Published in final edited form as:
Eur J Clin Nutr. 2008 July ; 62(7): 931–938. doi:10.1038/sj.ejcn.1602789.

Adjusting for energy intake in dietary pattern investigations using principal components analysis

K Northstone[1], **AR Ness**[2], **PM Emmett**[1], and **IS Rogers**[1]
[1]Department of Social Medicine, University of Bristol, Bristol, UK

[2]Department of Oral and Dental Science, University of Bristol, Bristol, UK

Abstract

Objective: The effect of energy adjustment on variables entered into principal component analysis (PCA) to derive dietary patterns has received little attention.

This study determines the effect of adjusting for energy on dietary patterns resulting from PCA and the subsequent effect on future outcomes.

Design and methods: As part of regular self-completion questionnaires, used in the Avon Longitudinal Study of Parents and Children, pregnant women were asked to record the frequency of consumption of a variety of food items. A total of 12 053 women completed the questionnaire. Individual dietary types were identified using PCA, before and after adjusting the food variables for energy intake. Associations with estimated nutrient intakes and with birthweight were examined for the two solutions and when energy adjustment was performed at a later stage of the analysis.

Results: Slight differences were seen in terms of the components extracted and the factor loadings obtained. The associations with nutrient intakes showed that there was a general reduction in the size of the correlation coefficients for the energy-adjusted components compared to the unadjusted components. There did not appear to be any difference in the size of the effects of the dietary pattern scores on birthweight, whether energy was adjusted for before entry into the PCA or after.

Conclusions: In this sample, it is not necessary to adjust for energy intake before entry into a PCA analysis to determine dietary patterns when using food frequency questionnaire data. Effects of energy intake can be determined at a later stage in the analytical process.

Keywords

dietary patterns; pregnancy; principal components analysis; ALSPAC; nutrient intake; energy adjustment

---

Correspondence: K Northstone, Department of Social Medicine, University of Bristol, 24 Tyndall Avenue, Bristol BS8 1TQ, UK. E-mail: kate.northstone@bristol.ac.uk.
Guarantor: K Northstone.
Contributors: KN was responsible for the data analysis and interpretation and drafted the manuscript. PME and IR were responsible for the design, collection and availability of the dietary data and revision of the manuscript. All authors contributed to the intellectual content and approved the final version of the manuscript.

Introduction

There has been a substantial increase over recent years in the use of dietary patterns to assess diet–disease associations. These are used as an alternative to studying the intake of individual food items, food groups or a nutrient specific approach. There are various approaches to obtaining dietary patterns, the most common being Principal components analysis (PCA). However, there is debate surrounding the lack of generalization of dietary patterns obtained using PCA across studies due to differences in methodology and the patterns observed (Martinez et al., 1998; Hu, 2002). The methods used to obtain dietary patterns need to become more uniform among researchers if PCA is to remain a valuable tool in nutritional epidemiology. One particular analytical decision, which has received little attention in the literature, is whether to adjust dietary intake variables for energy before entry into the PCA.

In most studies of diet and disease, the primary exposure of interest is relative (adjusted for energy) rather than absolute dietary intake. Owing to the high inter-correlation of dietary intake with energy, energy adjustment in dietary investigations reduces the variation in dietary intake resulting from differences in ‘body size, metabolic efficiency and physical activity’ (Willett, 1989). The most common measure of diet used to obtain dietary patterns via PCA is a food frequency questionnaire (FFQ), detailing the regularity of food consumption, as opposed to measuring actual intake. FFQs cannot measure energy intake accurately. Nevertheless, any association found between a disease outcome and a dietary pattern obtained from PCA that represents a diet high in energy-dense foods may not be a real effect of the foods themselves, rather an association with actual energy intake. It is therefore important to determine whether researchers need to enter energy-adjusted foods into PCAs or can simply adjust for energy when looking at any dietary pattern – disease association.

A small number of studies, using PCA based on FFQ data, have adjusted dietary variables for energy before entry into the PCA but they provide little or no justification for this choice and no suggestions as to what effect such an adjustment may have (Balder et al., 2003; Costacou et al., 2003; van Dam et al., 2003; Bamia et al., 2005; Velie et al., 2005; Martinez-Ortiz et al., 2006; Waijers et al., 2006). Only one study to date appears to have reported any comparison between unadjusted and adjusted data and this only considered differences in the extracted components, but did not investigate any subsequent impact on potential diet–outcome associations (Bamia et al., 2005). It is important to assess any differences in dietary patterns obtained using PCA based on unadjusted or energy-adjusted data. There is potential for different solutions to be obtained owing to the changes in correlations between food groups after adjusting for energy intake.

The purpose of this study is to examine the effect of energy adjustment on the dietary patterns extracted using PCA on data collected via FFQ. We have chosen to adjust for energy using the residual method as advised by Willett (1989), which is the most widely used method to obtain relative intakes and it therefore seems reasonable to apply this approach. This study also aims to determine any differences in the relationships with estimated nutrient intakes and finally, using birthweight as an example, to evaluate the effects of energy adjustment before PCA on a relevant outcome.

Subjects and methods

The Avon Longitudinal Study of Parents and Children (ALSPAC) is an ongoing population-based study designed to investigate the effects of environmental, genetic and other influences on the health and development of children (Golding et al., 2001). Pregnant women with an expected delivery date between 1 April 1991 and 31 December 1992 and who were resident in the former Avon Health Authority in southwest England were eligible for the study. A total cohort of 14 541 pregnancies resulted. Ethical approval for the study was obtained from the ALSPAC Law and Ethics committee and the local research ethics committees. The representative nature of the ALSPAC sample has been investigated by comparison with the 1991 National Census data of mothers with infants under 1 year of age who were resident in the county of Avon. The ALSPAC sample had a slightly greater proportion of mothers who were married or cohabiting, who were owner-occupiers and who had a car in the household. The study had a smaller proportion of ethnic minority mothers.

More detailed information on the ALSPAC study is available on the website http://www.alspac.bris.ac.uk.

The present study was based on a self-completion questionnaire completed at 32 weeks gestation. This contained a set of questions enquiring about the frequency of consumption of a wide variety of foods and drinks that have been shown to produce mean nutrient intakes (Rogers et al., 1998) similar to those obtained for women in the British National Diet and Nutritional survey for adults (Gregory et al., 1990).

For most food types, the women were given the following options to indicate how often they were currently consuming each food type: (i) never or rarely; (ii) once in 2 weeks; (iii) 1–3 times a week; (iv) 4–7 times a week; (v) more than once a day. Women were also asked to record how many cups of tea or coffee, the number of glasses of cola and the number of slices of bread they usually consumed daily. The usual type of bread (white or other) they used was also recorded.

The data were numerically transformed into times consumed per week as follows: (i) 0; (ii) 0.5; (iii) 2; (iv) 5.5 and (v) 10 times per week, to apply quantitative meaning to the frequency categories. Additionally, all data were standardized by subtracting the mean and dividing by the standard deviation for each variable, since tea, coffee, cola and bread were measured on a scale different from that of the other variables.

Weekly energy intake and other nutrients (selected on the basis of their potential significance during pregnancy) were estimated from the FFQ described above using the fifth edition of McCance and Widdowson's ‘The composition of Food’ and supplements (The Royal Society of Chemistry and MAFF, 1988, 1989, 1991a, b, 1992a, b, 1993) based on standard portion sizes. More detailed information on the methodology is published elsewhere (Rogers et al., 1998). Birthweight (g), as recorded in the delivery room, was abstracted from birth notifications.

Statistical methods

PCA with varimax rotation was performed on the 44 standardized food items. PCA reduces the data by forming linear combinations of the original observed variables, thereby grouping together correlated variables, which in turn identifies any underlying dimensions in the data. The coefficients defining these linear combinations are called ‘factor loadings’ and are the correlations of each food item with that component. The number of components that best represented the data was chosen on the basis of the scree plot (Cattell, 1966) and the interpretability of the factor loadings. Varimax rotation (Gorsuch, 1974; Kline, 1994) was applied. Rotation redistributes the explained variance for the individual components, thereby achieving a simpler structure, increasing the number of larger and smaller loadings.

Foods with loadings above 0.3 on a component were considered to have a strong association with that component and were deemed to be the most informative in describing the dietary patterns. Women were excluded from the PCA if they had more than 10 dietary items missing. If 10 or fewer items were missing, the assumption was made that the woman never consumed the item and it was given a value of 0. This ensured that no missing values were entered into the PCA, a method that cannot deal with missing data.

Two PCAs were performed. The first analysis used the unadjusted dietary variables indicating times consumed per week. The second analysis adjusted these weekly frequency of consumption variables for energy intake using the residual method (Willett, 1989). To test the validity of these solutions, each PCA was repeated in two randomly selected split-half samples. The results were highly comparable for the unadjusted and energy-adjusted data, both in terms of the magnitude of the factor loadings and the component scores obtained; the original solutions were therefore retained.

A component score was created for each woman for each of the components identified, for each method by multiplying the factor loadings by the corresponding standardized value for each food and summing across the food items. Pearson's correlation coefficients were calculated as a measure of the association between the unadjusted and energy-adjusted dietary patterns.

Correlation coefficients were calculated between each of the dietary scores and absolute intake of selected nutrients. Additionally, partial correlation coefficients were obtained after adjusting the nutrients for energy intake. The proportions of variance explained by the dietary pattern scores were calculated for absolute and energy-adjusted nutrient intake and for individual foods, by summing the squared correlations for each nutrient (Bland, 2000).

To determine whether the obtained dietary patterns differed in respect to their ability to predict an outcome, we used crude birthweight as a simple example. Univariable models were obtained by regressing birthweight on each of the dietary pattern scores individually, considered as unadjusted scores and those scores obtained from energy-adjusted data. Bivariable models including energy intake were then examined for the original unadjusted data. We chose not to adjust for other covariates known to be associated with birthweight, to clearly demonstrate any potential effects of energy adjustment.

All analyses were performed using SPSS for windows v.12.0.1.

Results

A total of 12 436 women returned the questionnaire completed at 32 weeks gestation (85.5% of the original sample, many of these had already been lost due to miscarriage); of these, 12 053 (96.9%) had sufficient dietary data available for the PCA.

Table 1 shows the factor loadings obtained from the unadjusted and energy-adjusted PCAs. Five dietary components were chosen to best describe the dietary patterns of the women, explaining 32.4% of the variance using the unadjusted data. For this unadjusted analysis, the first component had high loadings on salad, fruit, rice, pasta, breakfast cereals, fish, eggs, pulses, fruit juices, poultry and non-white bread and we have chosen to label ‘health conscious’. The high loading foods on the second component included green vegetables and root vegetables, potatoes, peas and to some extent red meat and poultry and we labelled it ‘traditional’, based on the familiar British ‘Meat and two veg’ diet. The third component loaded highly on high-fat processed foods, such as meat pies, sausages and burgers, fried foods, pizza, chips and crisps and we therefore labelled it ‘processed’. The fourth component, which we labelled ‘confectionery’, was characterized by high intakes of confectionery and other foods with high sugar content such as chocolate, sweets, biscuits, cakes and other puddings. Finally, the fifth component loaded highly on meat substitutes, pulses, nuts and herbal tea and high negative loadings were seen with red meat and poultry and was therefore labelled ‘vegetarian’.

Four components best described the energy-adjusted data (Table 1), with 26.9% of the variance being explained. Compared to the unadjusted data, the ‘processed’ component was lost. The other components remained in the same order, although there were slight differences in the size of some of the factor loadings. In particular, the food items that loaded highly on the ‘processed’ component using the unadjusted data had increased negative loadings on the ‘health-conscious’ component using the adjusted data. There were strong correlations between the analogous dietary patterns obtained using the two sets of data (Table 2), all >0.8. In addition, the ‘processed’ component obtained from the unadjusted data was negatively correlated with both the ‘health-conscious’ and ‘confectionery’ components obtained using the energy-adjusted data (−0.538 and −0.492, respectively).

Table 3 shows the correlations between the unadjusted and energy-adjusted dietary patterns and absolute intake of selected nutrients and partial correlations for the nutrients after adjusting for energy intake. Of particular interest is the lack of correlation between energy intake and the ‘Traditional’ and ‘Vegetarian’ dietary patterns. The majority of correlations are substantially smaller for the energy-adjusted dietary patterns compared to the unadjusted patterns, in particular there are large decreases in the correlations between fat and sugar intake with the ‘confectionery’ energy-adjusted component.

Parameter estimates for the regression of dietary pattern scores on birthweight are shown in Table 4. The first and second columns show the univariable results for the original unadjusted patterns and for the patterns adjusted for energy. The final column presents the parameter estimates for the unadjusted dietary pattern scores when energy intake is also included in the model. However, the effect sizes for the unadjusted ‘health conscious’ and ‘traditional’ patterns both increased after adjustment for energy intake and their respective confidence intervals were wider (column 1 vs column 3). The effect of the ‘health conscious’ and ‘traditional’ patterns extracted using the energy-adjusted data was larger than that extracted using the unadjusted data, although the width of the confidence intervals remained the same (column 1 vs column 2). There was very little difference in the results for the ‘processed’ and ‘vegetarian’ patterns, although the effect size for the ‘confectionery’ pattern was smaller when energy was adjusted for, either before the extraction of the principal components or including energy in the model with the unadjusted patterns.

Discussion

This study presents the dietary patterns obtained using PCA where food variables were treated as unadjusted or adjusted for energy intake using the residual method before entry in the PCA. There were differences observed in terms of the components extracted and the factor loadings obtained. Most notably, the ‘processed’ component in the unadjusted analysis was lost in the adjusted analysis; the food items that loaded highly on the original ‘processed’ component had increased negative loadings on the ‘health-conscious’ component in the adjusted analysis. The loss of the ‘processed’ component can be partly explained by the attenuation of the correlations between the food items after adjustment for energy. This process removes some of the variability in the data. The dietary patterns common to both sets of analyses that were obtained from the two solutions were highly correlated with each other.

Balder et al. (2003) examined the effect of different analytic decisions on the stability of dietary patterns, including adjusting for energy. They noted that it was those factors, obtained using unadjusted factors, having high loadings on energy-contributing foods were the ones that changed. They explain that as 'substitution' such that the replacement of white with brown bread and high-fat with low-fat diary products became increasingly important. This is a consequence of the requirement that adjusting for energy results in the food groups not being correlated with energy. Comparable results were seen in this study using the energy-adjusted data. In particular, white bread, which was highly positively loaded on the 'processed' component in the unadjusted solution, became highly negatively associated with the 'health-conscious' component in the adjusted solution. Likewise, all the other foods that loaded highly on the 'processed' component, which are high in fat, became negatively highly loaded on the 'health-conscious' component after adjustment for energy. Overall, Balder et al. (2003) reported that the patterns obtained using unadjusted food variables collected via FFQ were comparable to those using energy-adjusted data. They concluded that the dietary patterns were robust to energy adjustment and argued that dietary patterns are based on relative consumption of foods rather than actual intake. Similarly, Bamia et al. (2005) argued that the data they entered into their PCA were based on dietary choices as opposed to overall quantities of intakes.

This is the first study to present the two solutions and to take the analyses further by looking at associations with other variables. The associations with nutrient intakes showed that there was a general reduction in the size of the correlation coefficients for the energy-adjusted components compared to the unadjusted components. This is not unexpected, as adjusting for energy has already reduced the amount of variation. Only selected nutrients were used to illustrate the data a more detailed analysis of nutrient associations will be described elsewhere.

There did not appear to be any difference in the size of the effects of the dietary pattern scores on birthweight, whether energy was adjusted for before entry into the PCA or after. For example, the greatest effect was seen with the 'health-conscious' component. Adjusting for energy intake before the PCA resulted in a birthweight increase of 35 g per 1 s.d. increase in the 'health-conscious' score. This compared to a 40 g increase if the unadjusted score was used but energy intake was adjusted for in the regression model. In epidemiological studies, it is important to be able to clearly determine the effects the different variables have. With this in mind, we would recommend that energy is adjusted for at as late a stage as possible in the analytical process. In this way, the exact impact of energy intake on the outcome of interest is as transparent as it can be and any real effects of energy can be clearly determined. In addition, some outcomes of interest may not be directly affected by energy intake. Birthweight has been used as an illustration to examine the effects of energy adjustment, as with the nutrient analysis, more detailed investigations are currently underway and will be described elsewhere.

In this study, dietary intake was assessed using an unquantified FFQ, with no portion size information included. As such, the derived energy and nutrient information may be inaccurate, compared to the 'gold standard' method of collecting dietary data – weighed intakes. However, one major advantage of this study is the large sample size; weighed intake in such a sample would have been expensive and a greater burden for the participants, potentially resulting in a biased sample. The nutrient intakes in this sample have been shown to compare favourably with the intakes reported by the last Dietary and Nutritional Survey of British adults for all women aged 16–64 (Rogers et al., 1998). Several studies have compared the results of PCA using FFQs with those using weighed dietary records (Hu et al., 1999; Togo et al., 2003; Khani et al., 2004; McNaughton et al., 2005). They all concluded that the resulting factor loadings and dietary pattern scores were comparable. It has been contended that adjusting for energy intake estimated from an FFQ is inadequate (Kipnis et al., 2003; Jakes et al., 2004). FFQs can neither precisely measure energy intake nor can they precisely measure absolute intake. However, it has been argued that relative intakes perform better than absolute intakes using this method of dietary data collection, due to the fact that the errors in measuring energy and nutrients are strongly correlated and therefore cancel each other out (Willett, 2001).

It could be argued that this sample of pregnant women is not the ideal setting for this investigation. Pregnancy is a particularly vulnerable time in the life cycle and adequate nutrition is vital. However, the dietary patterns we report are similar to those we have reported in the resulting children at 4 and 7 years of age (Northstone et al., 2005) and to other UK adult studies (Barker et al., 1990; Gregory et al., 1990).

In conclusion, we found that there were differences in the dietary pattern solutions obtained using unadjusted or energy-adjusted data. However, these differences did not appear to have any major impact on the associations with nutrient intakes or birthweight. In this sample, we do not feel it necessary to adjust for energy intake when performing PCA analysis using FFQ data. We believe that it is sufficient to make adjustment when analysing the effects of the dietary patterns on the outcome of interest, although it is still important to present both unadjusted and adjusted results. Any effects of energy intake can be more precisely determined at a later stage of the analytical process. However, we recommend that researchers still measure the correlations between their dietary pattern scores and estimated energy intake and report these in their papers, to assess the likely impact of energy intake. The correlations between energy intake and the dietary patterns in this study showed that the ‘traditional’ and ‘vegetarian’ components appeared to be robust to energy intake. It would therefore be valuable for other studies to present these associations. It is important to replicate these analyses using data collected via other means, such as dietary diaries to determine whether the conclusions would be the same. It is also important to assess the effect of energy adjustment when using other methods to obtain dietary patterns, such as Reduced Rank Regression (RRR) (Hoffman et al., 2004). We chose to examine the effects of energy adjustment when performing PCA, as this is currently the most popular method of obtaining dietary patterns. Methods such as RRR that offer advantages over PCA when examining health outcomes can also be affected by the type of data that are used as response variables.

Acknowledgments

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council, the Wellcome Trust and the University of Bristol provided core support for ALSPAC. This work was also partially funded by a Wellcome Trust VIP award to KN and by the Arthritic Association supporting KN and PME. This publication is a work by us. We all contributed to the writing of the article. We have no conflict of interest.

References

Balder HF, Virtanen M, Brants HAM, Krogh V, Dixon LB, Tan F, et al. Common and country-specific dietary patterns in four European cohort studies. J Nutr. 2003; 133:4246–4251. [PubMed: 14652380]

Bamia C, Orfanos P, Ferrari P, Overvad K, Hundborg HH, Tjonneland A, et al. Dietary patterns among older Europeans: the EPIC-elderly study. Br J Nutr. 2005; 94:100–113. [PubMed: 16115339]

Barker ME, McClean SI, Thompson KA, Reid NG. Dietary behaviours and sociocultural demographics in Northern Ireland. Br J Nutr. 1990; 64:319–329. [PubMed: 2223737]

Bland, M. An Introduction to Medical Statistics. 3rd edn.. Oxford: Oxford University Press; 2000.

Cattell RB. The screen test for the number of factors. Multivariate Behav Res. 1966; 1:245–276.

Costacou T, Bamia C, Ferrari P, Riboli E, Trichopoulos D, Trichopoulou A. Tracing the Mediterranean diet through principal components and cluster analysis in the Greek population. Eur J Clin Nutr. 2003; 57:1378–1385. [PubMed: 14576750]

Golding J, Pembrey M, Jones R, ALSPAC Study Team. ALSPAC – The Avon Longitudinal Study of Parents and Children. Paediatr Perinat Epidemiol. 2001; 15:74–87. [PubMed: 11237119]

Gorsuch, RL. Factor Analyses. Philadelphia, PA: WB Saunders; 1974.

Gregory, J.; Foster, K.; Tyler, H.; Wiseman, M. The Dietary and Nutritional Survey of British Adults. Office of Population Censuses and Surveys, Chapter 13: Classification of Types of Diet. London: HMSO; 1990.

Hoffman K, S M, Schienkiewtz A, Nothlings U, Boeing H. Application of a new statistical method to derive dietary patterns in nutritional epidemiology. Am J Epidemiol. 2004; 159:935–944. [PubMed: 15128605]

Hu FB, Rimm E, Smith-Warner SA, Feskanich MJ, Stampfer MJ, Ascherio A, et al. Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. Am J Clin Nutr. 1999; 69:243–249. [PubMed: 9989687]

Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol. 2002; 13:3–9. [PubMed: 11790957]

Jakes RW, Day NE, Luben R, Welch A, Bingham S, Mitchell J, et al. Adjusting for energy intake – what measure to use in nutritional epidemiological studies? Int J Epidemiol. 2004; 33:1382–1386. [PubMed: 15333618]

Khani BR, Ye W, Terry P, Wolk A. Reproducibility and validity of major dietary patterns among Swedish women assessed with a food-frequency questionnaire. J Nutr. 2004; 134:1541–1545. [PubMed: 15173426]

Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, et al. Structure of dietary measurement energy: results of the OPEN Biomarker Study. Am J Epidemiol. 2003; 158:14–21. [PubMed: 12835281]

Kline, P. An Easy Guide to Factor Analysis. London: Routledge; 1994.

Martinez ME, Marshall JR, Sechrest L. Invited commentary: factor analysis and the search for objectivity. Am J Epidemiol. 1998; 148:17–19. [PubMed: 9663398]

Martinez-Ortiz JA, Fung TT, Baylin A, Hu FB, Campos H. Dietary patterns and risk of nonfatal acute myocardial infarction in Costa Rican adults. Eur J Clin Nutr. 2006; 60:770–777. [PubMed: 16465200]

McNaughton SA, Mishra GD, Bramwell G, Paul AA, Wadsworth MEJ. Comparability of dietary patterns assessed by multiple dietary assessment methods: results from the 1946 British Birth Cohort. Eur J Clin Nutr. 2005; 59:341–352. [PubMed: 15523484]

Northstone K, Emmett P, the ALSPAC Study Team. Multivariate analysis of diet in children at four and seven years of age and associations with socio-demographic characteristics. Eur J Clin Nutr. 2005; 59:751–760. [PubMed: 15841093]

Rogers I, Emmett P, the ALSPAC Study Team. Diet during pregnancy in a population of pregnant women in South West England. Eur J Clin Nutr. 1998; 52:246–250. [PubMed: 9578336]

The Royal Society of Chemistry and MAFF. Cereals and Oat Products. London: HMSO; 1988.

The Royal Society of Chemistry and MAFF. Milk Products and Eggs. London: HMSO; 1989.

The Royal Society of Chemistry and MAFF. The Composition of Foods. 5th edn.. London: HMSO; 1991a.

The Royal Society of Chemistry and MAFF. Vegetables, Herbs and Spices. London: HMSO; 1991b.

The Royal Society of Chemistry and MAFF. Fruits and Nuts. London: HMSO; 1992a.

The Royal Society of Chemistry and MAFF. Vegetable Dishes. London: HMSO; 1992b.

The Royal Society of Chemistry and MAFF. Fish and Fish Products. London: HMSO; 1993.

Togo P, Heitmann BL, Sorensen TIA, Osler M. Consistency of food intake factors by different dietary assessment methods and population groups. Br J Nutr. 2003; 90:667–678. [PubMed: 13129474]

van Dam RM, Grievink L, Ocke MC, Feskens EJM. Patterns of food consumption and risk factors for cardiovascular disease in the general Dutch population. Am J Clin Nutr. 2003; 77:1156–1163. [PubMed: 12716666]

Velie EM, Schairer C, Flood A, He JP, Khattree R, Schatzkin A. Empirically derived dietary patterns and risk of postmenopausal breast cancer in a large prospective cohort study. Am J Clin Nutr. 2005; 82:1308–1319. [PubMed: 16332665]

Waijers PMCM, Ocke MC, van Rossum CTM, Peeters PHM, Bamia C, Chloptsios Y, et al. Dietary patterns and survival in older Dutch women. Am J Clin Nutr. 2006; 83:1170–1176. [PubMed: 16685062]

Willett, WC. Nutr Epidemiol. 2nd edn.. New York: Oxford University press; 1989.

Willett W. Commentary: dietary diaries versus food frequency questionnaires – a case of undigestible data. Int J Epidemiol. 2001; 30:317–319. [PubMed: 11369736]

Table 1

Factor loadings of various food items in the principal dietary components obtained using PCA based on (1) unadjusted data and (2) data adjusted for energy using the residual method

Food item	***Health conscious***		***Traditional***		***Processed***		***Confectionery***		***'Vegetarian'***		***% Variance explained***	
(Variance explained)	***Unadjusted (10.6%)***	***Adjusted (11.0%)***	***Unadjusted (6.9%)***	***Adjusted (6.4%)***	***Unadjusted (4.4%)***	***Adjusted***	***Unadjusted (6.0%)***	***Adjusted (5.5%)***	***Unadjusted (4.5%)***	***Adjusted (4.0%)***	***Unadjusted (32.4%)***	***Adjusted (26.9%)***
White bread	**−0.535**	**−0.632**	0.075	−0.021	**0.367**	—	0.080	−0.097	−0.018	−0.018	42.7	41.0
Non-white bread	**0.615**	**0.671**	−0.049	−0.011	−0.323	—	−0.057	0.012	0.032	0.054	48.9	45.3
Bran based cereal	**0.365**	**0.386**	0.092	0.019	−0.126	—	−0.004	−0.103	0.009	0.022	15.8	16.0
Oat based cereal	**0.297**	0.295	0.113	0.017	−0.039	—	0.050	−0.105	0.140	0.128	12.5	11.5
Other breakfast cereal	−0.110	−0.188	−0.015	−0.092	0.139	—	0.221	0.066	−0.082	−0.094	8.7	5.7
Biscuits	0.108	0.040	0.023	−0.019	−0.007	—	**0.603**	**0.481**	−0.108	−0.131	38.8	25.0
Crispbreads/ crackers	0.218	0.196	0.088	0.084	−0.010	—	0.052	0.012	0.156	0.109	8.2	5.7
Puddings	0.265	0.163	0.064	−0.030	0.124	—	**0.389**	0.163	−0.112	−0.198	25.4	9.3
Cakes/buns	0.202	0.097	0.004	−0.057	0.086	—	**0.559**	**0.405**	−0.080	−0.156	36.7	20.1
Poultry	0.270	0.153	0.223	0.193	0.121	—	0.023	−0.082	**−0.535**	**−0.577**	42.4	40.0
Red Meat	0.147	0.011	0.219	0.173	0.166	—	0.101	−0.068	**−0.596**	**−0.649**	46.3	45.6
Meat pies	−0.105	**−0.395**	0.032	−0.101	**0.538**	—	0.087	−0.170	−0.118	−0.116	32.3	20.9
Offal	0.087	−0.025	0.091	0.060	0.248	—	−0.066	−0.126	0.087	−0.089	8.9	2.8
Sausages, burgers	−0.091	**−0.392**	−0.062	−0.094	**0.565**	—	0.029	−0.224	−0.169	−0.147	36.1	23.4
Fried foods	−0.094	**−0.415**	0.001	−0.049	**0.574**	—	0.164	−0.138	−0.009	0.043	36.5	19.6
Pizza	0.233	0.018	−0.105	−0.145	**0.349**	—	0.104	−0.103	0.105	0.143	19.8	2.0
Fish	**0.457**	**0.349**	0.155	0.120	0.133	—	−0.075	−0.190	−0.018	−0.061	25.7	20.8
Eggs	0.278	0.031	0.090	0.042	**0.403**	—	−0.027	**−0.349**	−0.016	0.044	24.9	12.6
Cheese	**0.443**	**0.360**	0.078	0.021	0.053	—	0.122	−0.085	0.026	0.070	22.1	14.2
Meat substitutes	0.180	0.134	0.066	0.059	0.124	—	−0.028	−0.037	0.577	0.546	38.6	32.1
Pulses	**0.356**	**0.342**	0.146	0.129	0.006	—	−0.055	−0.097	0.565	0.527	47.0	42.1
Nuts	0.278	0.253	0.116	0.086	0.051	—	0.052	−0.038	**0.531**	**0.476**	37.8	29.9

Food item	*Health conscious*		*Traditional*		*Processed*		*Confectionery*		*'Vegetarian'*		*% Variance explained*	
(Variance explained)	*Unadjusted (10.6%)*	*Adjusted (11.0%)*	*Unadjusted (6.9%)*	*Adjusted (6.4%)*	*Unadjusted (4.4%)*	*Adjusted*	*Unadjusted (6.0%)*	*Adjusted (5.5%)*	*Unadjusted (4.5%)*	*Adjusted (4.0%)*	*Unadjusted (32.4%)*	*Adjusted (26.9%)*
Chips	−0.255	**−0.569**	−0.057	−0.130	0.561	—	0.235	−0.058	−0.036	0.043	44.0	34.6
Roast potatoes	−0.271	**−0.473**	0.225	0.199	0.388	—	0.154	−0.006	−0.165	−0.156	32.6	28.8
Potatoes (not chips)	0.254	0.149	**0.321**	0.291	0.104	—	0.070	−0.099	−0.219	−0.217	23.1	16.4
Pasta	**0.578**	**0.451**	0.045	0.044	0.136	—	−0.070	−0.157	0.121	0.087	37.4	23.8
Rice	**0.543**	**0.438**	0.078	0.071	0.125	—	−0.120	−0.179	0.063	0.018	33.5	22.9
Baked beans	0.004	−0.230	0.049	0.042	**0.413**	—	0.081	−0.148	0.045	0.117	18.2	9.0
Leafy green vegetables	0.045	0.026	**0.809**	**0.796**	0.011	—	−0.015	−0.058	0.041	0.048	65.8	64.0
Other green vegetables	0.147	0.146	**0.799**	**0.789**	−0.043	—	−0.004	−0.025	0.054	0.056	66.5	64.8
Carrots	0.178	0.160	**0.704**	**0.688**	−0.020	—	0.023	−0.028	0.008	0.003	52.8	50.0
Other root vegetables	0.084	0.076	**0.606**	**0.600**	0.018	—	0.003	−0.012	0.106	0.098	38.6	37.6
Peas	0.174	0.024	**0.352**	**0.346**	0.190	—	0.063	−0.088	−0.104	−0.085	20.5	13.5
Salad	**0.420**	**0.404**	0.212	0.183	−0.078	—	−0.022	−0.086	0.100	0.094	23.8	21.3
Fresh fruit	**0.518**	**0.548**	0.182	0.154	−0.229	—	0.090	0.035	0.005	−0.006	36.2	32.5
Fruit juice	**0.488**	**0.474**	0.079	0.035	−0.090	—	0.085	−0.011	0.057	0.029	26.3	22.7
Cola	−0.209	**−0.307**	−0.081	−0.052	0.221	—	0.142	0.105	0.051	0.058	12.2	11.1
Tea	−0.100	−0.155	0.078	−0.009	0.156	—	0.029	−0.137	−0.037	−0.004	4.3	4.3
Coffee	−0.161	−0.185	0.053	0.030	0.105	—	0.002	−0.025	−0.037	−0.018	4.1	3.6
Herbal tea	0.186	0.223	0.068	0.070	−0.085	—	−0.057	−0.047	**0.302**	0.292	14.1	14.2
Sweets	−0.098	−0.171	0.071	0.045	0.069	—	**0.514**	**0.413**	0.061	0.055	28.7	20.5
Chocolate	0.000	−0.088	0.022	−0.016	0.036	—	**0.717**	**0.674**	0.058	0.067	51.9	46.7
Chocolate bars	−0.085	−0.214	−0.020	−0.079	0.096	—	**0.749**	**0.684**	0.021	0.053	57.8	52.3
Crisps	−0.101	−0.295	−0.041	−0.075	0.292	—	**0.381**	0.203	0.004	0.033	24.2	13.5

Abbreviation: PCA, principal component analysis.

Loadings >0.3 are shown in bold.

Table 2

Correlations between unadjusted and energy-adjusted dietary patterns obtained using PCA

Unadjusted	***Health conscious***	***Traditional***	***Processed***	***Confectionery***	***Vegetarian***
Adjusted					
Health conscious	0.869	−0.005	−0.534	−0.090	0.048
Traditional	−0.041	0.972	−0.093	−0.072	0.008
Confectionery	−0.185	−0.046	−0.492	0.803	0.060
Vegetarian	−0.039	−0.011	0.026	−0.008	0.982

Abbreviation: PCA, principal component analysis.

Table 3

Correlation coefficients between dietary pattern score (unadjusted and adjusted for energy) and daily absolute nutrient intakes and partial correlation coefficients between dietary pattern score and daily nutrient intakes adjusting for energy intake (*n*=12055)

Nutrient	***Health-conscious***		***Traditional***		***Processed***		***Confectionery***		***Vegetarian***		***% variance explained***	
	Absolute	**Adjusted**	**Absolute**	**Adjusted**	**Absolute**	**Adjusted**	**Absolute**	**Adjusted**	**Absolute**	**Adjusted**	**Absolute**	**Adjusted**
Energy (MJ)												
Unadjusted pattern	0.319		0.174		0.475		0.548		−0.067		66.2	
Adjusted pattern	0.000		0.000		0.000		0.000		0.000		0.0	
Total fat (g)												
Unadjusted pattern	0.210	−0.186	0.098	−0.136	0.532	0.273	0.522	0.081	−0.077	−0.038	61.4	13.6
Adjusted pattern	−0.121	−0.275	−0.048	−0.110	—	—	0.008	0.018	−0.011	−0.024	1.7	8.9
Sugar (g)												
Unadjusted pattern	0.050	−0.283	0.096	−0.046	0.190	−0.260	0.639	0.417	−0.007	0.061	45.6	32.7
Adjusted pattern	−0.098	−0.143	−0.041	−0.060	—	—	0.295	0.431	0.028	0.041	9.9	21.1
Protein (g)												
Unadjusted pattern	0.564	0.544	0.304	0.279	0.402	0.040	0.217	−0.449	−0.266	−0.357	67.0	70.4
Adjusted pattern	0.260	0.437	0.165	0.278	—	—	−0.255	−0.427	−0.217	−0.365	20.7	58.4
Carbohydrates (g)												
Unadjusted pattern	0.253	−0.131	0.147	−0.044	0.374	−0.215	0.573	0.206	0.006	0.189	55.4	14.3
Adjusted pattern	−0.007	−0.018	−0.023	−0.064	—	—	0.081	0.226	0.064	0.179	1.1	8.8
Calcium (mg)												
Unadjusted pattern	0.379	0.221	0.197	0.100	0.278	−0.151	0.317	−0.191	0.048	0.155	36.3	14.2
Adjusted pattern	0.174	0.271	0.033	0.052	—	—	−0.126	−0.196	0.110	0.172	5.9	14.4
Iron (mg)												
Unadjusted pattern	0.593	0.560	0.288	0.149	0.218	−0.287	0.289	−0.247	0.025	0.278	56.6	55.6
Adjusted pattern	0.375	0.556	0.154	0.228	—	—	−0.131	−0.194	0.061	0.091	18.5	40.7

[*] Correlation coefficients >0.25 $P<0.01$; correlation coefficients >0.2 and <0.25 $P<0.05$; correlation coefficients <0.2 $P>0.05$.

Table 4

Associations of dietary pattern scores (unadjusted and energy adjusted) with birthweight, with and without the inclusion of energy intake

Dietary pattern	*Parameter estimate (95% CI)*	*Parameter estimate (95% CI)*	*Parameter estimates (95% CI) unadjusted pattern score and energy in the same model*	
	Unadjusted -pattern score	*Energy-adjusted pattern score*	*Unadjusted pattern score*	*Energy intake*
Health conscious	29.26 (19.77, 38.75)	34.99 (25.46, 44.52)	39.87 (29.59, 50.16)	−1.33 (−2.04, −0.63)
Traditional	3.03 (−6.49, 12.56)	7.24 (−2.31, 16.80)	6.63 (−3.40, 16.66)	−0.54 (−1.22, 0.14)
Processed	−22.09 (−31.60, −12.58)	—	−18.03 (−29.74, −6.33)	0.09 (−0.67, 0.85)
Confectionery	−4.03 (−13.57, 5.51)	−1.05 (−10.60, 8.50)	−1.25 (−12.90, 10.41)	−0.42 (−1.21, 0.38)
Vegetarian	−17.94 (−27.47, −8.41)	−17.06 (−26.63, −7.48)	−15.47 (−25.44, −5.50)	−0.53 (−1.20, 0.14)