Phenotypic factors influencing the variation in response of circulating cholesterol level to personalised dietary advice in the Food4me study.

Laura Kirwan[1], Marianne C Walsh[1], Carlos Celis-Morales[2], Cyril F. M. Marsaux[3], Katherine M Livingstone[2], Santiago Navas-Carretero[4,5], Rosalind Fallaize[6], Clare O'Donovan[1], Clara Woolhead[1], Hannah Forster[1], Silvia Kolossa[7], Hannelore Daniel[7], George Moschonis[8] , Yannis Manios[8], Agnieszka Surwillo[9], Magdalena Godlewska[9], Iwona Traczyk[9], Christian A. Drevon[10], Mike J. Gibney[1], Julie A. Lovegrove[6], J. Alfredo Martinez[4,5], Wim H. M. Saris[3], John C. Mathers[2], Eileen R Gibney[1] and Lorraine Brennan[1] on behalf of the Food4me study

[1]UCD Institute of Food and Health, UCD School of Agriculture and Food Science, University College Dublin, Belfield, Dublin 4, Republic of Ireland

[2]Human Nutrition Research Centre, Institute of Cellular Medicine, Newcastle University, Newcastle Upon Tyne, UK

[3]Department of Human Biology, NUTRIM School of Nutrition and Translational Research in Metabolism, Maastricht University Medical Centre + (MUMC+), Maastricht, The Netherlands

[4]Department of Nutrition, Food Science and Physiology, Centre for Nutrition Research, University of Navarra, Pamplona

[5]CIBER Fisiopatogía de la Obesidad y Nutrición (CIBERobn), Instituto de Salud Carlos III, Madrid, Spain

[6]Hugh Sinclair Unit of Human Nutrition and Institute for Cardiovascular and Metabolic Research, University of Reading, Reading, UK

[7]ZIEL Research Center of Nutrition and Food Sciences, Biochemistry Unit, Technische Universität München, München, Germany

[8]Department of Nutrition and Dietetics, Harokopio University, Athens, Greece

[9]National Food & Nutrition Institute (IZZ), Warsaw, Poland

[10]Department of Nutrition, Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Oslo, Norway

Corresponding author: Lorraine Brennan, University College Dublin, Belfield, Dublin 4, Ireland. email: lorraine.brennan@ucd.ie, phone: +353 1 7162811

Keywords: Personalised nutrition, cholesterol, responders, phenotype, fatty acid profile

Short title: Cholesterol response to dietary advice

ABSTRACT

Individual response to dietary interventions can be highly variable. The phenotypic characteristics of those who will respond positively to personalised dietary advice are largely unknown. The objective of this study was to compare the phenotypic profiles of differential responders to personalised dietary intervention, with a focus on total circulating cholesterol. Subjects from the Food4Me multi-centre study were classified as responders or non-responders to dietary advice based on the change in cholesterol level from baseline to month 6, with lower and upper quartiles defined as the responder and non-responder groups, respectively. There were no significant differences between the demographic and anthropometric profiles of the groups. Furthermore, with the exception of alcohol, there was no significant difference in reported dietary intake, at baseline. However, there were marked differences in baseline fatty acid profiles. The responder group had significantly higher levels of stearic acid (18:0, p=0.034) and lower levels of palmitic acid (16:0, p=0.009). Total monounsaturated fatty acids (p=0.016) and total polyunsaturated fatty acids (p=0.008) also differed between the groups. In a stepwise logistic regression model, age, baseline total cholesterol, glucose, five fatty acids and alcohol intake were selected as factors that successfully discriminated responders from non-responders, with sensitivity of 82% and specificity of 83%. The successful delivery of personalised dietary advice may depend on our ability to identify phenotypes that are responsive. The results demonstrate the potential use of metabolic profiles in identifying response to an intervention and could play an important role in the development of precision nutrition.

INTRODUCTION

At a population level, generic dietary advice is provided using a 'one-size-fits-all' approach based on requirements for population groups [1], which ignores inter-individual differences , and therefore nutrient requirements. In addition, individuals' responses to dietary interventions can be highly variable [2; 3; 4]. Demographic characteristics such as sex and age, and factors such as adiposity, physical activity, metabolic profile, and genetic factors contribute to this variation [5]. This phenomenon is well recognised in the medical field with a current emphasis on precision medicine [6]. Considering the reported variation in response to dietary interventions there is now an emerging recognition that this should be considered in development of personalised or precision nutrition. [7; 8]. Personalised nutrition, or dietary advice that has been tailored to an individual, offers the possibility of improving health and reducing risk of diet-related diseases [9]. Many studies suggest that tailored dietary advice is more effective than generic advice, promoting greater improvements in dietary behaviours and related health outcomes such as body weight [9; 10]. A recent meta-analysis reported that personalised interventions were more effective than non-personalised advice, with participants receiving the personalised intervention reducing body weight by 1.8 kg more on average than those receiving the non-personalised advice [9]. However, these studies have not taken individual variability into account and in the longterm the effectiveness of the personalised dietary advice will depend on the ability to tailor advice taking into account knowledge about an individual's potential response to the intervention [11].

The concept of using metabolic profiles to identify responders to dietary interventions is relatively new (Brennan, 2015). However, a number of examples exist in the literature demonstrating the potential of such an approach. O'Sullivan *et al.* (2011) used k-means cluster analysis to identify responders and non-responders to a vitamin D intervention [12]. van Bochove *et al.* (2012) applied k-means clustering to lipoprotein profiles and identified three clusters, two of which responded positively to fenofibrate [13], while Elnenaei *et al.* (2011) identified responders and non-responders to vitamin D and Ca supplementation, based on a baseline metabolomic profile [14]. Metabolomic and transcriptomic profiles have also been used to discriminate between responders and non-responders to an n-3 PUFA supplementation [15]. The objective of this study is to investigate differences in the phenotype and in particular blood lipids of responders and non-responders to personalised nutrition, with a specific focus on changes in circulating cholesterol levels. Using data from the Food4Me personalised dietary intervention study, individuals with borderline high baseline total

cholesterol (> 5 mmol/L) were examined for factors that predict their response to the intervention.

MATERIALS AND METHODS

Subjects were participants in the Food4Me study, a 6-month, web-based randomised control trial conducted in 7 European countries. The aim of the study was to determine whether providing personalised dietary advice leads to improvements in dietary intakes and health outcomes relative to population-based public health messages. The 1,607 adult subjects were randomly assigned to one of four intervention treatment groups – level 0 (standard nonpersonalised dietary and physical activity guidelines), level 1 (personalised advice based on current diet and physical activity), level 2 (personalised advice based on current diet, physical activity and phenotype) and level 3 (personalised advice based on current diet, physical activity, phenotype and genotype) [16]. The control group received conventional, non-personalised advice and so are not considered for this analysis. The study protocol is detailed in Celis-Morales *et al.* [16].

All data were collected remotely following standardized operating procedures. At baseline, participants received study kits by post containing all necessary materials to perform measurements at home. Printed instructions were included and demonstration videos were available on the Food4Me website (http://www.food4me.org). Following measurements at baseline and 3 months, participants received a personalised report. The personalised feedback provided was based on a predefined set of algorithms, including anthropometric, physical activity (Levels 1-3), phenotypic (Levels 2 and 3), and genotypic (Level 3 only) data [16].

Demographic characteristics

The measurement of characteristics including age, country and sex and have been described elsewhere [16]. Having excluded the control group and those with normal total cholesterol levels at baseline (total cholesterol < 5 mmol/L), there were 151 males and 162 females, with a mean age of 46.8 years from 7 European countries, Germany (n=67), Greece (n=48), Ireland (n=39), Netherlands (n=54), Poland (n=30), Spain (n=43) and the United Kingdom (n=32). Subjects were classified as responders and non-responders based on the change in blood cholesterol from baseline to month 6. To achieve this the subjects were firstly stratified

into quartiles based on cholesterol response. Two of the groups, the lower and upper quartiles, were defined as the responders and non-responders, respectively. This resulted in n=78 responders and n=79 non-responders.

Anthropometric measurements

Body weight, height and waist circumference were self-measured and self-reported by participants via the Internet, as described previously [16]. They were provided with clear instructions in text and video format to facilitate accurate measurements and a validation study demonstrated the reliability of these internet-based self-reported anthropometric data [17]. Waist circumference was measured at the midpoint between the lower rib and the iliac crest using the same tape measure. Physical Activity was self-reported using the Baecke questionnaire online [18; 19] based on physical activity during the last month. Physical activity level scores (PAL) were calculated at baseline and month 6, according to the questionnaire protocol.

Dietary intake measurements

Habitual dietary intake was quantified using an online food frequency questionnaire (FFQ) including food items frequently consumed in each of the 7 recruitment sites. The Food4me FFQ has been compared to a paper based FFQ [20] and 4-day weighed food record [21] for both food group and nutrient intakes. Bland Altman analysis showed good agreement between the on-line and paper-based FFQ for both the nutrient and food group level. Cross-classification into exact plus adjacent quartiles ranged from 77 % to 97% at the nutrient level and 77% to 99% at the food group level. For comparison with the weighed food record the mean cross-classification into exact agreement plus adjacent was 80% and 78% for nutrient and food groups respectively. Importantly the energy intake estimated by the FFQ was in agreement with the weighed food record. Overall, indicating that overall the on-line FFQ was a suitable tool for assessing dietary intake.

Fatty acid and carotenoid profiles

Finger-prick blood samples were collected by participants using a test kit provided by Vitas Ltd, Oslo, Norway, as described previously [22]. Each participant filled two Dry Blood Spot

cards (equivalent to five drops of blood or 150 μL of blood per card) at each collection time point. The samples were sent to Vitas (Vitas Ltd, Norway) for measurements of total cholesterol, carotenoids, and 32 fatty acids (FA). The n-3 fatty acid index was calculated as the sum of eicosapentaenoic acid (20:5n-3) and docosahexaenoic acid (22:6n-3). The Δ5 desaturase index (D5D) and Δ6 desaturase index (D6D) are calculated based on key enzymes in the metabolism of polyunsaturated fatty acids. The D5D was calculated as the ratio of arachidonic acid (20:4n-6) to dihomoγlinoleic acid (20:3n-6) the D6D was calculated as the ratio of dihomoγlinoleic acid (20:3n-6) to linoleic acid (18:2n-6).

Ethics

This study was conducted according to the guidelines laid down in the Declaration of Helsinki. The Research Ethics Committees at each participating centre granted ethical approval for the study [16].

Statistical Analysis

The baseline demographic and phenotypic characteristics of the responders and non-responders were compared using generalised linear models. Models were fitted using the GLM (for continuous variables) and GENMOD (for categorical variables) procedures in SAS 9.3 (SAS Institute, Cary NC). To account for multiple comparisons, False Discovery Rate (FDR) adjusted p-values are presented for fatty acid profile data.

To assess whether baseline demographic or phenotypic characteristics can discriminate between responders and non-responders, a stepwise logistic regression procedure was applied in four stages. Firstly, only anthropometric characteristics were included (Model 1). Then baseline cholesterol was added to the model (Model 2). Thirdly, dietary intake data were added to the analysis (Model 3) and lastly, all demographic, anthropometric, dietary intake and biochemical characteristics were included (Model 4). At each stage, the stepwise procedure selected the characteristics that best discriminated between the two groups. Variables were tested using a bootstrapping approach to correct for overoptimism in model fitting. The ability of the models to classify responders and non-responders was assessed using area under the ROC curves. ROC comparisons were performed by using a contrast matrix to take differences of the areas under the empirical ROC curves.

RESULTS

Characteristics of responders and non-responders

Demographic characteristics did not differ significantly between the responder and non-responder groups by country (χ_6^2=5.0, p=0.544, **Table 1**), sex (χ_1^2=0.16, p=0.693, Table 1) or age (p=0.082, Table 1). There was also little difference between the responder and non-responder groups for the anthropometric characteristics measured at baseline (Table 1).

During the intervention period, both groups significantly reduced BMI, weight and waist circumference, with both groups exhibiting similar effect sizes (Table 1). The responders significantly increased their blood omega-3 index, whereas the non-responders did not (mean change Δ = 0.31 versus 0.14, p<0.001).

At baseline, the responders and non-responders had similar dietary intakes of most food groups, with the exception of alcohol (**Table 2**) for which the responders had lower intake (170g day^{-1} versus 258g day^{-1}, p=0.035). Post-intervention, the responders reported reduced intake of dairy (Δ= -59 g day^{-1}, Table 2) and both responders and non-responders reported significantly reduced red meat intake (Δ= -31 and -28 g day^{-1} respectively).

The percentage of participants receiving dietary advice for specific target nutrients was broadly similar (**Table S1**). The most common nutrient targeted at baseline was salt (73% of responders and 59% of non-responders). There was no difference in the percentage of responders and non-responders receiving a dietary message specifically targeted at cholesterol (24% versus 23%, p=0.816), although a greater number of non-responders received a message to increase physical activity (56% of responders versus 73% of non-responders, p=0.027). While the responders had a significant reduction in cholesterol, there was no significant change in physical activity during the intervention period for either group.

At baseline, the responders had higher total cholesterol level than the non-responders (6.09 mmol/L versus 5.54 mmol/L, p<0.001, Table 1). The fatty acid profiles differed between the responders and non-responders at baseline (**Table 3**). There was no difference between the groups for total saturated fatty acids (SFA, p=0.203), but the responders had lower palmitic acid (16:0, p=0.009). At baseline, the responders had significantly lower total monounsaturated fatty acids (MUFA, p=0.016), and in particular lower palmitoleic acid (16:1n-7, p=0.012) and cis-vaccenic acid (18:1n-7, p=0.001). At baseline, the responders had

higher total PUFA (p=0.008), in particular linoleic acid (18:2n-6, p=0.011), eicosadienoic acid (20:2n-6, p=0.006) and docosapentaenoic acid (DPA 22:5n-3, p=0.014). At baseline both groups had similar carotenoids profiles (**Table 4**).

Discriminating between responders and non-responders

When the stepwise logistic regression model was applied using demographic and anthropometric data, age and weight were selected as being important factors in discriminating responders from non-responders (Model 1, **Table 5**). The classification accuracy (as measured by the area under the ROC curve, **Figure 1**) was 0.61, indicating that the demographic and anthropometric data do not provide sufficient discriminatory power. As expected, the classification accuracy improved when the model was adjusted for baseline cholesterol, (Model 2 area under curve=0.76, Table 5, Figure 1). Including dietary intake data (Model 3) did not improve the discriminatory power, with none of the food groups being selected when tested in the stepwise model. When the additional biochemical data were added to the model (Model 4), the key variables selected were baseline levels of cholesterol, glucose, stearic acid, DPA, and eicosenoic acid, each with significant positive coefficients and EPA and trans fatty acids, with significant negative coefficients. Alcohol intake also had a significant negative coefficient in this model that included the biochemical variables. The coefficients of the final logistic regression discriminant model are detailed in **Table 6**. Increases in the variables with positive or negative coefficients were associated with increased or decreased probability of being a responder, respectively. The additional biochemical data significantly improved the classification accuracy (Model 4 area under curve=0.90, Table 5, Figure 1), with increases in the true positive rate (sensitivity) resulting in only a small trade-off with the false positive rate (1- sensitivity). For example, to achieve a sensitivity of 80% for Model 3, the false positive rate is only 10%. This compares with 67% for Model 1 and 44% for Model 2 (Figure 1). Furthermore, it is also worth noting that intervention group was not selected as a discriminant variable indicating that it did not contribute to classification as a responder or non-responder.

DISCUSSION

Identification of sub phenotypes that respond differently to dietary interventions has the possibility to significantly enhance delivery of personalised nutrition. In the current study a baseline phenotype characterised by age, alcohol intake, and levels of stearic acid, DPA, EPA, eicosenoic acid and trans fatty acids, was identified which could discriminate responders and non-responders in 90% of cases. Discriminant analysis has previously been used in dietary intervention studies to test whether metabolic profiles may be used to identify responders and non-responders. In a choline-depletion study, analysis of the baseline metabolomics profile predicted which participants developed liver dysfunction when deprived of dietary choline [23]. Mutch *et al*. 2007 classified responders and non-responders to dietary intervention using linear discriminant analysis on a gene expression snapshot [24]. Here we used a stepwise logistic regression model to select the individual factors that best classified the probability of being a responder. Incorporation of such information into dietary advice strategies has the potential to significantly enhance the success of interventions.

Wide inter-individual variation has been observed in the response of total, LDL and HDL cholesterol to dietary change [25; 26; 27] with little alterations in blood cholesterol for some participants despite significant changes in dietary fatty acid pattern and cholesterol intake [28]. This means that while the population response to a diet can be estimated, the responsiveness of a single individual will have as a result of dietary change is difficult to determine [29]. Lefevre *et al.* [30] observed that variability in the change of serum was related to anthropometric measurements including BMI, waist circumference and body fat percentage. Furthermore, there is a large body of evidence to support the genetic influence on response of plasma cholesterol to dietary interventions [25; 31; 32; 33]. The present study determined a profile which was responsive to dietary advice in terms of lowering cholesterol levels. Overall, this work in conjunction with the literature provides compelling evidence that individual variation and response to interventions needs to be incorporated into dietary advice strategies.

The most marked differences between the responder and non-responder phenotypes were found in their baseline fatty acid profiles. The responders had a lower mean percentage of *trans* fatty acids at baseline. TFAs have been found to increase LDL- and decrease HDL-cholesterol levels [34]. While the responders and non-responders did not differ in their total percentage of SFA, contributions of different SFAs differed. The responders had lower

palmitic acid (16:0) and higher stearic acid (18:0) than the non-responders. A review comparing the risk factors for stearic acid with other saturated fatty acids [35] reported that diets high in stearic acid have favourable effects on LDL cholesterol compared with palmitic acid. However, it has also been reported that stearic acid itself has no cholesterol-enhancing effect in clinically very well controlled exchange of single fatty acids, whereas palmitic, myristic and lauric acids have strong cholesterol-raising effects [36].

The responder group had lower total MUFA, in particular palmitoleic acid (16:1n-7) and cis-vaccenic acid (18:1n-7). A meta-analysis investigating the effects of MUFA on cardiovascular and diabetic risk factors observed no consistent evidence for a relationship between MUFA and total cholesterol [36; 37].The PUFA profiles differed between the responders and non-responders, with a more marked difference in the n-6 PUFAs. The responders had higher levels of linoleic acid (18:2n-6) and eicosadienoic acid (20:2n-6) compared to the non-responders at baseline. Linoleic acid, the primary n-6 PUFA, has been shown to have a cholesterol lowering effect [36; 38] and a recent meta-analysis reported a lower risk of coronary heart disease events and deaths with increasing linoleic acid intake [39]. While the total n-3 PUFAs did not differ between the two groups, the responders had a higher percentage of DPA (22:5n-3). Higher levels of DPA in human blood have been shown to be correlated with lower cholesterol [40]. Overall, the data supports the growing evidence that fatty acids patterns as opposed to single individual fatty acids are important in determining health. Moreover, it supports the importance of adequate intake of PUFAs.

The demographic profiles of the responders and non-responders did not differ, and at baseline, the groups also had similar anthropometric characteristics. Dietary intake at baseline was similar across the two groups, with only alcohol intake differing. As this was a study of the effects of personalised nutrition the dietary advice given to the participants differed between individuals. However, for all the participants, the percentage of subjects receiving dietary advice for specific target nutrients was generally similar. The strengths of this study were that it was a multi-country group with multiple time points allowing analysis of change in response to the intervention. Furthermore, the participants are well phenotyped. A limitation of the study is the unique study design involving personalised nutrition advice which makes replication and prospective analysis in an independent cohort difficult.

An objective of this study was to investigate whether the different types of data were useful in classifying whether an individual will respond to the dietary intervention. Our study has shown that baseline phenotypic data provided more classification power than anthropometric or dietary intake data in classifying responsiveness to personalised dietary advice. While the work identified particular predictive characteristics, it was not our aim to establish causative relationships between the variables. Our study has shown that, in principle, we can predict, *a priori* whether an individual's health status will improve in response to the consumption of a given food/diet. This strengthens the evidence base for the concept that intervention and dietary advice can be personalised with more confidence. Future work should examine the optimal method for incorporation of such data into dietary advice and should pave the way for precision nutrition.

ACKNOWLEDGEMENTS

This project was supported by the European Commission under the Food, Agriculture, Fisheries and Biotechnology Theme of the 7th Framework Programme for Research and Technological Development, grant number 265494. The authors' contributions are as follows: L.K. L.B., E.R.G. and M.C.W. derived the research question for this manuscript, drafted the manuscript and conducted statistical analysis; J. C. M. was the study director of the proof-of-principle study of Food4Me; H.D., I.T., C.A.D., M.G., J.A. L., Y.M., J.A.M. and W.H.M.S. contributed to the design of the proof-of-principle study and were principle investigators for their respective research centre; L.B., R.F., H.F., E.R.G., M.G., S.K., K.M.L., C.F.M.M., C.C.-M., G.M., S. N.-C., C. B.O.'D., A.S., M.C.W. and C.W. contributed to the study design and execution at the research centres. All authors read and approved the final version of the manuscript. C.A.D. is a founder, stock owner, board member and consultant for Vitas Ltd, Oslo, Norway. The other authors have no potential financial or personal conflicts of interest to declare.

FIGURE LEGENDS

Figure 1. ROC curves illustrating the performance of models M1, M2 and M4 at discriminating responders from non-responders. The selected variables in M3 were identical to M2 and so it has not been included. The diagonal reference line represents random discrimination, with points above the line indicating discrimination ability.

REFERENCES

1. de Roos B (2013) Personalised nutrition: ready for practice? *Proceedings of the Nutrition Society* **72**, 48-52.

2. van Ommen B, Keijer J, Kleemann R *et al.* (2008) The challenges for molecular nutrition research 2: quantification of the nutritional phenotype. *Genes & nutrition* **3**, 51-59.

3. Konstantinidou V, Ruiz LAD, Ordovás JM (2014) Personalized Nutrition and Cardiovascular Disease Prevention: From Framingham to PREDIMED. *Advances in Nutrition: An International Review Journal* **5**, 368S-371S.

4. Zeevi D, Korem T, Zmora N *et al.* (2015) Personalized Nutrition by Prediction of Glycemic Responses. *Cell* **163**, 1079-1094.

5. Lampe JW, Navarro SL, Hullar MA *et al.* (2013) Inter-individual differences in response to dietary intervention: integrating omics platforms towards personalised dietary recommendations. *The Proceedings of the Nutrition Society* **72**, 207-218.

6. Schork NJ (2015) Personalized medicine: Time for one-person trials. *Nature* **520**, 609-611.

7. Brennan L (2015) Metabotyping: moving towards personalised nutrition. In *Metabolomics as a Tool in Nutrition Research*, pp. 137-144: Woodhead publishing series in food science, technology and nutrition.

8. Kaput J, Morine M (2012) Discovery-based nutritional systems biology: developing N-of-1 nutrigenomic research. *International journal for vitamin and nutrition research Internationale Zeitschrift fur Vitamin- und Ernahrungsforschung Journal international de vitaminologie et de nutrition* **82**, 333-341.

9. Celis-Morales C, Lara J, Mathers JC (2015) Personalising nutritional guidance for more effective behaviour change. *Proceedings of the Nutrition Society* **74**, 130-138.

10. Curtis PJ, Adamson AJ, Mathers JC (2012) Effects on nutrient intake of a family-based intervention to promote increased consumption of low-fat starchy foods through education, cooking skills and personalised goal setting: the Family Food and Health Project. *British Journal of Nutrition* **107**, 1833-1844.

11. Ryan NM, O'Donovan CB, Forster H *et al.* (2015) New tools for personalised nutrition: The Food4Me project. *Nutrition Bulletin* **40**, 134-139.

12. O'Sullivan A, Gibney MJ, Connor AO *et al.* (2011) Biochemical and metabolomic phenotyping in the identification of a vitamin D responsive metabotype for markers of the metabolic syndrome. *Molecular nutrition & food research* **55**, 679-690.

13. van Bochove K, van Schalkwijk DB, Parnell LD *et al.* (2012) Clustering by plasma lipoprotein profile reveals two distinct subgroups with positive lipid response to fenofibrate therapy. *PloS one* **7**, e38072.
14. Elnenaei MO, Chandra R, Mangion T *et al.* (2011) Genomic and metabolomic patterns segregate with responses to calcium and vitamin D supplementation. *The British journal of nutrition* **105**, 71-79.
15. Rudkowska I, Paradis AM, Thifault E *et al.* (2013) Differences in metabolomic and transcriptomic profiles between responders and non-responders to an n-3 polyunsaturated fatty acids (PUFAs) supplementation. *Genes & nutrition* **8**, 411-423.
16. Celis-Morales C, Livingstone KM, Marsaux CF *et al.* (2015) Design and baseline characteristics of the Food4Me study: a web-based randomised controlled trial of personalised nutrition in seven European countries. *Genes & nutrition* **10**, 450.
17. Celis-Morales C, Livingstone KM, Woolhead C *et al.* (2015) How reliable is internet-based self-reported identity, socio-demographic and obesity measures in European adults? *Genes & nutrition* **10**, 1-10.
18. Baecke JA, Burema J, Frijters JE (1982) A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *The American journal of clinical nutrition* **36**, 936-942.
19. Marsaux CFM, Celis-Morales C, Livingstone KM *et al.* (2016) Changes in Physical Activity Following a Genetic-Based Internet-Delivered Personalized Intervention: Randomized Controlled Trial (Food4Me). *Journal of medical Internet research* **18**, e30.
20. Forster H, Fallaize R, Gallagher C *et al.* (2014) Online dietary intake estimation: the Food4Me food frequency questionnaire. *Journal of medical Internet research* **16**, e150.
21. Fallaize R, Forster H, Macready AL *et al.* (2014) Online dietary intake estimation: reproducibility and validity of the Food4Me food frequency questionnaire against a 4-day weighed food record. *Journal of medical Internet research* **16**, e190.
22. Hoeller U, Baur M, Roos FF *et al.* (2016) Application of dried blood spots to determine vitamin D status in a large nutritional study with unsupervised sampling: the Food4Me project. *British Journal of Nutrition* **115**, 202-211.
23. Sha W, da Costa KA, Fischer LM *et al.* (2010) Metabolomic profiling can predict which humans will develop liver dysfunction when deprived of dietary choline. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **24**, 2962-2975.

24. Mutch DM, Temanni MR, Henegar C *et al.* (2007) Adipose gene expression prior to weight loss can differentiate and weakly predict dietary responders. *PloS one* **2**, e1344.
25. Masson LF, McNeill G, Avenell A (2003) Genetic variation and the lipid response to dietary intervention: a systematic review. *The American journal of clinical nutrition* **77**, 1098-1111.
26. Beynen AC, Katan MB, Van Zutphen LF (1987) Hypo- and hyperresponders: individual differences in the response of serum cholesterol concentration to changes in diet. *Advances in lipid research* **22**, 115-171.
27. Jacobs DR, Anderson JT, Hannan P *et al.* (1983) Variability in individual serum cholesterol response to change in diet. *Arteriosclerosis, Thrombosis, and Vascular Biology* **3**, 349-356.
28. Cox C, Mann J, Sutherland W *et al.* (1995) Individual variation in plasma cholesterol response to dietary saturated fat. *BMJ (Clinical research ed)* **311**, 1260-1264.
29. Denke MA, Adams-Huet B, Nguyen AT (2000) Individual cholesterol variation in response to a margarine- or butter-based diet: A study in families. *JAMA* **284**, 2740-2747.
30. Lefevre M, Champagne CM, Tulley RT *et al.* (2005) Individual variability in cardiovascular disease risk factor responses to low-fat and low-saturated-fat diets in men: body mass index, adiposity, and insulin resistance predict changes in LDL cholesterol. *The American journal of clinical nutrition* **82**, 957-963; quiz 1145-1146.
31. Qi Q, Durst R, Schwarzfuchs D *et al.* (2015) CETP genotype and changes in lipid levels in response to weight-loss diet intervention in the POUNDS LOST and DIRECT randomized trials. *Journal of lipid research* **56**, 713-721.
32. Asztalos B, Lefevre M, Wong L *et al.* (2000) Differential response to low-fat diet between low and normal HDL-cholesterol subjects. *Journal of lipid research* **41**, 321-328.
33. Wallace AJ, Mann JI, Sutherland WH *et al.* (2000) Variants in the cholesterol ester transfer protein and lipoprotein lipase genes are predictors of plasma cholesterol response to dietary change. *Atherosclerosis* **152**, 327-336.
34. Hunter JE (2014) Health and nutrition update on trans fatty acids. *Lipid Technology* **26**, 199-201.
35. Hunter JE, Zhang J, Kris-Etherton PM (2010) Cardiovascular disease risk of dietary stearic acid compared with trans, other saturated, and unsaturated fatty acids: a systematic review. *The American journal of clinical nutrition* **91**, 46-63.

444 36. Müller H, Kirkhus B, Pedersen JI Serum cholesterol predictive equations with special
445 emphasis on Trans and saturated fatty acids. An analysis from designed controlled studies.
446 *Lipids* **36**, 783-791.
447 37. Schwingshackl L, Hoffmann G (2012) Monounsaturated Fatty Acids and Risk of
448 Cardiovascular Disease: Synopsis of the Evidence Available from Systematic Reviews and
449 Meta-Analyses. *Nutrients* **4**, 1989-2007.
450 38. Harris WS, Mozaffarian D, Rimm E *et al.* (2009) Omega-6 Fatty Acids and Risk for
451 Cardiovascular Disease: A Science Advisory From the American Heart Association Nutrition
452 Subcommittee of the Council on Nutrition, Physical Activity, and Metabolism; Council on
453 Cardiovascular Nursing; and Council on Epidemiology and Prevention. *Circulation* **119**, 902-
454 907.
455 39. Farvid MS, Ding M, Pan A *et al.* (2014) Dietary Linoleic Acid and Risk of Coronary
456 Heart Disease: A Systematic Review and Meta-Analysis of Prospective Cohort Studies.
457 *Circulation* **130**, 1568-1578.
458 40. Byelashov OA, Sinclair AJ, Kaur G (2015) Dietary sources, current intakes, and
459 nutritional role of omega-3 docosapentaenoic acid. *Lipid Technology* **27**, 79-82.

460

Table 1. Demographic and phenotypic profiles of responders and non-responders

		Responder				Non-responder				Responder vs Non- responder	
		n	%			n	%			Chi-sq	p-value
	Total	78				79					
Sex	Male	40	51.20%			43	55.13%				
	Female	38	48.70%			36	46.15%			0.16	0.693
Country	Germany	19	24.40%			23	29.11%				
	Greece	10	12.80%			11	13.92%				
	Ireland	11	14.10%			7	8.86%				
	Netherlands	16	20.50%			13	16.46%				
	Poland	4	5.10%			8	10.13%				
	Spain	11	14.10%			14	17.72%				
	United Kingdom	7	9.00%			3	3.80%			5.0	0.544
		Baseline		Δ		Baseline		Δ		p-value for difference	
		Mean	SEM	Mean	SEM	Mean	SEM	Mean	SEM	Baseline	Δ
Age (years)		45.1	1.35			48.2	1.15			0.082	
Height (m)		1.72	0.01			1.74	0.01			0.262	
Weight (kg)		78.1	1.66	**-1.7**	**0.39**	82.6	1.79	**-1.3**	**0.37**	0.065	0.429
BMI (kg/m^2)		26.4	0.52	**-0.6**	**0.14**	27.5	0.56	**-0.4**	**0.13**	0.17	0.495
Waist circumference (m)		0.9	0.015	**-0.02**	**0.005**	0.93	0.015	**-0.02**	**0.005**	0.091	0.764
Physical activity level		1.54	0.012	0.03	0.01	1.53	0.012	0.027	0.01	0.687	0.908
Total Cholesterol (mmol/L)		6.09	0.091	**-2.01**	**0.072**	5.54	0.063	0.47	0.06	**<.001**	**<0.001**
Glucose (mmol/L)		4.13	0.08	**-0.82**	**0.113**	3.88	0.111	**-0.23**	**0.111**	0.934	0.259
Omega-3 index		5.68	0.127	**0.31**	**0.096**	5.69	0.13	0.14	0.109	0.068	**<0.001**

Measurements at baseline and mean change (Δ) between baseline and month 6 are presented as mean ± standard error. P-values were obtained from generalised linear models including the responder group as a factor. Bolded p-values are significant at the 5% level.

Table 2. Baseline dietary intake (g day^{-1}) and change from baseline to month 6 for responders and non-responders

	Responders				Non-responders				P-value for difference	
	Baseline		Δ		Baseline		Δ		Baseline	Δ
	Mean	SEM	Mean	SEM	Mean	SEM	Mean	SEM		
Fruit	445	45	6	33	387	32	37	31	0.289	0.487
Vegetables	234	22	-6	22	229	14	4	15	0.851	0.71
Whole grains	169	18	6	15	127	14	22	12	0.064	0.413
Oily fish	23	3	9	6	21	3	4	3	0.691	0.42
Red meat	95	11	**-31**	**10**	85	7	**-28**	**5**	0.424	0.763
Dairy	337	30	**-59**	**27**	286	28	-16	24	0.214	0.247
Nuts	7	2	0	1	6	1	1	1	0.694	0.638
Alcohol	170	23	-22	25	258	34	-16	35	**0.035**	0.892

Dietary intake at baseline and mean change (Δ) between baseline and month 6 are presented as mean ± standard error. Bolded mean changes are significant at the 5% level. P-values were obtained from generalised linear models including responder group as a factor. Bolded p-values are significant at the 5% level.

6 Table 3. Mean % of blood total fatty acid at baseline for responders and non-responders and mean change from baseline to month 6

			Responders				Non-responders				P-value for difference			
			Baseline		Δ		Baseline		Δ		Baseline		Δ	
			Mean	SEM	Mean	SEM	Mean	SEM	Mean	SEM	p-value	FDR	p-value	FDR
		Trans fatty acids	0.79	0.027	0.04	0.035	0.9	0.028	-0.06	0.039	**0.007**	**0.040**	0.059	0.295
SFA	(14:0)	Myristic	0.78	0.042	-0.07	0.046	0.85	0.048	-0.02	0.044	0.281	0.351	0.398	0.807
SFA	(15:0)	Pentadecyclic	0.21	0.006	0	0.005	0.2	0.006	**0.01**	**0.006**	0.166	0.239	0.046	0.295
SFA	(16:0)	Palmitic	22.89	0.157	-0.04	0.188	23.63	0.229	-0.46	0.36	**0.009**	**0.040**	0.308	0.807
SFA	(17:0)	Margaric	0.32	0.005	**-0.01**	**0.006**	0.31	0.008	-0.01	0.009	0.799	0.799	0.617	0.807
SFA	(18:0)	Stearic	12.81	0.118	**0.67**	**0.156**	12.44	0.129	**0.68**	**0.275**	0.034	0.076	0.978	0.978
SFA	(20:0)	Arachidic	0.2	0.007	**0.15**	**0.032**	0.19	0.007	**0.17**	**0.027**	0.639	0.710	0.621	0.807
MUFA	(16:1n-7)	Palmitoleic	1.26	0.056	-0.08	0.041	1.49	0.072	-0.02	0.059	**0.012**	**0.040**	0.436	0.807
MUFA	(18:1n-9)	Oleic n9	19.21	0.278	-0.34	0.245	19.9	0.241	**-0.84**	**0.328**	0.063	0.126	0.225	0.807
MUFA	(18:1n-7)	Cis-vaccenic	1.34	0.021	**0.22**	**0.046**	1.48	0.039	0.03	0.056	**0.001**	**0.020**	0.01	0.200
MUFA	(20:1)	Eicosenoic	0.26	0.006	**-0.01**	**0.006**	0.25	0.007	-0.01	0.008	0.167	0.239	0.686	0.807
n-3 PUFA	(18:3n-3)	□-linolenic ALA	0.33	0.013	0	0.017	0.34	0.018	0	0.018	0.528	0.621	0.943	0.978
n-3 PUFA	(20:5n-3)	Eicosapentaenoic EPA	0.73	0.045	0.04	0.038	0.82	0.061	0.07	0.056	0.208	0.277	0.661	0.807
n-3 PUFA	(22:5n-3)	Docosapentaenoic DPA	1.41	0.039	0.03	0.027	1.28	0.035	0.01	0.036	**0.014**	**0.040**	0.67	0.807

n-3 PUFA	(22:6n-3)	Docosahexaenoic DHA	2.96	0.1	**0.31**	**0.069**	3.01	0.095	0.09	0.076	0.696	0.733	0.041	0.295
n-6 PUFA	(18:2n-6)	Linoleic	19.92	0.259	**-0.6**	**0.221**	18.96	0.266	-0.61	0.311	**0.011**	**0.040**	0.969	0.978
n-6 PUFA	(18:3n-6)	γlinolenic GLA	0.2	0.01	0.01	0.012	0.23	0.014	0	0.012	0.101	0.182	0.669	0.807
n-6 PUFA	(20:2n-6)	Eicosadienoic	0.22	0.004	0	0.003	0.21	0.004	0	0.004	**0.006**	**0.040**	0.431	0.807
n-6 PUFA	(20:3n-6)	Dihomoγlinolenic DGLA	1.58	0.036	**-0.07**	**0.029**	1.46	0.037	-0.04	0.033	0.024	0.060	0.467	0.807
n-6 PUFA	(20:4n-6)	Arachidonic ARA	8.66	0.152	0.08	0.156	8.32	0.146	-0.15	0.184	0.109	0.182	0.351	0.807
Desaturase index D5D		ARA/ DGLA	5.7	0.17	**0.35**	**0.12**	5.93	0.07	0.12	0.17	0.338		0.102	
Desaturase index D6D		DGLA/ Linoleic	0.08	0.002	0	0.002	0.08	0.001	0.002	0.002	0.442		0.393	
		SFA	37.2	0.22	**0.7**	**0.294**	37.63	0.255	0.38	0.594	0.203		0.639	
		MUFA	22.07	0.299	-0.21	0.256	23.07	0.282	**-0.83**	**0.372**	**0.016**		0.166	
		PUFA	36	0.338	-0.21	0.361	34.66	0.365	-0.63	0.565	**0.008**		0.533	
		PUFA n-3	5.43	0.156	0.37	0.2	5.46	0.155	0.23	0.2	0.872		0.419	
		PUFA n-6	30.58	0.315	-0.57	0.317	29.18	0.313	-0.43	0.317	**0.002**		0.747	
		n-3 / n-6	0.18	0.006	-0.017	0.004	0.19	0.006	0.009	0.004	0.181		0.221	

Fatty acid percentage at baseline and mean change (Δ) between baseline and month 6 are presented as mean ± standard error. P-values were obtained from generalised linear models including responder group as a factor. FDR adjusted p-values control for false discovery rate. Bolded p-values are significant at the FDR 5% level. The Δ5 desaturase (D5D) was calculated as the ratio of arachidonic acid (20:4n-6) to dihomoγlinoleic acid (20:3n-6). The Δ6 desaturase (D6D) was calculated as the ratio of dihomoγlinoleic acid (20:3n-6) to linoleic acid (18:2n-6).

Table 4. Mean blood carotenoid levels (μmol/L) for responders and non-responders at baseline

	Responders				Non-responders				P-value for difference	
	Baseline		Δ		Baseline		Δ		Baseline	Δ
	Mean	SEM	Mean	SEM	Mean	SEM	Mean	SEM		
Lutein	0.23	0.013	**-0.03**	**0.012**	0.25	0.014	**-0.03**	**0.01**	0.328	0.74
Zeaxanthin	0.06	0.004	**-0.02**	**0.004**	0.05	0.003	**-0.01**	**0.003**	0.525	0.282
βCryptoxanthin	0.24	0.028	**-0.08**	**0.02**	0.19	0.018	-0.01	0.017	0.098	**0.022**
αCarotene	0.14	0.014	-0.01	0.018	0.11	0.011	0.01	0.008	0.146	0.448
βCarotene	0.45	0.035	-0.05	0.034	0.4	0.035	0.02	0.021	0.276	0.098
Lycopene	0.55	0.027	**-0.07**	**0.026**	0.54	0.033	-0.01	0.038	0.863	0.225
Total Carotenoids	1.67	0.078	**-0.21**	**0.072**	1.54	0.08	-0.03	0.071	0.263	0.082

Carotenoid levels at baseline and mean change (Δ) between baseline and month 6 are presented as mean ± standard error. P-values were obtained from generalised linear models containing responder group as a factor. Bolded p-values are significant at the 5% level.

Table 5. Examining the ability to classify responders and non-responders.

	Area	SE	p-value[a]	Asymptotic 95% Confidence Interval	p-value[b]
M1: Anthropometric data only	0.61	0.045	0.014	0.53 – 0.70	
M2: M1 plus baseline cholesterol	0.76	0.037	<0.001	0.69 – 0.836	0.0007
M3: M2 plus dietary intake data	0.76	0.037	<0.001	0.69 – 0.836	0.999
M4: M3 plus biochemical data	0.90	0.026	<0.001	0.85 – 0.95	0.0003

Area under the ROC curve (AUC). The area measures the accuracy, or discrimination ability, to classify responders and non-responders.

Area under the curve is presented as area ± standard errors.

[a] Null hypothesis: true area = 0.5

[b] P-value for comparison of C-statistic versus previous model

Table 6. List of discriminating parameters.

			Estimate	Standard Error	t-value	p-value
		Constant	30.56	6.347	23.17	0.001
		Baseline cholesterol	2.95	0.583	25.55	0.001
		Baseline glucose	1.02	0.354	8.34	0.10
		Age	-0.06	0.0232	6.67	0.016
SFA	(18:0)	Stearic acid	0.62	0.253	6.03	0.025
MUFA	(20:1)	Eicosenoic acid	13.53	5.16	6.88	0.007
n-3 PUFA	(22:5n-3)	Docosapentaenoic acid (DPA)	4.51	1.04	18.76	0.001
n-6 PUFA	(20:5n-3)	Eicosapentaenoic acid (EPA)	-2.73	0.717	14.53	0.001
		Trans Fatty acids	-3.03	1.054	8.27	0.010
		Alcohol intake	0.0033	0.0011	8.25	0.042

Stepwise logistic regression discriminant analysis. Estimates are on the logit scale. This is the final model selected using stepwise selection procedure including all demographic, anthropometric, dietary intake, fatty acids and carotenoids as potential predictors. The logistic regression model estimates the probability of being a responder. A positive coefficient for an independent variable implies an increased probability of being a responder with increasing values of the variable.

Figures

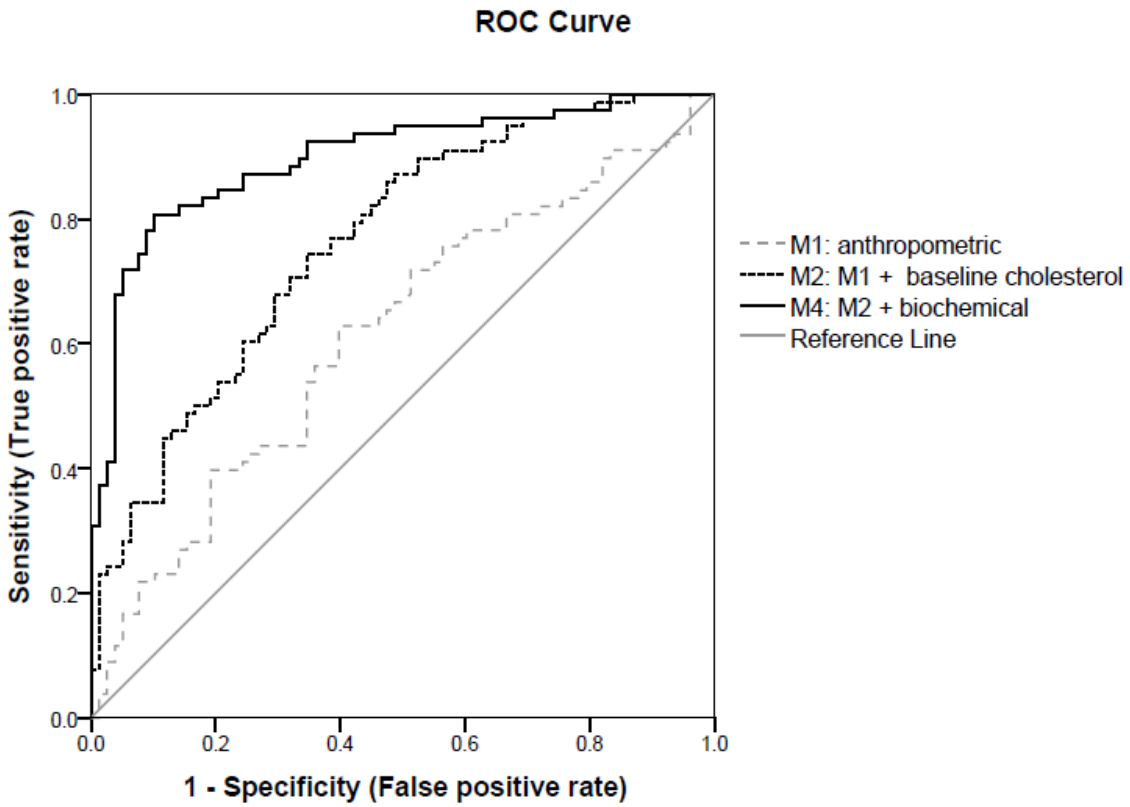

Figure 1. ROC curves illustrating the performance of models M1, M2 and M4 at discriminating responders from non-responders. The selected variables in M3 were identical to M2 and so it has not been included. The diagonal reference line represents random discrimination, with points above the line indicating discrimination ability.

Supplementary Material

Table S1. Percentage of participants receiving dietary advice for specific target nutrients at baseline

Message	Responders	Non-responders	p-value *
Nutrient			
Salt	73%	59%	0.073
Saturated fat	37%	33%	0.575
Fibre	32%	29%	0.690
Carotenoids	28%	30%	0.765
Folate	26%	27%	0.893
Cholesterol	24%	23%	0.816
Unsaturated fat	21%	19%	0.810
Omega 3	18%	23%	0.453
Reduce total fat	14%	11%	0.611
Increase calcium	6%	16%	0.056
Body weight			
Increase PAL	56%	73%	**0.027**
Bodyweight & cholesterol	55%	65%	0.229
Reduce BMI	31%	43%	0.113
Reduce waist circumference	18%	34%	**0.022**

* P-values were obtained from logistic regression models including responder group as a factor. Bolded p-values are significant at the 5% level