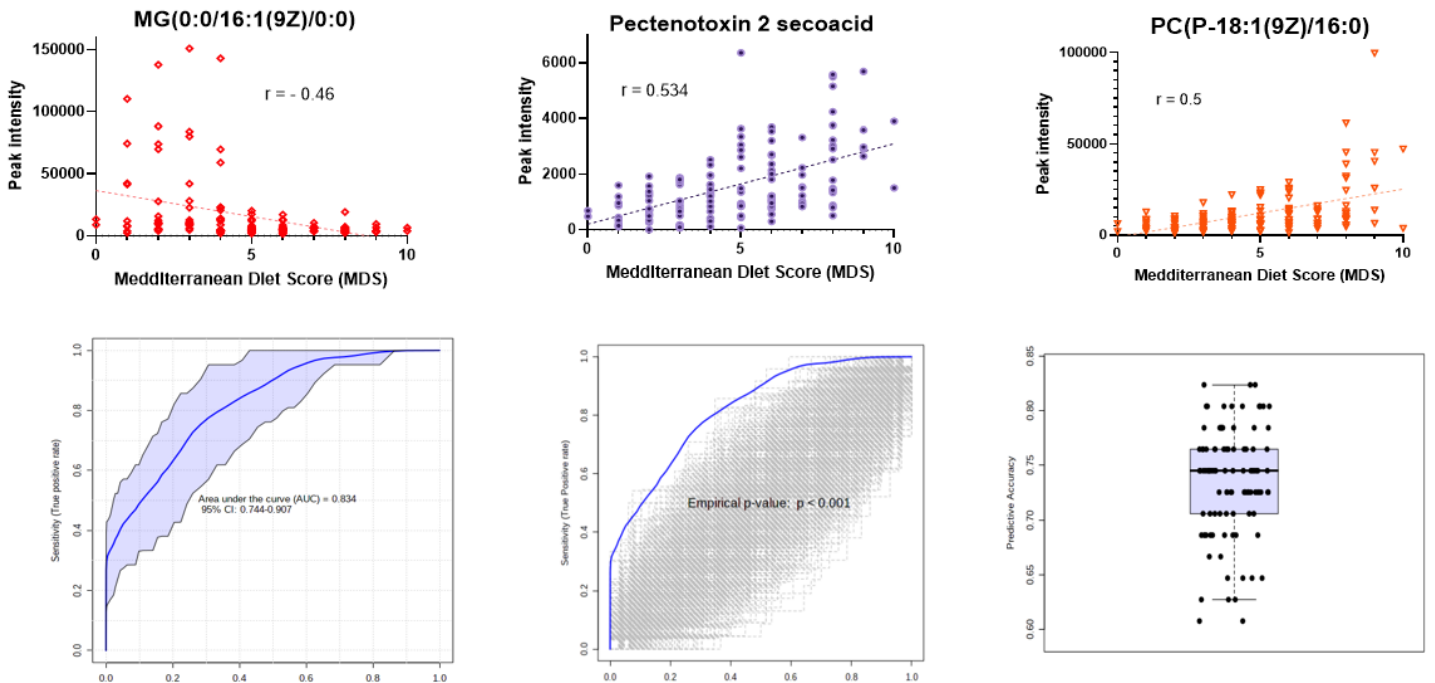


Figure 2

Optimised biomarker performance using logistic regression. Figure shows three individual plasma biomarkers: MG(0:0/16:1(9Z)/0:0), Pectenotoxin 2 secoacid and PC(P-18:1(9Z)/16:0), which were optimised as panel of plasma biomarkers with the resulting logistic regression curve. The correlation of each biomarker is shown against MDS with Spearman r values shown. Also shown is the resulting ROC curve (AUC=0.834) with 95% CI (0.744-907). Also shown is the outcome of 1000 permutation tests, each of which re-assigns the Low/High MDS labels randomly to each sample and performs random sub-sampling cross-validation. None of the 1000 permutation tests were better than the original with the empirical p-value reported as $p < 0.001$.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SUPPLEMENTARYINFORMATION.docx](#)