Public Health Nutrition: 16(5), 848–857 doi:10.1017/S1368980011002473

Dietary patterns in Irish adolescents: a comparison of cluster and principal component analyses

Áine P Hearty* and Michael J Gibney
UCD Institute of Food & Health, Room 3.02, Agriculture & Food Science Centre, University College Dublin, Belfield, Dublin 4, Republic of Ireland

Submitted 1 September 2010: Accepted 8 September 2011: First published online 13 October 2011

Abstract

Objective: Pattern analysis of adolescent diets may provide an important basis for nutritional health promotion. The aims of the present study were to examine and compare dietary patterns in adolescents using cluster analysis and principal component analysis (PCA) and to examine the impact of the format of the dietary variables on the solutions.

Design: Analysis was based on the Irish National Teens Food Survey, in which food intake data were collected using a semi-quantitative 7 d food diary. Thirty-two food groups were created and were expressed as either g/d or percentage contribution to total energy. Dietary patterns were identified using cluster analysis (*k*-means) and PCA.

Setting: Republic of Ireland, 2005–2006.

Subjects: A representative sample of 441 adolescents aged 13–17 years.

Results: Five clusters based on percentage contribution to total energy were identified, 'Healthy', 'Unhealthy', 'Rice/Pasta dishes', 'Sandwich' and 'Breakfast cereal & Main meal-type foods'. Four principal components based on g/d were identified which explained 28 % of total variance: 'Healthy foods', 'Traditional foods', 'Sandwich foods' and 'Unhealthy foods'.

Conclusions: A 'Sandwich' and an 'Unhealthy' pattern are the main dietary patterns in this sample. Patterns derived from either cluster analysis or PCA were comparable, although it appears that cluster analysis also identifies dietary patterns not identified through PCA, such as a 'Breakfast cereal & Main meal-type foods' pattern. Consideration of the format of the dietary variable is important as it can directly impact on the patterns obtained for both cluster analysis and PCA.

Keywords
Adolescents
Dietary patterns
Cluster analysis
Principal component analysis

The diets of adolescents have received much publicity in recent years from surveys which indicate that this population group is consuming diets low in vitamins and minerals and high in fat and sugar(1–3). These findings are of concern because such diets may impact on dental health(4,5), physical development(6), academic achievement(7) and the risk of developing chronic diseases, such as CVD(8–10) and osteoporosis(11). Without effective intervention, poor eating habits developed during adolescence may 'track' into adulthood(6,12,13). Therefore it is important to provide a basis for nutritional health promotion in adolescents by documenting their dietary practices and identifying areas where strategies can be focused.

One possible strategy is to examine dietary patterns using statistical methods such as cluster analysis or factor analysis (specifically, principal component analysis (PCA)). These approaches avoid the well-documented limitations of the single food or nutrient approach(14,15), and take into account the cumulative effect of multiple foods. Most importantly, they provide a realistic measure of dietary exposure for epidemiological research(9,16–19). Both of these methods are statistically quite different. PCA is an exploratory approach used to identify dietary patterns in a population(20,21). New dietary/food pattern variables are derived on the basis of the correlation matrix of the original food variables and individuals receive a factor score for each of the derived factors. In cluster analysis, dietary data are reduced into patterns based on individual differences in dietary intakes. Cluster analysis creates patterns that are mutually exclusive, as each individual can belong to only one cluster(22).

In the literature there are only limited studies that have examined dietary patterns of adolescents, and there are none to our knowledge that have compared cluster analysis and PCA in elucidating these patterns. In relation to adults, these methods have been compared in only a few studies(23–26) and in general it was found that both methods produced comparable dietary patterns. Comparisons of

*Corresponding author: Email Aine.Hearty@ucd.ie

statistical methods used to derive dietary patterns are necessary in order to ensure that the most appropriate method is used to explore and quantify dietary patterns in adolescents. Therefore, the aims of the present study were: (i) to examine dietary patterns in a representative sample of Irish adolescents using both cluster analysis and PCA; (ii) to examine the impact of the format of the dietary variables on the pattern solutions (i.e. expressed as either g/d of each food group or the percentage contribution to total energy intake (%TE) from each food group); and (iii) to directly compare the dietary pattern solutions obtained from both methods.

Methods

Study sample details

The National Teens Food Survey (NTFS)[27] was conducted in the Republic of Ireland, in 2005–2006, on a representative sample of 441 teenagers aged 13–17 years (Table 1). All respondents completed a 7 d semi-weighed food diary, which was used to collect information on all foods and beverages consumed. Respondents were provided with detailed instructions on how to record information on the amount and types of all foods and beverages consumed per eating occasion per day over the 7 d period. Foods were quantified using the following methods: direct weighing of foods, manufacturer's information, photographic food atlas, published food portion sizes and household measures. Food and nutrient analysis was conducted using WISP (Weighed Intake Software Program) version 3·0 (Tinuviel Software, Anglesey, UK). WISP uses data from McCance & Widdowson's *The Composition of Foods*[28] plus supplemental volumes to generate nutrient intake data. Questionnaires were used to obtain further health and lifestyle information and these were completed by the respondents and their parents/guardians.

Food groups

Total food intake was associated with 1761 food composition codes; therefore, to ease interpretation of clusters and factors, these codes were aggregated into thirty-two food groups (online Supplemental Table A). These were established by grouping similar foods together based on their nutrient profiles and on recognised classification systems (e.g. apples and oranges were grouped into a 'fruit' group; peanuts and potato crisps were grouped into a 'savoury snack' group). Only edible fraction weights of all foods were considered (e.g. weight of banana without skin). Foods expressed as the dry weight version were corrected to represent the amount as consumed. Two beverage food groups were created to differentiate between the types of beverages consumed: 'low-calorie' (e.g. water, tea, coffee, sugar-free drinks and sugar-free squashes) and 'high-calorie' (e.g. soft drinks and squashes). Whole milk, low-fat milk and fruit juice remained in separate food groups. For both cluster analysis and PCA, food groups were expressed in both mean g/d intake and as %TE.

Table 1 Demographic characteristics of the sample: Irish adolescents aged 13–17 years, National Teens Food Survey (NTFS), Republic of Ireland, 2005–2006

	All			
	n	%	Boys (*n*)	Girls (*n*)
Total	441	100	224	217
Age range				
13–14 years	188	42·6	95	93
15–17 years	253	57·4	129	124
Season of survey				
Winter (September–February)	187	42·4	80	107
Summer (March–August)	254	57·6	144	110
Geographical location				
Country/village	137	31·1	66	71
Small town	80	18·1	47	33
Large town	62	14·1	34	28
City	162	36·7	77	85
Parents' social class*				
Professional worker	216	50·1	115	101
Non-manual worker	79	18·3	41	38
Skilled manual worker	85	19·7	38	47
Unskilled worker	51	11·8	25	26
Parents' education level*				
Intermediate	85	19·6	38	47
Secondary	169	39·0	83	86
Tertiary	179	41·3	99	80

*Responses do not always add up to 441 as complete demographic information was not obtained from each participant. The teenager was assigned the higher social class and education level category of both parents.

Identification of dietary patterns

Cluster analysis

Cluster analysis was performed using the *k*-means algorithm, which provides a measure of Euclidean distance from each record to the cluster centre and from each cluster to the others[22]. A series of steps was taken to select the most suitable number of clusters. First, several runs on the continuous food group variables were conducted with a varying number of clusters (i.e. two to eight). For each run, cluster proximities for each cluster centre were examined and the number of iterations per cluster was increased to ensure minimum error in cluster membership and that the model had converged to a solution. Clusters were also run without outliers to help find the best cluster solutions. Finally, the resulting clusters, based on food group variables expressed in g/d and %TE, were examined for sensible patterns (refer to online Supplemental Table B for comparisons between the cluster solutions).

Principal component analysis

PCA was used to extract dietary factors on the basis of the correlation of the thirty-two food groups consumed. Determination of which factors to extract should be a multiple-step process; in the present study, scree plots were examined and principal components (PC) with eigenvalues of $\geq$1·5 were retained in order to extract the

https://doi.org/10.1017/S1368980011002473 Published online by Cambridge University Press

main dietary factors. The extracted factors were then orthogonally rotated by the varimax method so that factors were uncorrelated to allow for increased ease of interpretation. Further to this, factor loadings were saved for each PC for each respondent. Each rotated PC was interpreted based on food groups with loadings of $\geq 0\cdot 2$ or $\leq -0\cdot 2$, as these loadings have been previously recognised as making a significant contribution to a dietary pattern[29]. Food groups were also expressed as either g/d or %TE (refer to online Supplemental Table B). This was to allow for a comparison of the dietary factors based on differing input variables.

Statistical analyses

The SPSS® statistical software package version 15 (SPSS Inc., Chicago, IL, USA) was used for data manipulation and basic statistical analysis of the data sets. The SPSS Clementine® statistical software package version 9·0 (SPSS Inc.) was used to conduct the data reduction analysis, and this software standardises all variables prior to analysis. Differences in the mean percentage contribution of each food group across clusters, and in mean nutrient intake across clusters, were evaluated using one-way ANOVA. One-way ANOVA was also used to test for significant differences in mean nutrient intakes across quartiles of PC. Where statistically different effects were encountered ($P < 0\cdot 05$), comparisons of means were made using the Scheffé *post hoc* multiple comparisons test. For values that did not comply with Levene's test for homogeneity of variance, the Tamhane *post hoc* multiple comparison test was used[30]. Mean PC scores were also calculated and computed for each cluster solution.

In order to determine whether the dietary patterns derived by both methods were comparable, binary logistic regression analysis was used. A model was constructed which predicted the odds of being in quartile 4 of each of the four PC (dependent variable) based on membership of one of the five clusters (independent variable). Confounding factors were adjusted for (i.e. sex, age group, social class and smoking), and models for different scenarios based on these confounders were run. Odds ratios and the corresponding 95% confidence intervals were calculated for each model.

Results

Cluster analysis

The five-cluster solution derived based on %TE provided the most sensible and defined clusters, so this solution was chosen for the remainder of analysis presented here. The solution based on g/d variables produced two dietary clusters, which were not as easy to interpret as the solutions produced using the %TE variables (online Supplemental Table B). The values for %TE per food group per cluster are depicted in Table 2, and the five clusters were labelled as 'Healthy', 'Unhealthy', 'Rice/Pasta dishes', 'Sandwich' and 'Breakfast cereal & Main meal-type foods'. Mean daily nutrient intakes were also compared across clusters (Table 3). Dietary patterns were also analysed separately for males and females. For males, a two-cluster solution was revealed, and the clusters were labelled as 'Unhealthy' and 'Healthy'. For females a four-cluster solution was revealed and these were labelled as 'Light meal/Snacks', 'Unhealthy & Meat', 'Healthy & Desserts', 'Healthy & Savouries' (data not shown). A summary of the cluster profiles for the total adolescent sample follows.

Cluster 1: 'Healthy'

This cluster was characterised by providing a relatively high %TE from breakfast cereals, fruit, low-fat spreads, low-fat milk, vegetables, wholemeal bread and yoghurt. This cluster had a high %TE from carbohydrate (relative to Clusters 4 and 5) and from protein (relative to Clusters 2 and 4), and a low %TE from both total and saturated fat (relative to Clusters 2, 4 and 5). It also had the highest fibre of all the clusters, and it had a high Fe intake (relative to Cluster 2), Zn intake (relative to Clusters 2 and 3), folate intake (relative to Clusters 2, 3 and 4) and vitamin C intake (relative to Cluster 2).

Cluster 2: 'Unhealthy'

This cluster was characterised by providing a relatively high %TE from chips, confectionery, high-calorie beverages, meat products and savoury snacks. This cluster was associated with having the highest intake of added sugar. It was also found to have low Fe and folate intakes (relative to Clusters 1 and 5) and low vitamin C intake (relative to Cluster 1).

Cluster 3: 'Rice/Pasta dishes'

This was the least prevalent cluster (10·0%) and was characterised by providing a relatively high %TE from eggs, poultry dishes, rice & pasta, and savouries. This cluster was associated with having a low energy intake (relative to Clusters 1, 2 and 5). It also had low fibre and Ca intakes (relative to Clusters 1, 4 and 5) and a low Zn intake (relative to Clusters 1 and 5).

Cluster 4: 'Sandwich'

This was the most prevalent cluster (27·2%) and it was characterised by providing a relatively high %TE from butter & spreads, cheese, red meat, sauces, sugar & preserves, and white bread. This cluster was associated with having a low %TE from carbohydrate (relative to Clusters 1, 2 and 3) and a high %TE from total fat (relative to Clusters 1, 3 and 5) and from saturated fat (relative to Clusters 1 and 3). It also had a low added sugar intake (relative to Cluster 2).

Cluster 5: 'Breakfast cereal & Main meal-type foods'

This cluster was characterised by providing a relatively high %TE from breakfast cereal, desserts, potatoes, poultry and whole milk. This cluster was associated with having a

https://doi.org/10.1017/S1368980011002473 Published online by Cambridge University Press

Table 2 The dietary profile of the five clusters observed, as described by the mean percentage contribution of each food group variable to total energy intake: Irish adolescents aged 13–17 years, National Teens Food Survey (NTFS), Republic of Ireland, 2005–2006

	Total (*n* 441)		Cluster 1: 'Healthy' (*n* 63; 14·3 %)		Cluster 2: 'Unhealthy' (*n* 112; 25·4 %)		Cluster 3: 'Rice/Pasta dishes' (*n* 44; 10 %)		Cluster 4: 'Sandwich' (*n* 120; 27·2 %)		Cluster 5: 'Breakfast cereal & Main meal-type foods' (*n* 102; 23·1 %)	
Food group	Mean	SD	Mean*	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Breakfast cereal	6·3	5·5	**9·5**[a]	5·8	3·4[b]	3·6	5·3[b]	4·9	4·0[b]	3·3	**10·4**[a]	5·7
Biscuits & cakes	4·8	5·2	6·2[a,b]	6·7	3·6[a]	3·9	3·6[a]	4·5	6·2[b]	5·5	4·3[a,b]	4·7
Butter & spreads	3·0	3·4	2·3[a]	2·8	2·1[a]	2·5	2·6[a]	3·3	**4·7**[b]	4·2	2·5[a]	2·8
Cheese	2·2	2·9	2·5[a,c]	2·6	1·3[b]	2·2	1·5[a,b]	2·4	**3·8**[c]	3·8	1·3[b]	1·8
Chips	7·4	5·7	4·7[a]	4·7	**11·7**[b]	6·6	8·2[c]	5·3	6·5[a,c]	4·4	5·3[a]	3·7
Confectionery	6·8	5·6	4·7[a]	4·2	**8·9**[b]	6·6	6·6[a,b]	6·2	6·6[a,b]	5·8	5·9[a]	4·1
Desserts	2·4	3·3	2·0[NS]	2·3	2·3	3·4	2·3	2·9	2·5	3·5	**2·9**	3·6
Eggs	0·8	1·5	0·8[NS]	1·6	0·8	1·7	**1·2**	1·7	0·8	1·3	0·6	1·1
Fish	0·9	1·8	1·2[NS]	1·5	0·5	1·2	1·1	3·1	1·0	1·8	0·9	1·6
Fruit	1·5	2·4	**4·2**[a]	4·5	0·6[b]	0·9	1·0[b]	1·3	1·7[c]	1·6	0·9[b]	1·3
Fruit juice†	1·8	2·4	1·8[NS]	2·2	1·5	2·3	2·1	2·6	2·1	2·8	1·6	2·2
High-calorie beverages	4·5	4·1	2·6[a]	2·7	**8·3**[b]	4·3	5·9[c]	4·7	2·7[a]	2·3	2·9[a]	2·6
Low-fat spreads	0·4	1·0	**0·8**[NS]	1·6	0·3	0·8	0·3	0·7	0·4	1·1	0·3	0·8
Low-calorie beverages	0·1	0·3	0·1[NS]	0·1	0·1	0·3	0·1	0·1	0·1	0·3	0·1	0·3
Low-fat milk	1·0	2·5	**4·5**[a]	4·7	0·3[b,c]	1·1	0·7[b,c]	1·7	0·6[b]	1·5	0·1[c]	0·4
Meat products	6·8	5·5	3·8[a]	3·0	**9·7**[b]	6·6	7·1[b,c]	5·6	6·6[c]	5·0	5·5[c]	4·4
Potatoes	3·0	2·6	2·8[NS]	2·2	2·9	2·7	2·5	2·5	3·0	2·9	**3·3**	2·6
Poultry dishes	1·3	2·3	0·9[a]	1·5	0·7[a]	1·3	**6·8**[b]	2·2	0·7[a]	1·2	0·7[a]	1·2
Poultry	2·1	2·3	2·7[a]	2·2	1·9[a,b]	2·3	1·5[b]	1·8	1·4[b]	1·8	**2·8**[a]	2·8
Red meat	3·6	3·3	3·0[NS]	2·2	3·8	3·7	3·3	3·1	**4·1**	3·8	3·1	2·7
Red meat dishes	2·5	3·2	3·0[NS]	2·5	2·1	2·8	1·9	2·8	2·5	2·9	3·0	4·2
Rice & pasta	2·4	3·3	2·8[a]	2·3	1·3[b]	2·0	**4·1**[a]	5·1	2·6[a]	3·0	2·4[a,b]	3·9
Sauces	1·8	2·2	1·9[a,b]	2·0	1·4[a]	1·7	1·5[a,b]	2·1	**2·5**[b]	3·0	1·4[a]	1·6
Savouries	5·1	5·1	4·6[NS]	4·7	5·5	5·2	**5·9**	5·3	3·9	4·1	5·9	6·0
Savoury snacks	3·7	4·0	2·3[a,c]	2·5	**6·6**[b]	4·8	4·0[a]	4·3	3·2[a]	3·3	1·9[c]	2·5
Soups	0·5	1·0	0·5[NS]	1·0	0·3	0·7	0·6	1·3	0·7	1·2	0·4	1·0
Sugar & preserves	1·4	2·4	1·4[NS]	1·9	1·1	2·2	0·6	0·8	**1·6**	2·4	1·6	3·2
Vegetables	2·0	2·2	**2·8**[a]	2·3	1·4[b,c]	1·5	1·7[a,b,c]	2·5	2·4[a,c]	2·8	1·7[c]	1·5
White bread	9·7	5·5	7·5[a]	4·8	8·8[a]	4·6	8·3[a]	5·8	**12·7**[b]	5·1	9·0[a]	5·8
Wholemeal bread	2·2	3·6	**6·9**[a]	5·8	0·8[b]	1·8	1·8[b,c]	2·4	1·8[c]	2·6	1·4[b,c]	2·4
Whole milk	6·9	6·5	3·0[a]	4·1	4·9[b]	4·2	4·6[a,b]	4·4	5·1[b]	3·8	**14·5**[c]	7·0
Yoghurt	1·0	1·7	**2·0**[a]	2·2	0·5[b]	1·0	0·8[b,c]	1·5	1·2[a,c]	1·7	0·9[b,c]	1·8

[a,b,c]Mean values within a row (across clusters) with unlike superscript letters were significantly different using one-way ANOVA ($P < 0·05$).
*Values in bold signify the main food groups contributing to the cluster (percentage of total energy).
†Fruit juice here refers to intakes of 100 % fruit juice only, excluding sweetened fruit juice drinks which are included in the beverage food groups.

high energy intake (relative to Clusters 2 and 3), a high %TE from saturated fat (relative to Clusters 1 and 3) and a high intake of Ca (relative to Clusters 2, 3 and 4).

Principal component analysis

For the PCA, it was decided to use the solution determined from food groups in the g/d format as this produced components that explained greater variance than when they were based on %TE. The eigenvalues revealed four major dietary patterns, which explained 28·0 % of the total variance. These were labelled as 'Healthy foods', 'Traditional foods', 'Sandwich foods' and 'Unhealthy foods'. The analysis was repeated separately for boys and girls. Five PC were extracted for boys, which together explained 33·9 % of total variance, and were labelled as 'Healthy foods', 'Traditional foods', 'Snack foods', 'White bread' and 'Rice/Pasta dishes'. Four PC were extracted for girls which explained 30·0 % of total variance, and were labelled as 'Healthy foods', 'Unhealthy foods', 'Traditional foods' and 'Snack foods' (Table 4). For each PC, quartiles of the total component weights were calculated and compared across nutrient intakes, and the highest quartile (Q4) for each is presented in Table 5. The findings for each PC follow.

PC 1: 'Healthy foods'

This was characterised by positive loadings for wholemeal bread, fruit, fish, breakfast cereal and vegetables. In relation to Q4, this PC had a low %TE from total fat (relative to PC 3 and 4) and from saturated fat (relative to PC 3). It also had high fibre intake (relative to PC 3 and 4), high folate intake (relative to PC 4), low added sugar intake (relative to PC 4) and overall it had the highest vitamin C intake ($P < 0·05$).

PC 2: 'Traditional foods'

This was characterised by having positive loadings for potatoes, red meat and vegetables. In relation to Q4, this PC had a high %TE from protein (relative to PC 3 and 4).

https://doi.org/10.1017/S1368980011002473 Published online by Cambridge University Press

Table 3 Comparison of selected macro- and micronutrient intakes across dietary clusters: Irish adolescents aged 13–17 years, National Teens Food Survey (NTFS), Republic of Ireland, 2005–2006

	Total		Cluster 1: 'Healthy'		Cluster 2: 'Unhealthy'		Cluster 3: 'Rice/Pasta dishes'		Cluster 4: 'Sandwich'		Cluster 5: 'Breakfast cereal & Main meal-type foods'	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Energy (kJ)	8287	2433	8734^{a,c}	3112	7935^{a}	1955	7071^{b}	2392	8229^{a,b,c}	2403	8988^{c}	2236
Energy (kcal)	1980·6	581·6	2087·4^{a,c}	743·9	1896·8^{a}	467·2	1690·1^{b}	571·8	1966·7^{a,b,c}	574·4	2148·3^{c}	534·5
Protein (%TE)	14·8	2·6	16·2^{a}	2·4	13·3^{b}	2·2	14·8^{a,c}	2·3	14·5^{c}	2·4	15·7^{a}	2·5
Carbohydrate (%TE)	49·0	5·1	51·4^{a}	5·4	49·6^{a,c}	4·7	50·0^{a,c}	5·7	47·2^{b}	4·7	48·6^{b,c}	4·8
Total fat (%TE)	35·6	5·0	31·9^{a}	5·1	36·4^{b,c}	4·4	34·6^{a,b}	5·4	37·8^{c}	4·6	35·0^{b}	4·5
Saturated fat (%TE)	14·0	2·7	12·0^{a}	2·9	13·9^{b,c}	2·2	12·8^{a,b}	2·3	14·8^{c}	2·5	14·8^{c}	2·7
Added sugar (g)	65·7	31·6	59·5^{a}	31·4	80·4^{b}	34·8	59·2^{a}	29·3	57·1^{a}	26·0	66·2^{a}	29·6
Englyst fibre (g)	11·6	5·2	16·2^{a}	8·9	10·0^{b}	2·8	9·4^{b}	3·8	11·4^{c}	3·9	11·8^{c}	4·1
Ca (mg)	906·4	406·4	1092·2^{a}	458·7	699·6^{b}	274·9	670·9^{b}	288·0	878·2^{c}	351·0	1153·4^{a}	413·8
Fe (mg)	12·4	11·7	16·4^{a}	16·5	9·5^{b}	5·6	11·9^{a,b}	15·7	11·5^{a,b}	10·4	14·5^{a}	11·5
Zn (mg)	8·7	3·7	10·1^{a}	4·1	7·6^{b}	3·3	6·9^{b}	3·6	8·7^{a,b}	3·4	9·8^{a}	3·7
Folate (μg)	275·8	150·5	374·7^{a}	177·6	212·2^{b}	110·3	227·9^{b}	156·8	257·3^{b}	137·7	326·8^{a}	135·9
Vitamin C (mg)	95·3	93·9	146·5^{a}	164·4	76·3^{b}	61·9	87·1^{a,b}	88·8	92·7^{a,b}	71·0	91·1^{a,b}	79·0

%TE, percentage of total energy intake.
^{a,b,c}Mean values within a row (across clusters) with unlike superscript letters were significantly different using one-way ANOVA ($P<0·05$).

PC 3: 'Sandwich foods'

This was characterised by having positive loadings for white bread, sugar, butter and cheese. In relation to Q4, this PC had a high %TE from total and saturated fat (relative to PC 1).

PC 4: 'Unhealthy foods'

This was characterised by having positive loadings for meat products, chips, sauces and high-calorie beverages. In relation to Q4, this PC had a low %TE from protein (relative to PC 1 and 2), low intakes of fibre, folate and vitamin C (relative to PC 1) and a high intake of added sugar (relative to PC 1).

Comparison of cluster solutions and principal components

Mean PC scores were calculated and computed across the five cluster solutions (Fig. 1). The most striking features of this are in relation to Clusters 1, 2 and 4. Cluster 1 ('Healthy') scored highest for PC 1 ('Healthy foods') and lowest for PC 4 ('Unhealthy foods'), indicating that close similarities exist between the dietary pattern 'Healthy' derived by both methods and that it is most different from the 'Unhealthy' PC. Cluster 2 ('Unhealthy') scored highest for PC 4 ('Unhealthy foods'), also indicating that both cluster analysis and PCA derived very similar patterns. Cluster 4 ('Sandwich') scored highest for PC 3 ('Sandwich foods'), which also illustrates close similarities. Cluster 3 ('Rice/Pasta dishes') was most different from PC 2 ('Traditional foods'), which is expected as rice and pasta are not typically associated with a traditional Irish diet. Cluster 5 ('Breakfast cereal & Main meal-type foods') scored highest for PC 2 ('Traditional foods'), which is expected due to the prevalence of main meal-type foods in both dietary patterns. However, this cluster also included breakfast cereals, which is probably why it is most different from PC 4 ('Unhealthy foods'), indicating that this cluster represents a dietary pattern not fully identified using PCA alone.

In order to compare cluster solutions with PC derived from the same data set, logistic regression was performed. The logistic regression model was adjusted for sex, age group, social class and smoking status, which were the variables shown to have most impact on the model fit (Table 6). The logistic regression model attempts to predict the odds of being in Q4 of each PC, compared with the other quartiles, from the five clusters with Cluster 1 ('Healthy') as the reference category. Compared with the reference category, all of the other clusters had a lower odds of predicting PC 1 ('Healthy foods'), with Cluster 2 ('Unhealthy') having the lowest odds (OR = 0·08). In relation to Q4 of PC 2 ('Traditional foods'), none of the clusters were found to be significant predictors of it (i.e. $P>0·05$). In relation to PC 3 ('Sandwich foods'), Cluster 4 ('Sandwich') had the highest odds compared with the reference category for predicting this (i.e. OR = 3·97). Finally, in relation to PC 4 ('Unhealthy foods'), Cluster 2

https://doi.org/10.1017/S1368980011002473 Published online by Cambridge University Press

Table 4 Factor loadings* for each food group per retained PC, for the total sample and split for boys and girls: Irish adolescents aged 13–17 years, National Teens Food Survey (NTFS), Republic of Ireland, 2005–2006

	Total sample†				Boys‡					Girls§			
Food group	PC 1	PC 2	PC 3	PC 4	PC 1	PC 2	PC 3	PC 4	PC 5	PC 1	PC 2	PC 3	PC 4
Wholemeal bread	0·78	–	–	–	0·73	–	–	–	–	0·71	–	–	–
Fruit	0·74	–	–	–	0·76	–	–	–	–	0·64	–	–	0·42
Breakfast cereal	0·58	0·26	–	–	0·66	–	–	–	–	0·64	–	–	–
Fish	0·55	–	–	–	0·58	–	–	–	–	–	–	–	–
Vegetables	0·39	0·39	–	–	0·48	0·44	–	–	–	–	−0·33	–	0·57
Low-calorie beverages	0·25	–	0·23	–	0·35	–	0·20	0·42	–	0·26	–	0·23	0·42
Sugar	0·26	–	0·40	–	–	–		0·61	–	0·34	–	0·36	–
Potatoes	–	0·80	–	–	0·31	0·72	–	–	–	–	–	0·79	–
Red meat	–	0·74	–	–	–	0·82	–	–	–	–	–	0·64	–
Rice & pasta	–	−0·23	–	–	–	−0·22	–	–	0·77	–	−0·22	−0·23	–
Savouries	–	−0·23	–	0·28	–	–	–	–	–	–	0·22	–	–
Cheese	–	−0·22	0·30	–	–	–	0·48	–	–	0·21	–	−0·31	–
Poultry dishes	–	−0·22	–	–	–	–	–	–	–		–	–	–
Whole milk	–	0·20	–	–	–	–	–	–	–	0·25	–	0·24	–
Butter & spreads	–	–	0·78	–	–	0·27	–	0·66	–	–	–	–	–
White bread	–	–	0·77	–	–	–	–	0·73	–	–	–	–	–
Meat products	–	–	–	0·64	–	–	–	–	–	−0·24	0.25	–	–
Chips	–	–	–	0·64	–	–	–	–	−0·25	–	0·39	−0·21	–
Sauces	–	–	–	0·54	–	–	–	–	–	–	–	–	–
High-calorie beverages	–	–	–	0·51	–	–	–	−0·22	–	–	0·33	–	–
Eggs	–	–	–	–	–	0·36		–	–	–	–	–	–
Yoghurts	–	–	–	–	–	–	0·74	–	–	–	–	–	0·74
Biscuits & cake	–	–	–	–	–	–	0·75	–	–	0·23	–	0·30	–
Low-fat milk	–	–	–	–	–	–	0·21	–	–	0·20	–	–	–
Low-fat spreads	–	–	–	–	–	–	0·23	–	−0·20	–	–	–	–
Red meat dishes	–	–	–	–	–	–	–	–	0·84	–	–	–	–
Poultry	–	–	–	–	–	–	–	–	–	0·39	–	–	–
Confectionery	–	–	–	–	–	–	–	–	–	–	0·75	–	–
Savoury snacks	–	–	–	–	–	–	–	–	–	–	0·74	–	–
Fruit juice	–	–	–	–	–	–	–	–	–	–	–	−0·26	–
Desserts	–	–	–	–	–	–	–	–	–	–	–	–	0·55
% Variance	10·36	6·91	5·64	5·12	10·38	6·91	6·08	5·47	5·03	11·63	6·83	5·79	5·75

PC, principal component (extraction method: principal component analysis, rotation method: varimax with Kaiser normalisation).
*Factor loadings are only displayed for values ≤−0·2 or ≥0·2; some food groups excluded as they did not load onto any factor retained.
†Total sample: PC 1 = 'Healthy foods'; PC 2 = 'Traditional foods'; PC 3 = 'Sandwich foods'; PC 4 = 'Unhealthy foods'.
‡Boys: PC 1 = 'Healthy foods'; PC 2 = 'Traditional foods'; PC 3 = 'Snack foods'; PC 4 = 'White bread'; PC 5 = 'Rice/Pasta dishes'.
§Girls: PC 1 = 'Healthy foods'; PC 2 = 'Unhealthy foods'; PC 3 = 'Traditional foods'; PC 4 = 'Snack foods'.

('Unhealthy') had the highest odds of all clusters for predicting this PC compared with the reference category (OR = 7·47).

Discussion

Food choice is a complex process because it is deeply embedded in culture and is influenced by many factors internal and external to the person[31,32]. Food choice in adolescence is particularly important because nutrient needs are higher at this stage than at any other time in the life cycle and food intakes at this stage may influence future health[11]. In the literature, there are not as many studies of dietary patterns in adolescents as there are for adults, and the present study is the first one that the authors are aware of in Irish adolescents. The present study explored dietary patterns in Irish adolescents using cluster analysis and PCA methods. A 'Sandwich' pattern was the most prevalent cluster (27·2 %), indicating the importance of light meals and sandwiches to the diets of adolescents. This 'Sandwich' pattern was identified by both cluster analysis and PCA and had the highest percentage contributions to total and saturated fat intakes of all of the patterns. Unlike many studies on dietary patterns in adults which have often noted a 'Traditional' dietary pattern[22], in these adolescents a 'Traditional' pattern was revealed by PCA but not by cluster analysis. Cluster 5 was a modified version of the 'Traditional' pattern, containing main-meal foods along with breakfast cereal. This indicates that PCA and cluster analysis have identified slightly different dietary patterns. These differences identified by the two methods are most evident with respect to Cluster 3, 'Rice/Pasta dishes', as it appears to not have been identified through PCA and is very different from the PC 'Traditional foods' and 'Sandwich foods'.

Both methods produced a 'Healthy' and an 'Unhealthy' dietary pattern. The 'Healthy' pattern as identified by cluster analysis was present in 14·0 % of adolescents, while the 'Unhealthy' pattern was present in 25·0 %. The 'Unhealthy' pattern was associated with having high intake of added sugar, and low intakes of fibre, Fe, folate

https://doi.org/10.1017/S1368980011002473 Published online by Cambridge University Press

Table 5 Comparison of micro- and macronutrient intakes across Q4 of four PC of dietary patterns (*n* 110): Irish adolescents aged 13–17 years, National Teens Food Survey (NTFS), Republic of Ireland, 2005–2006

	Q4 of PC 1: 'Healthy foods'		Q4 of PC 2: 'Traditional foods'		Q4 of PC 3: 'Sandwich foods'		Q4 of PC 4: 'Unhealthy foods'	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Energy (kJ)	9380^{NS}	2416	9543	2301	9632	2442	9620	2008
Energy (kcal)	2241·9^{NS}	577·4	2280·8	549·9	2302·0	583·7	2299·2	479·9
Carbohydrate (%TE)	49·8^{NS}	5·3	48·3	5·8	48·0	5·4	48·0	4·7
Protein (%TE)	15·5^{a}	2·4	15·8^{a}	2·3	14·5^{b}	2·3	14·3^{b}	2·3
Total fat (%TE)	34·1^{a}	5·3	35·5^{a,b}	5·5	37·2^{b}	5·3	36·9^{b}	4·8
Saturated fat (%TE)	13·3^{a}	3·0	14·2^{a,b}	3·1	14·7^{b}	3·2	13·9^{a,b}	2·3
Englyst fibre (g)	15·6^{a}	7·1	13·9^{a,b}	6·3	13·3^{b}	6·0	12·5^{b}	4·6
Added sugar (g)	66·5^{a}	30·5	71·6^{a,b}	33·6	72·4^{a,b}	32·6	80·1^{b}	31·2
Ca (mg)	1063·8^{NS}	407·4	1053·3	443·0	1064·0	396·1	964·8	407·8
Fe (mg)	14·9^{NS}	12·1	14·5	11·0	15·0	15·4	13·5	11·2
Zn (mg)	10·3^{NS}	3·9	10·8	3·4	10·1	4·0	9·6	3·5
Folate (μg)	363·7^{a}	179·3	326·3^{a,b}	146·5	325·8^{a,b}	174·2	299·2^{b}	159·0
Vitamin C (mg)	130·9^{a}	102·5	98·4^{b}	73·7	94·7^{b}	93·2	91·3^{b}	62·3

Q4, quartile 4; PC, principal component; %TE, percentage of total energy intake.
^{a,b}Mean values within a row (across PC) with unlike superscript letters were significantly different using one-way ANOVA ($P<0·05$).

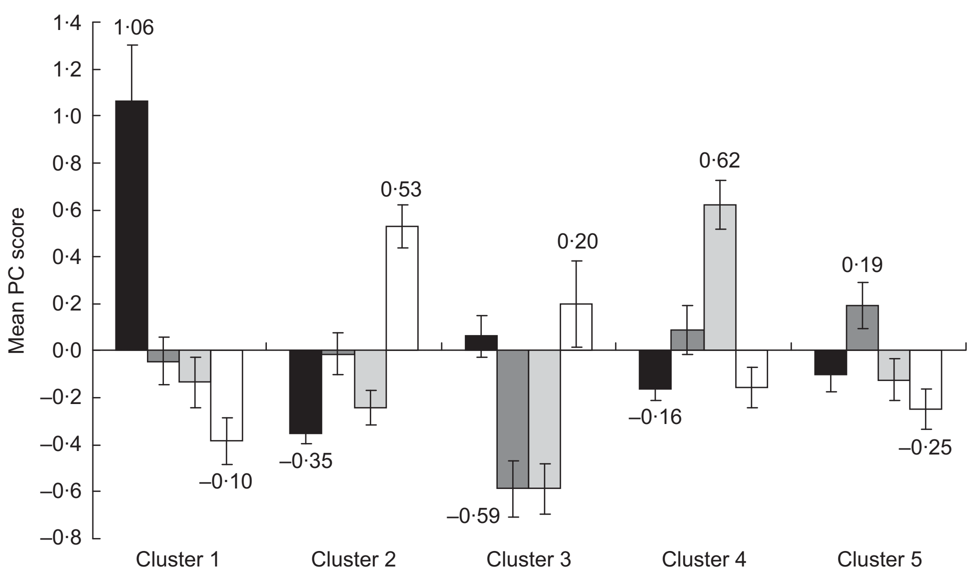

Fig. 1 Mean principal component (PC) score (■, PC 1 = 'Healthy foods'; ▩, PC 2 = 'Traditional foods'; ▨, PC 3 = 'Sandwich foods'; □, PC 4 = 'Unhealthy foods') compared across each of the five clusters of dietary patterns (Cluster 1 = 'Healthy', Cluster 2 = 'Unhealthy', Cluster 3 = 'Rice/Pasta dishes', Cluster 4 = 'Sandwich', Cluster 5 = 'Breakfast cereal & Main meal-type foods'). Values are means with their standard deviations represented by vertical bars; the values of the highest and lowest PC mean scores for each cluster are also indicated. Irish adolescents aged 13–17 years (*n* 441), National Teens Food Survey (NTFS), Republic of Ireland, 2005–2006

and vitamin C, and it was 7·5 times more likely to predict the 'Unhealthy' PC than the 'Healthy' cluster. Anderson *et al.*[33] explored dietary intake in 15-year-old adolescents in Scotland and found that less than one-third of the sample consumed a diet similar to that promoted by health education. A study by Lambert *et al.*[5] described the choice of beverages and desserts made by 7–16-year-old boys in a school cafeteria in the UK. They found that buns and cookies were ten times more popular than fresh fruits and yoghurts. Sugary soft drinks were twenty times more popular than fresh fruit drinks and milk. They concluded that, despite the availability of 'healthy' options, the boys showed a clear preference for products high in fat and/or sugar.

The most popular a posteriori methods of data reduction are cluster analysis and factor analysis. There is no 'gold standard' technique for evaluating dietary patterns, so therefore the decision should be made according to the expertise of the research group and the format of the output required. For example, clusters can be more easily profiled on selected attributes (such as sociodemographic characteristics) than can patterns derived by PCA[34].

https://doi.org/10.1017/S1368980011002473 Published online by Cambridge University Press

Table 6 Comparison of dietary patterns derived from cluster analysis and principal component analysis using binary logistic regression: Irish adolescents aged 13–17 years, National Teens Food Survey (NTFS), Republic of Ireland, 2005–2006

	Q4 of PC 1: 'Healthy foods'			Q4 of PC 2: 'Traditional foods'			Q4 of PC 3: 'Sandwich foods'			Q4 of PC 4: 'Unhealthy foods'		
	P	OR	95 % CI	*P*	OR	95 % CI	*P*	OR	95 % CI	*P*	OR	95 % CI
Adjusted model*												
Cluster 1: 'Healthy' (reference)	0·000			0·041			0·000			0·000		
Cluster 2: 'Unhealthy'	0·000	0·08	0·03, 0·20	0·326	1·51	0·66, 3·46	0·831	0·91	0·38, 2·17	0·000	7·47	3·05, 18·32
Cluster 3: 'Rice/Pasta dishes'	0·057	0·44	0·19, 1·02	0·130	0·35	0·09, 1·36	0·283	0·51	0·15, 1·74	0·170	2·11	0·73, 6·15
Cluster 4: 'Sandwich'	0·000	0·21	0·11, 0·43	0·052	2·13	0·99, 4·58	0·000	3·97	1·85, 8·51	0·236	1·71	0·70, 4·17
Cluster 5: 'Breakfast cereal & Main meal-type foods'	0·000	0·19	0·09, 0·40	0·288	1·53	0·70, 3·35	0·699	1·17	0·52, 2·65	0·881	0·93	0·36, 2·38

Q4, quartile 4; PC, principal component.
*Model adjusted for sex, age group, social class and smoking.

Factors identified also do not refer to identifiable groups within the population and hence do not give an indication of the prevalence of a particular type of diet[35]. On the other hand, the actual procedure for PCA is more straightforward and logical as cluster analysis places considerable burden on the user in terms of selecting and deciding on the appropriate number of clusters. Nevertheless, caution needs to be applied to the subjective criteria involved when using PCA (such as selection based on the eigenvalue), as this can have a direct impact on the number and type of dietary patterns revealed in the data. The four PC retained for further examination in the present study account for only 28·0 % of the total variance, therefore the possibility remains that other dietary patterns exist. Schulze *et al.*[36] investigated the effect of PCA on dietary patterns in a sample of the Potsdam cohort of the European Prospective Investigation into Cancer and Nutrition. They concluded that PCA may explain food and nutrient intake quite differently, and in some cases PCA may not uncover all dietary patterns in the data set. Another issue which should always be considered is the format of the food group variable. In the present study, when the food group variables were expressed as %TE, the most sensible cluster solutions were found, as only two cluster solutions were obtained with the g/d variable and these were difficult to interpret as distinct dietary patterns. However, for PCA, when the food group variables were expressed as g/d, more sensible PC were produced. When the variables were expressed as %TE, five PC were produced; some of these had only a few foods with high factor scores and therefore were difficult to adequately define. It should be borne in mind that the formats of the dietary variable used may have contributed to the slightly different patterns identified by either method. However, the focus of the present paper was first to identify the most appropriate format of the variable to use for either method separately and then to compare the resulting patterns.

As with any dietary survey where food intake is self-reported, misreporting could have been an issue with this adolescent sample. However, in order to obtain a valid measure of food intake strict protocols were adhered to during the data collection phase. These included adequate training for the respondent in completing the food diary, three or four visits by the fieldworker during the data entry phase, and face-to-face evaluation of the food diary at the end of the 7 d period.

There exist many studies on factors influencing food choices of adolescents[31,32,37,38]. Three key influences in adolescence have been described: (i) individual influences, such as snacking and dieting[35,37,39]; (ii) social environmental influences, such as family and peers[32,38]; and (iii) physical environmental influences on accessibility and availability of foods[31]. Eating family meals may result in adolescents having more healthy diets. A recent study on Australian adolescents aged 14 years found that a 'healthy' dietary pattern was positively associated with better family functioning, independent of family income and maternal education[38].

https://doi.org/10.1017/S1368980011002473 Published online by Cambridge University Press

In the available literature on dietary patterns in adolescents, it appears that unhealthy and snack food/convenience-type patterns predominate in this population subgroup, with poor compliance with national guidelines[6,33,37,40,41]. These studies corroborate the current findings. In general, there have been many reports of the relatively high prevalence of snacking in the adolescent population[6,8,40]. In the USA, snacking prevalence increased from 77 % to 84 % between 1977 and 1996[42] and snacks were found to contribute 20–23 % of total daily energy and 15–19 % of total daily fat intakes. A study on American adolescents aged 14–19 years found that less than half of them achieved the US food pyramid dietary recommendations[37]. Positive intakes (related to dietary diversity scores) were associated with the adolescent obtaining more of his/her meals and snacks from home. Among males, pyramid scores were positively predicted by snacking behaviour, indicating that snacks can make valuable contributions to adolescent diets, and may help them achieve the high requirements for energy and other nutrients. However, increased snacking also led to increased sugar scores among both genders. Martens *et al.*[41] also found a high prevalence of high-fat snack intake in a sample of Dutch adolescents aged 12–14 years, and found that adolescents' attitudes were the most important determinant of health-related eating behaviour. Another barrier in the promotion of healthy food choices in adolescents may be advertising. The less healthy products are often heavily branded and strongly promoted throughout society, e.g. chocolate and confectionery products[5,31].

A few studies have examined whether dietary patterns formed in adolescence 'track' into adulthood. Mikkila *et al.*[12] examined dietary patterns from the Young Finns Study, which was a prospective cohort of children and adolescents aged 3–18 years at baseline, with a 21-year follow-up. They found that the two patterns identified at baseline were clearly identified at all time points over the 21-year period. The small differences in the food groups associated with each pattern did not alter the overall food behaviour style, but rather they reflected the change in food variety and culture during the time of the study. They found that tracking was stronger among those who were adolescents at baseline than among those who were younger children. Ritchie *et al.*[6] examined a 10-year cohort of dietary patterns in African-American and Caucasian female adolescents. No African-American girls and only 12 % of Caucasian girls had diets classified as 'healthy'. They did not find that dietary patterns fluctuated above what would be expected over the 10 years. Li and Wang[13] also tracked changes in dietary intake patterns in adolescents over a 1-year period. They found moderate tracking in the 'Western' diet ($r = 0{\cdot}5$), but weaker tracking in two healthier dietary patterns ($r = 0{\cdot}31$–$0{\cdot}36$). Tracking was affected by gender, body weight status and physical activity levels. More recently, a study of the stability of dietary patterns in US adolescents revealed that patterns were relatively stable over 5 years, and that the dietary patterns identified ('vegetable', 'fruit', 'sweet/salty snack food' and 'starchy food') differed from those usually found in adults[43]. These results suggest that food behaviour and food choices are established early in life and show long-term stability. Therefore, nutrition education should be targeted at the adolescent population group as it could have a vital role in the prevention of many diet-related diseases in adulthood.

In conclusion, it appears that a 'Sandwich' and an 'Unhealthy' pattern are the main dietary patterns in this sample of Irish adolescents, and this corroborates the findings from other studies of dietary patterns of this population subgroup. For future studies, it is important that the format of the dietary variable is considered, as the present study found that it can directly impact on the patterns obtained for both cluster analysis and PCA. Patterns derived from either method were comparable, although it appears that cluster analysis also identifies dietary patterns not identified through PCA, such as a 'Breakfast cereal & Main meal-type foods' pattern.

Acknowledgements

The authors would like to acknowledge the Irish Department of Agriculture, Fisheries and Food for providing funding for further analysis of the National Teens Food Survey (NTFS). There are no conflicts of interest for this research. All authors have provided their permission and all those who contributed significantly to this work are listed. A.P.H. carried out the analysis for the paper and drafted the manuscript. M.J.G. provided guidance for the analysis and reviewed the final drafts of the manuscript.

References

1. Cavadini C, Siega-Riz AM & Popkin BM (2000) US adolescent food intake trends from 1965 to 1996. *Arch Dis Child* **83**, 18–24.
2. Thane CW, Bates CJ & Prentice A (2003) Risk factors for low iron intake and poor iron status in a national sample of British young people aged 4–18 years. *Public Health Nutr* **6**, 485–496.
3. Gibson S & Neate D (2007) Sugar intake, soft drink consumption and body weight among British children: further analysis of National Diet and Nutrition Survey data with adjustment for under-reporting and physical activity. *Int J Food Sci Nutr* **58**, 445–460.
4. Campain AC, Morgan MV, Evans RW *et al.* (2003) Sugar–starch combinations in food and the relationship to dental caries in low-risk adolescents. *Eur J Oral Sci* **111**, 316–325.
5. Lambert N, Plumb J, Looise B *et al.* (2005) Using smart card technology to monitor the eating habits of children in a school cafeteria: 3. The nutritional significance of beverage and dessert choices. *J Hum Nutr Diet* **18**, 271–279.
6. Ritchie LD, Spector P, Stevens MJ *et al.* (2007) Dietary patterns in adolescence are related to adiposity in young adulthood in black and white females. *J Nutr* **137**, 399–406.

https://doi.org/10.1017/S1368980011002473 Published online by Cambridge University Press

7. Fu ML, Cheng L, Tu SH *et al.* (2007) Association between unhealthful eating patterns and unfavorable overall school performance in children. *J Am Diet Assoc* **107**, 1935–1943.
8. Nicklas TA, Webber LS, Thompson B *et al.* (1989) A multivariate model for assessing eating patterns and their relationship to cardiovascular risk factors: the Bogalusa Heart Study. *Am J Clin Nutr* **49**, 1320–1327.
9. Rasanen M, Lehtinen JC, Niinikoski H *et al.* (2002) Dietary patterns and nutrient intakes of 7-year-old children taking part in an atherosclerosis prevention project in Finland. *J Am Diet Assoc* **102**, 518–524.
10. Kim JA, Kim SM, Lee JS *et al.* (2007) Dietary patterns and the metabolic syndrome in Korean adolescents: 2001 Korean National Health and Nutrition Survey. *Diabetes Care* **30**, 1904–1905.
11. Lytle LA (2002) Nutritional issues for adolescents. *J Am Diet Assoc* **102**, 3 Suppl., S8–S12.
12. Mikkila V, Rasanen L, Raitakari OT *et al.* (2005) Consistent dietary patterns identified from childhood to adulthood: the cardiovascular risk in Young Finns Study. *Br J Nutr* **93**, 923–931.
13. Li J & Wang Y (2008) Tracking of dietary intake patterns is associated with baseline characteristics of urban low-income African-American adolescents. *J Nutr* **138**, 94–100.
14. Williams DE, Prevost AT, Whichelow MJ *et al.* (2000) A cross-sectional study of dietary patterns with glucose intolerance and other features of the metabolic syndrome. *Br J Nutr* **83**, 257–266.
15. Park SY, Murphy SP, Wilkens LR *et al.* (2005) Dietary patterns using the Food Guide Pyramid groups are associated with sociodemographic and lifestyle factors: the multiethnic cohort study. *J Nutr* **135**, 843–849.
16. Greenwood DC, Cade JE, Draper A *et al.* (2000) Seven unique food consumption patterns identified among women in the UK Women's Cohort Study. *Eur J Clin Nutr* **54**, 314–320.
17. Maskarinec G, Novotny R & Tasaki K (2000) Dietary patterns are associated with body mass index in multi-ethnic women. *J Nutr* **130**, 3068–3072.
18. Sichieri R (2002) Dietary patterns and their associations with obesity in the Brazilian city of Rio de Janeiro. *Obes Res* **10**, 42–48.
19. Kim MK, Sasaki S, Sasazuki S *et al.* (2004) Prospective study of three major dietary patterns and risk of gastric cancer in Japan. *Int J Cancer* **110**, 435–442.
20. van Dam RM (2005) New approaches to the study of dietary patterns. *Br J Nutr* **93**, 573–574.
21. Martinez ME, Marshall JR & Sechrest L (1998) Invited commentary: factor analysis and the search for objectivity. *Am J Epidemiol* **148**, 17–19.
22. Moeller SM, Reedy J, Millen AE *et al.* (2007) Dietary patterns: challenges and opportunities in dietary patterns research an Experimental Biology workshop, April 1, 2006. *J Am Diet Assoc* **107**, 1233–1239.
23. Kant AK, Graubard BI & Schatzkin A (2004) Dietary patterns predict mortality in a national cohort: the National Health Interview Surveys, 1987 and 1992. *J Nutr* **134**, 1793–1799.
24. Newby PK, Muller D & Tucker KL (2004) Associations of empirically derived eating patterns with plasma lipid biomarkers: a comparison of factor and cluster analysis methods. *Am J Clin Nutr* **80**, 759–767.
25. Bamia C, Orfanos P, Ferrari P *et al.* (2005) Dietary patterns among older Europeans: the EPIC-Elderly study. *Br J Nutr* **94**, 100–113.
26. Crozier SR, Robinson SM, Borland SE *et al.* (2006) Dietary patterns in the Southampton Women's Survey. *Eur J Clin Nutr* **60**, 1391–1399.
27. Irish Universities Nutrition Alliance (date unknown) National Irish dietary surveys: methodology and results. http://www.iuna.net/?p=29 (accessed July 2011).
28. Food Standards Agency (2002) *McCance and Widdowson's The Composition of Foods*, 6th summary ed. Cambridge: Royal Society of Chemistry.
29. Togo P, Heitmann BL, Sorensen TI *et al.* (2003) Consistency of food intake factors by different dietary assessment methods and population groups. *Br J Nutr* **90**, 667–678.
30. Coakes SJ & Steed LG (1999) *SPSS without Anguish Versions 7·0, 7·5, 8·0 for Windows*. Brisbane: John Wiley and Sons.
31. Story M, Neumark-Sztainer D & French S (2002) Individual and environmental influences on adolescent eating behaviors. *J Am Diet Assoc* **102**, 3 Suppl., S40–S51.
32. Contento IR, Williams SS, Michela JL *et al.* (2006) Understanding the food choice process of adolescents in the context of family and friends. *J Adolesc Health* **38**, 575–582.
33. Anderson AS, Macintyre S & West P (1994) Dietary patterns among adolescents in the west of Scotland. *Br J Nutr* **71**, 111–122.
34. Pryer JA, Cook A & Shetty P (2001) Identification of groups who report similar patterns of diet among a representative national sample of British adults aged 65 years of age or more. *Public Health Nutr* **4**, 787–795.
35. Kant AK (2004) Dietary patterns and health outcomes. *J Am Diet Assoc* **104**, 615–635.
36. Schulze MB, Hoffmann K, Kroke A *et al.* (2001) Dietary patterns and their association with food and nutrient intake in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam study. *Br J Nutr* **85**, 363–373.
37. Cusatis DC & Shannon BM (1996) Influences on adolescent eating behavior. *J Adolesc Health* **18**, 27–34.
38. Ambrosini GL, Oddy WH, Robinson M *et al.* (2009) Adolescent dietary patterns are associated with lifestyle and family psycho-social factors. *Public Health Nutr* **12**, 1807–1815.
39. Dugdale AE, Townsend B & Rigsby B (1988) The eating and snacking patterns of young people in Brisbane, Australia. *J Hum Nutr Diet* **1**, 95–104.
40. Aranceta J, Perez-Rodrigo C, Ribas L *et al.* (2003) Socio-demographic and lifestyle determinants of food patterns in Spanish children and adolescents: the enKid study. *Eur J Clin Nutr* **57**, Suppl. 1, S40–S44.
41. Martens MK, van Assema P & Brug J (2005) Why do adolescents eat what they eat? Personal and social environmental predictors of fruit, snack and breakfast consumption among 12–14-year-old Dutch students. *Public Health Nutr* **8**, 1258–1265.
42. Zizza C, Siega-Riz AM & Popkin BM (2001) Significant increase in young adults' snacking between 1977–1978 and 1994–1996 represents a cause for concern. *Prev Med* **32**, 303–310.
43. Cutler GJ, Flood A, Hannan P *et al.* (2009) Major patterns of dietary intake in adolescents and their stability over time. *J Nutr* **139**, 323–328.