Metabotyping for the development of tailored dietary advice solutions in a European population: the Food4Me study

Clare B. O'Donovan[1], Marianne C. Walsh[1], Clara Woolhead[1], Hannah Forster[1], Carlos Celis-Morales[2], Rosalind Fallaize[3], Anna L. Macready[3], Cyril F. M. Marsaux[4], Santiago Navas-Carretero[5,6], S. Rodrigo San-Cristobal[5], Silvia Kolossa[7], Lydia Tsirigoti[8], Christina Mvrogianni[8], Christina P. Lambrinou[8], George Moschonis[8], Magdalena Godlewska[9], Agnieszka Surwillo[9], Iwona Traczyk[10], Christian A. Drevon[11], Hannelore Daniel[7], Yannis Manios[8], J. Alfredo Martinez[5,6,12], Wim H. M. Saris[4], Julie A. Lovegrove[3], John C. Mathers[2], Michael J. Gibney[1], Eileen R. Gibney[1] and Lorraine Brennan[1]*

[1]*School of Agriculture & Food Science, Institute of Food & Health, University College Dublin, Dublin 4, Ireland*
[2]*Human Nutrition Research Centre, Institute of Cellular Medicine, Newcastle University, Newcastle NE4 5PL, UK*
[3]*Hugh Sinclair Unit of Human Nutrition and Institute for Cardiovascular and Metabolic Health, University of Reading, Reading RG6 6AR, UK*
[4]*Department of Human Biology, NUTRIM School of Nutrition and Translational Research in Metabolism, Maastricht University Medical Centre + (MUMC +), Maastricht 6200 MD, The Netherlands*
[5]*Department of Nutrition, Food Science and Physiology, University of Navarra, 31008 Pamplona, Spain*
[6]*Fisiopatología de la Obesidad y Nutrición (CIBERobn), Instituto de Salud Carlos III, 28029 Madrid, Spain*
[7]*Research Center of Nutrition and Food Sciences (ZIEL), Biochemistry Unit, Technische Universität München, Munich 85354, Germany*
[8]*Department of Nutrition and Dietetics, Harokopio University, Athens 17671, Greece*
[9]*National Food & Nutrition Institute, Warsaw 02-903, Poland*
[10]*Department of Human Nutrition, Faculty of Health Science, Medical University of Warsaw, Warsaw 02-091, Poland*
[11]*Department of Nutrition, Institute of Basic Medical Sciences, University of Oslo, 0372 Oslo, Norway*
[12]*Instituto de Investigación Sanitaria de Navarra (IDISNA), 31008 Pamplona, Spain*

Abstract

Traditionally, personalised nutrition was delivered at an individual level. However, the concept of delivering tailored dietary advice at a group level through the identification of metabotypes or groups of metabolically similar individuals has emerged. Although this approach to personalised nutrition looks promising, further work is needed to examine this concept across a wider population group. Therefore, the objectives of this study are to: (1) identify metabotypes in a European population and (2) develop targeted dietary advice solutions for these metabotypes. Using data from the Food4Me study (*n* 1607), *k*-means cluster analysis revealed the presence of three metabolically distinct clusters based on twenty-seven metabolic markers including cholesterol, individual fatty acids and carotenoids. Cluster 2 was identified as a metabolically healthy metabotype as these individuals had the highest Omega-3 Index (6·56 (sd 1·29) %), carotenoids (2·15 (sd 0·71) μM) and lowest total saturated fat levels. On the basis of its fatty acid profile, cluster 1 was characterised as a metabolically unhealthy cluster. Targeted dietary advice solutions were developed per cluster using a decision tree approach. Testing of the approach was performed by comparison with the personalised dietary advice, delivered by nutritionists to Food4Me study participants (*n* 180). Excellent agreement was observed between the targeted and individualised approaches with an average match of 82 % at the level of delivery of the same dietary message. Future work should ascertain whether this proposed method could be utilised in a healthcare setting, for the rapid and efficient delivery of tailored dietary advice solutions.

Key words: Personalised nutrition: Dried blood spots: Cluster analyses: Metabotyping: Targeted nutrition

Early definitions of personalised nutrition were gene focused; however, in recent times, the definition has been extended and now incorporates the concept of levels[1]. This reworked definition of personalised nutrition now includes the following: level 1 personalised advice, which involves delivering personalised advice based on dietary intake; level 2 personalised

---

Abbreviation: DBS, dry blood spot.

* Corresponding author: L. Brennan, email lorraine.brennan@ucd.ie

advice, which involves personalised advice based on diet and phenotypic markers such as blood markers and BMI; and level 3 personalised advice, which builds on the previous levels and includes diet, phenotype and genotype information[2]. Although such definitions focus on personalised advice delivered at an individual level, there is an emerging concept that has gained momentum in recent years, where dietary advice can be tailored to specific groups of individuals and is referred to as targeted nutrition[3–5].

These groups of individuals have similar characteristics and are referred to as metabotypes[6]. There are numerous examples of metabotyping in the medical literature where it has been utilised to sub-group patients with diseases with differential symptomology[7–10]. For example, several studies have used cluster analysis to identify sub-groups of patients with characteristic phenotypes of asthma, a disease that is very heterogeneous in nature[11–14]. Metabotyping has also been used to identify groups of individuals with differing responses to drug treatments[15–17] and dietary interventions[18–20].

However, although there are many examples of identifying groups of similar individuals in the population[7–9,21], the evidence base for developing tailored health solutions for these groups is weak. Previous work from our group demonstrated a framework for the delivery of targeted nutrition advice to metabolically similar groups or metabotypes in the population[22]. In this study, three distinctly different metabotypes were identified on the basis of four routinely measured markers of metabolic health including blood TAG, total cholesterol, HDL-cholesterol and glucose (n 896). Using a decision tree approach, targeted dietary advice messages were developed on the basis of the characteristics of each cluster. Good agreement was observed between the targeted dietary advice method and an individualised method without the need for collection of detailed dietary data[22]. Overall, this previous work demonstrated the potential of the metabotyping approach to deliver appropriate tailored dietary advice at a group level with minimal data collection required.

In the current study, this concept is further advanced using data from the Food4Me study, a personalised nutrition intervention study[23]. In Food4Me, participants received personalised advice based on the three levels of personalisation, delivered by trained nutritionists, and thus provides a valuable resource for testing the targeted nutrition approach[23]. Therefore, the objectives of this study were to: (1) identify metabotypes in a European population group and (2) develop targeted dietary advice solutions for these metabotypes and compare them with personalised dietary advice given within the Food4Me study.

Methods

Study design and ethical approval

As part of the Food4Me project (CinicalTrials.gov no.: NCT01530139, https://clinicaltrials.gov/ct2/show/NCT01530139), a proof-of-principle study was conducted, which compared the effectiveness of personalised nutrition advice, based on the three levels of personalisation, on health-related outcomes, compared

with generic healthy eating advice. This was an internet-based study, designed to emulate a personalised nutrition service, and was conducted in seven research centres across Europe (Ireland, UK, the Netherlands, Spain, Germany, Poland, Greece). Ethical approval was obtained from the Research Ethics Committees at each university or research centre. This study was conducted according to the guidelines laid down in the Declaration of Helsinki, and all procedures involving human subjects were approved by the Research Ethics Committees at each university or research centre. Participants (>18 years) (n 1607) were randomised into one of four groups: control group received general European-based healthy eating guidelines; level 1 participants received personalised advice based on their dietary intake; level 2 participants received personalised advice based on their diet and phenotype; and level 3 participants received advice on their diet, phenotype and genotype. More details on the overall study design and the main outcomes of the personalised nutrition randomised controlled trial can be found elsewhere[23,24]. Written informed consent was obtained from all subjects.

Data collection and personalised feedback

All data were self-collected by participants using detailed instructions provided by researchers and online video demonstrations. A more detailed description of the data collection methods is reported elsewhere[23]. In brief, habitual dietary intake was assessed using the online Food4Me FFQ, which was previously developed and validated for the purposes of the study[25,26]. The foods included in the FFQ were aggregated to form thirty-two food groups. The list of the foods contributing to each of the food groups is found in the online Supplementary Table S1. Participants were provided with a measuring tape to perform anthropometric measures including weight (kg), height (m) and circumferences including waist (cm), hip (cm) and thigh (cm); all were collected according to standard previously published protocols[23]. A validation study was conducted to assess the accuracy of these measurements and strong correlation coefficients (0·983 for BMI and 0·993 for weight) were observed between the self-reported measurements and measurements performed face-to-face by researchers[27].

Metabolic markers were measured by finger-prick blood samples collected by participants using a collection pack provided by Vitas Ltd. Participants were asked to fast 8 h before collection in the morning and to fill two dry blood spot (DBS) cards (five drops of blood or 150 μl of blood per card). Once filled, cards were left to dry for 2–4 h at room temperature and placed in an airtight aluminium bag with a drying sachet and returned by post to their corresponding recruiting centre. The samples were then sent via courier service to Vitas, where the following metabolic markers were measured: total cholesterol, carotenoids (lutein, zeaxanthin, β-cryptoxanthin, α-carotene, β-carotene, lycopene) and twenty fatty acids, as shown in Table 1. The metabolic markers were measured using the following methods: cholesterol (LC-UV), carotenoids (HPLC-diode array detector-MS/MS) and fatty acids (GC-flame ionisation detector).

Participants randomised to levels 1, 2 and 3 received personalised reports based on decision trees to allow for the delivery of systematic tailored advice. The personalised reports were

sent via email at months 0, 3 and 6. Standard operating procedures were developed for use of the decision trees and these were standardised across the seven recruitment centres to ensure consistency in the personalised advice given across all centres. Those individuals in level 1 received feedback based on their current dietary intake and physical activity levels. Level 2 participants received feedback based on their current diet, physical activity levels and phenotypic measures such as anthropometry and metabolic markers (total cholesterol, carotenoids including lutein, zeaxanthin, β-cryptoxanthin, α-carotene, β-carotene, lycopene and twenty fatty acids, as shown in Table 1). Level 3 participants received the same feedback as level 2 with the addition of genotypic information (information on *FTO/FADS1/TCF7L2/ApoE(e4)/MTHFR*). The final section of the report contained a personalised goals section where participants were given three nutrient-related goals. The personalised goals were selected by a predefined ranking system, where those nutrients and metabolic markers that most warranted change were prioritised. Participants were asked to focus on making changes to these three nutrients in the personalised reports in line with the patient-centred counselling models for facilitating behaviour change[28].

Development and testing of targeted dietary advice

Targeted dietary advice was developed for each cluster based on the characteristics of the cluster and using a decision tree process. Two decision trees were developed per cluster based on the following: (1) metabolic markers and anthropometric information and (2) dietary information. This resulted in forty-nine messages for cluster 1, twenty messages for cluster 2 and twenty-four messages for cluster 3. As there are no defined cut-offs for total saturated fat (%) from DBS data, cluster 1 was described as high saturated fat, cluster 2 as low saturated fat and cluster 3 as medium saturated fat based on the mean values across the clusters, as shown in Table 1.

The appropriateness of the targeted dietary advice developed per cluster was then tested by comparison with the three nutrient-related goals, which were delivered to all level 2 participants (*n* 180) by trained nutritionists, as part of their personalised feedback reports. The development of these personalised feedback reports has been published elsewhere[29]. The agreement between the two methods was assessed based on the following questions:

(1) How many of the nutrient-related goals given as part of the personalised advice reports within the Food4Me study were given as part of the targeted dietary advice derived from this study?

(2) How many dietary messages were given as part of the targeted dietary advice in comparison with the personalised advice within Food4Me? (That is number of messages given as per the targeted dietary advice.)

Statistics

Baseline data were analysed using SPSS software package version 20.0 for Windows (SPSS, Inc.). In all, twenty-seven metabolic markers including cholesterol and individual carotenoids and fatty acids from the DBS analysis were chosen for clustering, as presented in Table 1. A full data set was available for 1354 participants (see online Supplementary Table S3). Following standardisation using *z*-scores, two-step cluster analysis revealed the presence of three clusters, and *k*-means cluster analysis was then used to characterise the clusters. The differences between the clusters were assessed using one-way ANOVA with Bonferroni *post hoc* tests. χ^2 distributions were used to assess categorical variables across the clusters including sex and country. As age, sex, BMI and country were significantly different across the clusters, these variables were controlled for in the general linear models with Bonferroni *post hoc* tests. *P* values were also adjusted for multiple comparisons using the Bonferroni approach.

Results

Characterisation of the clusters

Three clusters were identified in the Food4Me population (Table 1). Cluster 1 (*n* 326) was the group with the highest cholesterol, highest circulating *trans*-fatty acids (0·85 (SD 0·25)%) and lowest Omega-3 Index (5·16 (SD 0·93)%). Cluster 2 (*n* 433) was the most metabolically healthy group as they had the highest average Omega-3 Index (6·56 (SD 1·29)%), highest total carotenoid concentrations (2·15 (SD 0·71) μM) and lowest total saturated fat. Cluster 3 subjects (*n* 595) had the lowest average cholesterol concentrations (4·25 (SD 0·78) mM) and highest stearic acid (Table 1). Age was significantly different across the groups, with clusters 1 and 2 being older on average (Table 2). BMI and waist circumference were also significantly different across the clusters. Cluster 1 had the highest BMI of 27·7 (SD 5·3) kg/m² and waist circumference of 0·93 (SD 0·14 m, whereas participants in cluster 2 had the lowest BMI and waist circumference (Table 2). With the exception of the Netherlands and the UK, the distribution of nationality differed significantly across the clusters.

Reported dietary intakes across the clusters are presented in Table 3. There were no differences in total energy intake and macronutrients across the clusters. However, percentage energy contribution from alcohol and PUFA was found to be significantly different (*P* = 0·048). Furthermore, intakes of many micronutrients differed significantly across the clusters, including fat-soluble vitamins A, D and E, as well as some water-soluble vitamins such as folate, and vitamin C. Participants in cluster 1 had the higher percentage contribution of energy from alcohol (4·2 (SD 4·5)%) compared with individuals in clusters 2 and 3. The diets of cluster 2 participants were considered to be healthier as these individuals had the highest intakes of dietary fibre (32 (SD 15) g), fat-soluble vitamins D and E, folate and vitamin C.

Intakes of the food groups savouries (*P* = 1·27 × 10⁻⁴), fruit (*P* = 1·39 × 10⁻⁸), fish, fish dishes and products (*P* = 8·16 × 10⁻⁴) differed significantly between the clusters, as illustrated in Table 4. Similar to their nutrient intakes, participants in cluster 2 had the healthiest food intakes with the lowest intakes of savouries (11 (SD 13) g) and white bread/rolls/scones/croissants (34 (SD 73) g) and highest intakes of yogurt (91 (SD 107) g), fruit (355 (SD 306) g), fish, fish dishes and products (71 (SD 53) g).

Table 1. Clustering variables and other metabolites*
(Mean values and standard deviations)

Clustering variables	Cluster 1 (n 326)		Cluster 2 (n 433)		Cluster 3 (n 595)		
	Mean	SD	Mean	SD	Mean	SD	P
Total cholesterol (mM)	5.01†[2,3]	1.02	4.79[1,3]	0.91	4.25‡[1,2]	0.78	1.30×10^{-37}
Fatty acids (%)							
Myristic acid C14:0	1.07†[2,3]	0.52	0.69[1,3]	0.24	0.59‡[1,2]	0.20	5.06×10^{-92}
Pentadecyclic acid C15:0	0.22[3]	0.07	0.22[3]	0.06	0.18‡[1,2]	0.04	1.03×10^{-32}
Palmitic acid C16:0	24.77†[2,3]	1.87	22.62‡[1,3]	1.48	22.94[1,2]	1.45	1.39×10^{-76}
Palmitoleic acid C16:1	1.82†[2,3]	0.56	1.16[1,3]	0.37	1.02‡[1,2]	0.34	2.22×10^{-137}
Margaric acid C17:0	0.30‡[2]	0.06	0.34†[1,3]	0.06	0.31[2]	0.06	1.27×10^{-16}
Stearic acid C18:0	12.07‡[2,3]	1.12	12.79[1,3]	1.00	13.59†[1,2]	1.12	8.18×10^{-82}
cis-vaccenic acid C18:1 cis	1.52†[2,3]	0.32	1.42‡[1]	0.25	1.43[1]	0.23	1.47×10^{-7}
Oleic acid C18:1	20.72†[2,3]	2.4	18.06‡[1,3]	1.65	18.80[1,2]	1.86	1.02×10^{-70}
Arachidic acid C20:0	0.18‡[2,3]	0.06	0.20[1,3]	0.07	0.23[1,2]	0.09	1.02×10^{-19}
Eicosenoic acid C20:1	0.25‡[2,3]	0.06	0.27[1,3]	0.06	0.28†[1,2]	0.06	2.54×10^{-14}
Total saturated fat§	37.91‡[2,3]	2.5	36.11†[1,3]	1.97	37.11[1,2]	1.92	2.68×10^{-30}
Trans-fatty acids	0.85†[2,3]	0.25	0.79[1,3]	0.24	0.75‡[1,2]	0.21	1.07×10^{-10}
α-Linolenic acid C18:3n-3	0.39†[3]	0.19	0.37[3]	0.12	0.28‡[1,2]	0.12	6.28×10^{-39}
EPA C20:5n-3	0.66[2,3]	0.35	1.06†[1,3]	0.65	0.55‡[1,2]	0.27	1.68×10^{-67}
DPA C22:5n-3	1.24‡[2,3]	0.34	1.56†[1,3]	0.33	1.35[1,2]	0.37	2.81×10^{-37}
DHA C22:6n-3	2.57‡[2,3]	0.72	3.53†[1,3]	0.88	2.87[1,2]	0.76	1.90×10^{-60}
Omega-3 Index‖	5.16[2,3]	0.93	6.56†[1,3]	1.29	5.41[1,2]	0.92	5.14×10^{-80}
Linoleic acid C18:2n-6	17.55‡[2,3]	2.09	20.10†[1,3]	2.28	19.64[1,2]	2.10	8.71×10^{-58}
γ-Linolenic acid C18:3n-6	0.24†[2,3]	0.10	0.16‡[1,3]	0.07	0.18[1,2]	0.07	1.10×10^{-36}
Eicosadienoic acid C20:2 n6	0.21‡[2,3]	0.04	0.22[1,3]	0.04	0.24†[1,2]	0.04	8.04×10^{-38}
Dihomo-γ-linolenic acid C20:3n-6	1.54[2,3]	0.34	1.41‡[1,3]	0.32	1.60†[1,2]	0.33	6.00×10^{-19}
Arachidonic acid C20:4n-6	7.93‡[2,3]	1.46	8.64[1,3]	1.23	9.44†[1,2]	1.34	3.29×10^{-56}
Carotenoids (μM)							
α-Carotene	0.08[2]	0.07	0.21†[1,3]	0.17	0.08[2]	0.05	4.30×10^{-84}
β-Carotene	0.28[2]	0.17	0.66†[1,3]	0.36	0.27‡[2]	0.14	1.48×10^{-132}
β-Cryptoxanthin	0.14‡[2]	0.12	0.29†[1,3]	0.22	0.16[2]	0.11	1.04×10^{-48}
Lutein	0.20[2,3]	0.09	0.29†[1,3]	0.15	0.18‡[1,2]	0.08	4.56×10^{-55}
Lycopene	0.53[2]	0.24	0.65†[1,3]	0.31	0.50‡[2]	0.23	7.68×10^{-19}
Zeaxanthin	0.05[2]	0.03	0.06†[1,3]	0.04	0.04‡[2]	0.03	1.12×10^{-18}
Total carotenoids¶	1.28[2]	0.46	2.15†[1,3]	0.71	1.21‡[2]	0.40	1.90×10^{-145}

[1,2,3] Superscripts indicate which cluster is significantly different using Bonferroni *post hoc* tests for pairwise comparison between clusters.
* Subjects were clustered based on the above variables. One-way ANOVA was used to examine the differences across the clusters.
† Highest values across the clusters.
‡ Lowest values across the clusters.
§ Total saturated fat was calculated as the sum of myristic acid (C14:0), palmitic acid (C16:0) and stearic acid (C18:0).
‖ Omega-3 Index was calculated by the following formula: Omega-3 Index = 1.4473 + 0.8303 × (EPA + DPA + DHA).
¶ Total carotenoids was calculated by the following formula: total carotenoids = α-carotene + β-carotene + lutein + zeaxanthin + β-cryptoxanthin + lycopene.

Table 2. Demographical information across the clusters*
(Mean values and standard deviations; percentages and numbers)

Demographics	Cluster 1		Cluster 2		Cluster 3		
	Mean	SD	Mean	SD	Mean	SD	P
Age (years)	44†[3]	13	43	13[3]	36‡[1,2]	12	7.90×10^{-24}
BMI (kg/m^2)	27.7†[2,3]	5.3	23.9‡	3.9[1,3]	25.4[1,2]	4.7	3.04×10^{-26}
WC (m)	0.93†[2,3]	0.14	0.82‡	0.11[1,3]	0.85[1,2]	0.13	8.48×10^{-32}
Sex (M/F)	161/165[3]		141/292[3]		266/329[1,2]		3.87×10^{-6}
Frequency	%	n	%	n	%	n	
Germany	19.9	65	19.2	83	7.9	47	1.19×10^{-8}
Greece	13.2	43	5.5	24	20.0	119	2.39×10^{-10}
Ireland	15.6	51	20.6	89	10.6	63	5.34×10^{-5}
The Netherlands	16.3	53	14.1	61	14.6	87	0.693
Poland	15.0	49	17.1	74	10.6	63	0.008
Spain	8.3	27	8.5	37	21.3	127	1.08×10^{-10}
UK	11.7	38	15.0	65	15.0	89	0.325

WC, waist circumference; M, male; F, female.
[1,2,3] Superscripts indicate which cluster is significantly different using Bonferroni *post hoc* tests for pairwise comparison between clusters.
* One-way ANOVA was used to examine the differences across the clusters with exception of sex and country where χ^2 was used instead.
† Highest values across the clusters.
‡ Lowest values across the clusters.

Table 3. Dietary intakes across the clusters*
(Mean values and standard deviations)

Nutrients	Cluster 1		Cluster 2		Cluster 3		
	Mean	SD	Mean	SD	Mean	SD	P†
Energy (kJ)	11 370	5064	10 270	4015	10 816	4592	1·00
Total fat (%)	35·9	5·8	36·1	6·3	35·7	5·8	1·00
SFA (%)	14·5	3·2	14·0	3·3	14·0	3·0	1·00
MUFA (%)	13·6	3·0	13·7	3·3	13·9	3·1	0·864
PUFA (%)	5·7[2]	1·4	6·1[1,3]	1·5	5·6§[2]	1·4	0·048
Protein (%)	16·7	3·5	17·1	4·0	17·2	3·5	0·576
Carbohydrate (%)	45·3	7·3	46·1	7·9	46·3	7·4	1·00
Sugars (%)	20·6	6·2	22·1	6·3	20·6	5·3	1·00
Alcohol (%)	4·2‡[2,3]	4·5	2·9§[1]	3·5	3·1[1]	3·4	$1·85 \times 10^{-3}$
Salt (g)‖	8	4	7	3	7	4	1·00
Fibre (g)‖	30[2]	14	32‡[1,3]	15	28§[2]	15	$9·60 \times 10^{-6}$
Vitamin A (µg)‖	1720[3]	1150	1884‡[3]	1048	1500§[1,2]	900	$1·98 \times 10^{-5}$
Vitamin D (µg)‖	6[2]	9	8‡[1,3]	18	5§[2]	5	$3·33 \times 10^{-7}$
Vitamin E (mg)‖	15[2]	13	20‡[1,3]	35	15§[2]	21	$1·20 \times 10^{-2}$
Carotene (µg)‖	6243[2]	10534	7313‡[1,3]	5257	5015§[2]	3557	$1·02 \times 10^{-6}$
Retinol (µg)‖	1340	10175	665	511	664	576	1·00
Thiamine (mg)‖	4	10	5	9	4	6	0·600
Riboflavin(mg)‖	4	7	4	8	3	6	1·00
Folate (µg)‖	424[2]	200	443‡[1,3]	221	405§[2]	211	$2·26 \times 10^{-3}$
Vitamin B₆ (mg)‖	4	9	5	11	4	11	0·264
Vitamin B₁₂ (µg)‖	19	91	19	71	15	67	0·144
Vitamin C (mg)‖	219[2]	230	270‡[1,3]	325	192§[2]	195	$1·48 \times 10^{-8}$
Ca (mg)‖	1328	656	1261	549	1289	635	1·00
Fe (mg)‖	18	11	17	8	17	8	1·00

[1,2,3] Superscripts indicate which cluster is significantly different using Bonferroni *post hoc* tests for pairwise comparison between clusters.
* P values provided by general linear models controlling for age, sex, BMI and country where appropriate. Bonferroni *post hoc* tests were used to examine pairwise comparisons between groups with the exception of vitamin D, where Fisher's least significant difference *post hoc* tests were used instead.
† General linear models were calculated on log-transformed values where necessary and adjusted for multiple comparisons.
‡ Highest values across the clusters.
§ Lowest values across the clusters.
‖ Adjusted for energy (kJ). Dietary intakes were examined across the three clusters.

Clusters also differed in terms of supplement users ($P = 9·31 \times 10^{-8}$), with the highest percentage found in cluster 2 (54·3 %), in which the individuals also had the highest Omega-3 Index.

Development of the targeted dietary advice

Targeted dietary advice was developed on the basis of the characteristics (anthropometric, metabolic and nutrient intake data) of each cluster using a decision tree method (see Fig. 1). Two decision trees were constructed per cluster: a combined metabolic and anthropometric decision tree and a dietary decision tree. Ranges of the metabolic markers and nutrients were calculated for each of the clusters and these values were then used to determine whether individuals in each cluster were within the desirable or high/low range for that particular variable, as shown in Table 5. The cut-offs used in the current study were based on those used within the Food4Me study (online Supplementary Table S2), but were simplified for the purposes of the development of the targeted dietary advice. For the targeted dietary advice, the cut-offs were set as either 'desirable' or 'high/low'(Table 5), whereas in Food4Me the cut-offs were developed using a more complex gradation scale (online Supplementary Table S2). This information was then used to construct the branches of each of the decision trees per cluster. Using this method, dietary advice was developed based

on four metabolic markers (total cholesterol, total saturated fat, Omega-3 Index and carotenoids) and five key nutrients (salt, dietary fibre, Fe, Ca and folate). The online Supplementary Fig. S1(a) and (b) demonstrate the metabolic and anthropometric decision tree and dietary decision trees for cluster 2, respectively. Furthermore, examples of targeted messages from each of the decision trees for cluster 2 are given in the online Supplementary Fig. 1(a) and (b). Combining the information from the characteristics of each cluster and the messages from the decision trees resulted in a comprehensive message. An example of advice for a participant in cluster 1 is given in the online Supplementary Table S4.

Comparison of the targeted dietary advice and personalised feedback reports

Level 2 participants from Food4Me (*n* 180) were selected to test the appropriateness of the targeted dietary advice developed within this study. Excellent agreement was found between the personalised advice delivered by trained nutritionists in Food4Me and the targeted method developed in this study, with an average match of 82 % in relation to the dietary messages given (Table 6). Examining the clusters individually, good agreement was also found with an average match of 83 % for cluster 1, 74 % for cluster 2 and 88 % for cluster 3 for the dietary messages given. The number of messages given as part

Table 4. Food group intakes across the clusters*
(Mean values and standard deviations)

Food groups (g)	Cluster 1 Mean	SD	Cluster 2 Mean	SD	Cluster 3 Mean	SD	P†
Rice, pasta and grains	81	62	75	58	90	76	0·608
Savouries	24[2]	26	11‡[1,3]	13	28§[2]	33	$1·27 \times 10^{-4}$
White bread/rolls/scones/croissants	55[2]	92	34‡[1,3]	73	71§[2]	120	$1·36 \times 10^{-5}$
Wholemeal and brown bread	102	126	100	122	85	156	1·00
Breakfast cereals and porridge	53	82	76	109	48	66	1·00
Biscuits, cakes and pastries	70	133	68	80	70	88	0·512
Whole milk	45	175	32	93	36	106	1·00
Low-fat and skimmed milks	166	217	164	218	184	225	1·00
Other milks, milk-based beverages and other beverages	196	314	163	221	189	300	1·00
Creams, ice creams and desserts	12	15	9	21	7	10	0·128
Cheese	36	37	37	37	32	35	0·288
Yogurts	79[2]	96	91§[1,3]	107	75‡[2]	128	0·032
Egg and egg dishes	31	38	31	41	32	40	1·00
Butter, fat spreads and hard cooking fats	13	19	9	12	8	11	1·00
Low-fat spreads and oils	10	10	9	9	10	11	1·00
Potatoes	60	67	52	48	51	69	1·00
Chips and processed potatoes	25	27	18	21	25	34	1·00
Vegetables and vegetable dishes	190	145	225	194	146	114	0·480
Fruit juices and smoothies	125	176	126	171	114	158	1·00
Fruit	253[2]	216	355§[1,3]	306	228‡[2]	189	$1·39 \times 10^{-8}$
Savoury snacks	9	13	9	13	10	14	1·00
Fish, fish dishes and products	55‡[2]	40	71§[1,3]	53	66[2]	57	$8·16 \times 10^{-4}$
Red meat	43	37	30	39	41	36	0·704
Poultry	34	35	30	39	35	33	1·00
Meat products	49	51	34	49	45	52	0·832
Red meat dishes	34	58	30	37	33	34	1·00
Alcoholic beverages	211	257	129	180	156	207	0·064
Sugar syrups, preserves and sweeteners	11	14	9	12	10	15	1·00
Confectionery	29	48	21	24	25	30	1·00
Soups, sauces and condiments	100	73	91	72	94	85	1·00
Low-energy beverages	556	530	604	509	434	479	0·128
High-energy beverages	34	73	12	33	44	169	0·064
Supplement users (%)	37·2[2]		54·3[1,3]		37·5[2]		$9·31 \times 10^{-8}$

[1,2,3] Superscripts indicate which cluster is significantly different using Bonferroni *post hoc* tests for pairwise comparison between clusters.
* *P* values provided by general linear models controlling for age, sex, BMI and country where appropriate. Frequency of supplement users was assessed using χ^2 analysis.
† General linear models were calculated on logged values where necessary and adjusted for multiple comparisons.
‡ Lowest values across the clusters.
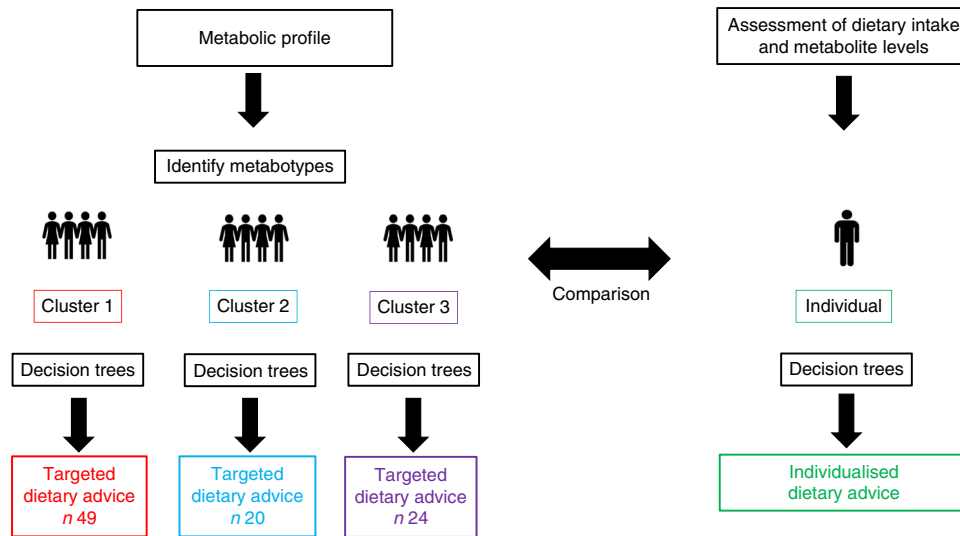§ Highest values across the clusters.

Fig. 1. Development of targeted dietary advice and comparison with individualised advice. Overview of the process for the delivery of targeted advice and comparison with individualised dietary advice using data from the Food4Me Study. Individuals are placed in metabotypes on the basis of their metabolic profiles. In this example, three distinctly different clusters are identified (cluster 1 had high cholesterol, high *trans*-fat and low *n*-3; cluster 2 had high *n*-3 and high total carotenoids; cluster 3 had low cholesterol and high stearic acid). Dietary advice encompasses characteristics of the metabotype and the decision trees, which include dietary factors not captured by the metabolites and anthropometric characteristics. The appropriateness of the targeted dietary advice was then compared with the individualised dietary advice of randomly selected Food4Me participants (*n* 180).

Table 5. Range of values across the clusters and cut-offs used for the development of the targeted dietary advice

	Cluster 1	Cluster 2	Cluster 3		Cut-offs	
Total cholesterol (mmol/l)	3·99–6·03	3·88–5·70	3·47–5·03		Desirable <5	High >5
Total carotenoids (μM)	0·82–1·74	1·437–2·863	0·807–1·613		Desirable >1·5	Low <1·5
Total SFA (%)	High	Low	Medium		NA	
Omega-3 Index (%)	4·23–6·09	5·27–7·85	4·49–6·33		Desirable ≥8	Low <4
Dietary fibre (g)	16·21–43·29	17·09–47·71	13·62–42·92	Males: 18–50 years	Desirable ≥38	Low <38
				>50 years	Desirable ≥30	Low <30
				Females: 18–50 years	Desirable ≥25	Low <25
				>50 years	Desirable ≥21	Low <21
Salt (g)	3·92–11·84	3·94–10·00	3·48–11·44	18–50 years	Desirable ≤3·75	High >3·75
				51–70 years	Desirable ≤3·25	High >3·25
				>70 years	Desirable <3	High >3
Folate (μg)	224·42–623·24	221·88–664·04	193·66–615·80		Desirable ≥320	Low <320
Ca (mg)	671·48–1984·10	712·24–1809·72	654·41–1924·15	Males: 18–70 years	Desirable ≥800	Low <800
				>70 years	Desirable ≥1000	Low <1000
				Females: 18–50 years	Desirable ≥800	Low <800
				>50 years	Desirable ≥1000	Low <1000
Fe (mg)	6·94–29·22	8·83–25·45	9·15–24·75	Males: >18 years	Desirable ≥ 6	Low <6
				Females: 18–50 years	Desirable ≥8·1	Low <8·1
				>50 years	Desirable ≥5	Low <5
BMI (kg/m^2)	22·43–32·93	20·09–27·79	20·66–30·04	Normal: 18·5–24·99	Overweight: ≥25	Obese: ≥30
WC (m)	0·79–1·07	0·71–0·93	0·72–0·98	Males	Desirable <102	High ≥102
				Females	Desirable <88	High ≥88

WC, waist circumference.

Table 6. Agreement between the proposed targeted dietary advice and the individualised dietary advice method adopted within the Food4Me study*

	Agreement between targeted and individualised methods (%)
Cluster 1	83
Cluster 2	74
Cluster 3	88
Total	82

* The agreement between the targeted and individualised method is at the level of the delivery of the same dietary message.

Table 7. Number of messages given as per the targeted dietary advice* (Numbers and percentages)

Number of messages given	No. of times	%
2	13	7
3	46	26
4	50	28
5	51	28
6	20	11

* Number of dietary messages given using the proposed targeted dietary advice method.

of the targeted dietary advice is depicted in Table 7. In general, more messages were given using the targeted approach compared with the individualised method used in Food4Me, where a restriction to three nutrient-related goals was imposed.

Discussion

The present study demonstrates a successful method for the delivery of targeted nutrition advice using a combination of metabotyping and decision trees. Excellent agreement between this method and that of a personalised method delivered by a team of trained nutritionists and dietitians in the Food4Me study was found, with an average match of 82 %, at the level of agreement of the same dietary message given. To the best of our knowledge, this is the first study to identify metabotypes in the European population and to develop tailored dietary solutions appropriate for participants from diverse cultures and dietary intakes. This work paves the way for further development of this approach and potential delivery of personalised nutrition advice to large population groups.

Using cluster analysis, three distinctly different metabotypes were identified based on a range of blood-based metabolic markers. Individuals in cluster 1 were found to have an unhealthy metabolic profile as these individuals had the highest cholesterol levels, highest saturated fat levels and lowest Omega-3 Index. On the other hand, individuals in cluster 2 were identified as the healthiest group and had the lowest saturated fat levels, highest carotenoid concentrations and highest Omega-3 Index. Subjects in cluster 3 were found to be the youngest with the lowest cholesterol levels and poor dietary quality in terms of fruit intake fibre levels and carotenoid levels. These findings are similar to previously published studies on metabotypes[6,17]. Morris et al.[6] identified four metabotypes consisting of four different responses to an oral glucose tolerance test (OGTT). Classification of individuals based on their response curves to an OGTT revealed an 'at-risk' metabolic phenotype, which had the highest BMI, TAG levels, C-reactive protein, C-peptide, insulin and homeostatic model assessment of insulin resistance (HOMA-IR) score[6]. In a similar manner, van Bochove et al.[17] identified three clusters based on their lipoprotein profiles and reported one cluster in which individuals did not respond favourably to fenofibrate treatment. In our previous study, one cluster with a metabolically unfavourable profile and another cluster in which the individuals were relatively healthy with respect to a range of metabolic markers were also identified[22]. The consistency of identification of clusters across a range of studies adds validation to the approach and supports the clusters found in the present study.

An important finding from the current study is the evidence that there was a relationship between the metabolic profiles of each cluster and the corresponding nutrient and food group intakes of those clusters. For example, in line with their high carotenoid concentrations, participants in cluster 2 were also found to have the highest intakes of vitamin C, folate and dietary fibre. Similarly, individuals in cluster 2 had the highest intake of the fish, fish dishes and products, which was also reflected in their metabolic profile as this group had the highest average Omega-3 Index. However, individuals in cluster 2 had the highest intakes of supplements, which were likely to contribute to their high *n*-3 levels. The agreement between the metabolic profiles and dietary intake supports the concept of using blood-based metabotypes as a basis for targeted nutrition advice.

Good agreement between the proposed framework and the individualised advice delivered in the Food4Me study was observed. In Food4Me, personalised dietary advice was delivered by trained dietitians and nutritionists across seven research centres in Europe and was based on a decision tree method, which resulted in 295 possible dietary messages[23]. In the final section of the personalised reports, participants were given three key pieces of dietary advice that they were encouraged to focus on, which were selected based a priority system, developed specially for the purposes of the Food4Me study[23]. In contrast to this, a more simplified method is proposed here, in which blood-based metabolic data in conjunction with minimal dietary information could be used to deliver tailored dietary advice. This more simplified approach showed an average match of 82 % at the level of the dietary advice given, with the actual advice delivered within the Food4Me study. On the basis of this, it is proposed that tailored dietary advice could be given based primarily on the metabolic markers and information on the intakes of five key nutrients.

A framework for the delivery of targeted dietary advice in the Irish population, by the identification of three diverse metabotypes, and development of tailored dietary advice based on decision trees was previously presented[22]. In the current paper, a similar method to identify metabotypes was used, but we have advanced this concept by the inclusion of a broader range of metabolic data. Furthermore, the decision trees for the delivery of the advice were expanded to include specific key nutrients. Furthermore, the decision trees were designed to accommodate sex-specific recommendations. This approach has potential to improve public health through the provision of tailored dietary advice to patients, in a quick and efficient manner, with minimal effort required by healthcare providers.

In this study, the metabolic markers were collected using DBS cards by the participants in their own homes. Collection of samples by DBS has a number of advantages including reduced costs, possibility of collection of large sample sizes, no blood processing and minimal storage facilities required[30,31]. This presents another opportunity for the proposed framework to be adopted in the community setting where community health nurses could deliver the targeted dietary advice. Community nurses are suitable candidates to deliver tailored advice as they routinely see patients who may benefit from dietary/lifestyle change, have regular contact with patients over long periods, visit patients in their own homes and can involve their family in the intervention, and visit those who may not be physically capable of attending their doctor[32]. Chan *et al.* conducted a study to investigate the scope for risk management practices by nurses based in the community[32]. They reported that levels of obesity and prevalence of risk factors including smoking status and low physical activity levels were higher in the individuals (*n* 804) who took part in the study, compared with the general population, and that the majority of individuals with at least one risk factor had not received advice or been referred in the past 3 months[32]. This suggests that there is considerable scope to deliver dietary and lifestyle interventions in the community. In addition, when provided with appropriate training, community nurses were shown to be confident in assessing lifestyle factors such as smoking, anthropometric measures and dietary intake[33]. It is envisaged that the proposed framework, in our study, could easily be adopted by nurses in the community setting, to deliver tailored dietary advice with minimal training required, and have the potential to reach many more individuals who could benefit from tailored dietary advice.

A major strength of this study is its applicability to the European population. Furthermore, good agreement was reported between the proposed targeted method and an individualised method delivered by a team of nutritionists across seven research centres in the Food4Me study. A limitation of this study is that the dietary intake data were collected using an online FFQ, which assessed dietary intake of the previous month. Furthermore, the dietary advice developed did not take into account cooking abilities, likes/dislikes or cost of meals.

The present study developed a framework for the identification of metabotypes in the European population and the development of tailored dietary advice. Good agreement was found in comparison with an individualised personalised nutrition approach, which has been used to deliver advice across seven countries. The demonstration of this approach in a pan European study offers significant credibility to the framework. In our previous study, we envisaged translation of this approach for use by healthcare professionals and the present study further supports such a concept. With this in mind, future work should test this framework in such a setting to ascertain whether the advice is effective in motivating changes in diet and lifestyle factors.

Acknowledgements

Supplementary material

For supplementary material/s referred to in this article, please visit https://doi.org/10.1017/S0007114517002069

References

1. Ronteltap A, van Trijp H, Berezowska A, et al. (2013) Nutrigenomics-based personalised nutritional advice: in search of a business model? Genes Nutr **8**, 153–163.
2. Gibney MJ & Walsh MC (2013) The future direction of personalised nutrition: my diet, my phenotype, my genes. Proc Nutr Soc **72**, 219–225.
3. Brennan L (2008) Session 2: Personalised nutrition. Metabolomic applications in nutritional research. Proc Nutr Soc **67**, 404–408.
4. Kaput J (2008) Nutrigenomics research for personalized nutrition and medicine. Curr Opin Biotechnol **19**, 110–120.
5. O'Donovan CB, Walsh MC, Gibney MJ, et al. (2015) Can metabotyping help deliver the promise of personalised nutrition? Proc Nutr Soc **75**, 106–114.
6. Morris C, O'Grada C, Ryan M, et al. (2013) Identification of differential responses to an oral glucose tolerance test in healthy adults. PLOS ONE **8**, e72890.
7. Erro R, Vitale C, Amboni M, et al. (2013) The heterogeneity of early Parkinson's disease: a cluster analysis on newly diagnosed untreated patients. PLOS ONE **8**, e70244.
8. Richette P, Clerson P, Perissin L, et al. (2013) Revisiting comorbidities in gout: a cluster analysis. Ann Rheum Dis **74**, 142–147.
9. Viniol A, Jegan N, Hirsch O, et al. (2013) Chronic low back pain patient groups in primary care – a cross sectional cluster analysis. BMC Musculoskelet Disord **14**, 294.
10. Castaldi PJ, Dy J, Ross J, et al. (2014) Cluster analysis in the COPDGene study identifies subtypes of smokers with distinct patterns of airway disease and emphysema. Thorax **69**, 415–422.
11. Haldar P, Pavord ID, Shaw DE, et al. (2008) Cluster analysis and clinical asthma phenotypes. Am J Respir Crit Care Med **178**, 218–224.
12. Moore WC, Meyers DA, Wenzel SE, et al. (2010) Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. Am J Respir Crit Care Med 181, pp. 315–323.
13. Newby C, Heaney LG, Menzies-Gow A, et al. (2014) Statistical cluster analysis of the British Thoracic Society Severe refractory Asthma Registry: clinical outcomes and phenotype stability. PLOS ONE **9**, e102987.
14. Park HW, Song WJ, Kim SH, et al. (2015) Classification and implementation of asthma phenotypes in elderly patients. Ann Allergy Asthma Immunol **114**, 18–22.
15. Clayton TA, Lindon JC, Cloarec O, et al. (2006) Pharmaco-metabonomic phenotyping and personalized drug treatment. Nature 440, pp. 1073–1077.
16. Winnike JH, Li Z, Wright FA, et al. (2010) Use of pharmaco-metabonomics for early prediction of acetaminophen-induced hepatotoxicity in humans. Clin Pharmacol Ther **88**, 45–51.
17. 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, pp. e38072.
18. 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. Mol Nutr Food Res **55**, 679–690.
19. Heinzmann SS, Merrifield CA, Rezzi S, et al. (2011) Stability and robustness of human metabolic phenotypes in response to sequential food challenges. J Proteome Res **11**, 643–655.
20. Vazquez-Fresno R, Llorach R, Perera A, et al. (2016) Clinical phenotype clustering in cardiovascular risk patients for the identification of responsive metabotypes after red wine polyphenol intake. J Nutr Biochem **28**, 114–120.
21. Botelho PB, Fioratti CO, Abdalla DS, et al. (2010) Classification of individuals with dyslipidaemia controlled by statins according to plasma biomarkers of oxidative stress using cluster analysis. Br J Nutr **103**, 256–265.
22. O'Donovan CB, Walsh MC, Nugent AP, et al. (2015) Use of metabotyping for the delivery of personalised nutrition. Mol Nutr Food Res **59**, 377–385.
23. 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 Nutr **10**, 450.
24. Celis-Morales C, Livingstone KM, Marsaux CF, et al. (2016) Effect of personalized nutrition on health-related behaviour change: evidence from the Food4me European randomized controlled trial. Int J Epidemiol **46**, 578–588.
25. Forster H, Fallaize R, Gallagher C, et al. (2014) Online dietary intake estimation: the Food4Me food frequency questionnaire. J Med Internet Res **16**, e150.
26. 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. J Med Internet Res **16**, e190.
27. 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 Nutr **10**, 476.
28. Rosal MC, Ebbeling CB, Lofgren I, et al. (2001) Facilitating dietary change: the patient-centered counseling model. J Am Diet Assoc **101**, 332–341.
29. Forster H, Walsh MC, O'Donovan CB, et al. (2016) A dietary feedback system for the delivery of consistent personalized dietary advice in the web-based multicenter Food4Me Study. J Med Internet Res **18**, e150.
30. McDade TW, Williams S & Snodgrass JJ (2007) What a drop can do: dried blood spots as a minimally invasive method for integrating biomarkers into population-based research. Demography **44**, 899–925.
31. McDade TW (2014) Development and validation of assay protocols for use with dried blood spot samples. Am J Hum Biol **26**, 1–9.
32. Chan BC, Laws RA, Williams AM, et al. (2012) Is there scope for community health nurses to address lifestyle risk factors? the community nursing SNAP trial. BMC Nurs **11**, 4.
33. Chan BC, Jayasinghe UW, Christl B, et al. (2013) The impact of a team-based intervention on the lifestyle risk factor management practices of community nurses: outcomes of the community nursing SNAP trial. BMC Health Serv Res **13**, 54.