Capturing changes in dietary patterns among older adults: a latent class analysis of an ageing Irish cohort

Janas M Harrington[1],*, Darren L Dahly[1], Anthony P Fitzgerald[1], Mark S Gilthorpe[2] and Ivan J Perry[1]

[1]Department of Epidemiology & Public Health, University College Cork, Fourth Floor, Western Gateway Building, Western Road, Cork, Republic of Ireland: [2]Centre for Epidemiology & Biostatistics, University of Leeds, Leeds, UK

Abstract

Objective: Data-driven approaches to dietary patterns are under-utilized; latent class analyses (LCA) are particularly rare. The present study used an LCA to identify subgroups of people with similar dietary patterns, explore changes in dietary patterns over a 10-year period and relate these dynamics to socio-demographic factors and health outcomes.
Design: The 1998 baseline and 2008 follow-up of the Cork and Kerry Diabetes and Heart Disease Study. Diets were assessed with a standard FFQ. LCA, under the assumption of conditional independence, was used to identify mutually exclusive subgroups with different dietary patterns, based on food group consumption.
Setting: Republic of Ireland.
Subjects: Men and women aged 50–69 years at baseline (*n* 923) and at 10-year follow-up (*n* 320).
Results: Three dietary classes emerged: Western, Healthy and Low-Energy. Significant differences in demographic, lifestyle and health outcomes were associated with class membership. Between baseline and follow-up most people remained 'stable' in their dietary class. Most of those who changed class moved to the Healthy class. Higher education was associated with transition to a healthy diet; lower education was associated with stability in an unhealthy pattern. Transition to a healthy diet was associated with higher CVD risk factors at baseline: respondents were significantly more likely to be smokers, centrally obese and to have hypertension (non-significant).
Conclusions: LCA is useful for exploring dietary patterns transitions. Understanding the predictors of longitudinal stability/transitions in dietary patterns will help target public health initiatives by identifying subgroups most/least likely to change and most/least likely to sustain a change.

The traditional approach to nutritional epidemiology is to relate single nutrients to particular diseases, a classic example being the link between vitamin C and scurvy[1]. This reductionist approach however has many limitations, particularly when applied to outcomes with complex, multifactoral aetiologies such as obesity, diabetes and CVD. Focusing on individual nutrients ignores the fact that people eat foods and meals that include many nutrients and that there are possible synergies among these[2]. It is also difficult to address collinearly among individual nutrient intakes and the approach can encourage multiple testing in search of modest effect sizes[2], which in turn leads to a high number of type 1 errors in published research[3].

Consequently, there has been a recent emphasis on research which identifies holistic dietary patterns in populations as determinants of disease[2,4–16]. Investigating dietary patterns better reflects the real-life scenario that foods are not consumed in isolation and allows us to look at the impact of overall diet on health. Furthermore, studying dietary patterns could have greater public health impact since results may be more easily translated into public health messages than research on specific nutrients[17,18]. However, the key challenge to dietary pattern research lies in reducing the complexity of human diets down to a manageable set of observed patterns in a principled manner.

In 1982, Schwerin *et al.*[19] introduced the concept of food patterns when they examined food habits and health. Slattery *et al.*[20,21] expanded on this and laid the foundations for future data-driven, empirical approaches work when they used factor analysis to identify dietary

Corresponding author: Email j.harrington@ucc.ie

patterns and assess the relationship between these patterns and colon cancer. Methodological development has continued as various research groups have turned to other statistical models[16,22–29]. As such, the use of multivariate statistical models to identify dietary patterns is now fairly well established in nutritional epidemiology[18]. Methods aimed at reducing the dimensionality of dietary data, such as factor analysis or principal components analysis, are more common. Cluster analyses aimed at identifying mutually exclusive groups of people based on a shared dietary pattern are less so, even though they may lead to more widely interpretable results. Among cluster methods, algorithmic approaches such as *k*-means clustering have been more popular, while model-based cluster methods such as latent class analysis (LCA) have been used sparingly[28,30,31] despite their relative advantages. Further, we are only aware of one published study that employed latent class modelling to explore dietary changes over time[29].

Using latent class modelling of multiple food-item intakes, we aimed to describe the change in dietary patterns over a 10-year period among a cohort of middle-aged Irish adults. We were particularly interested in the implications of the stability (or change) in the patterns over time in relation to health outcomes.

Methods

Study population

Data are from the Cork and Kerry Diabetes and Heart Disease Study, a cohort of adults recruited from seventeen general practices in the Cork and Kerry region of Ireland in 1998. All practices affiliated with the Cork training programme for general practice were invited and seventeen of eighteen practices agreed to participate in the study. These practices were broadly representative of the socio-economic profile of the area, with six urban and twelve rural practices. Fifteen practices were in Co. Cork and two were in Co. Kerry. The average list size is approximately 2500 people and this includes both public and private patients.

In summary, a stratified random sampling approach by age and sex was employed across four age strata between the ages of 50 and 69 years. Individuals with CVD, known diabetes mellitus or other diseases or those receiving medication were included in the sampling process. With an anticipated prevalence of type 2 diabetes of 5 %, we estimated that 1000 individuals would be needed to estimate the prevalence with a precision of ±3 %. The sampling strategy for the original cohort has been detailed previously[32–34].

In total, 1473 potential participants were identified and invited to participate in the study. Of these, 1018 attended for the physical examination yielding a 69·1 % response rate. Allowing for those who could not attend by reason of being hospitalized (*n* 6), out of the country (*n* 5), no longer alive (*n* 3), outside the target age group (*n* 2), too

confused (*n* 1) and untraceable (*n* 3), the effective response rate was 70·1 %.

The baseline study was originally designed as a cross-sectional survey, but in 2008 additional funding was obtained to re-contact survey participants and undertake a 10-year rescreen. As the study was not designed as a cohort study, contact had not been maintained with participants during the decade between the first survey and the rescreen. Before making contact with the individual participants, contact was made with the general practices to check whether the participants were still alive and to ascertain whether the rescreening examination was appropriate. Allowing for mortality rate of 12 % (*n* 121), 16 % (*n* 166) loss to follow-up (not currently with the general practice, moved from area) and a further 4 % (*n* 41) classified as too unwell to participate, an available sample of 690 individuals were invited to participate in the rescreen. After follow-up with and review of vital status and causes of death through linkage with the National Cancer Registry, the rescreen response rate was 52 % (*n* 359). Non-responders for the follow-up were slightly older than the responders (aged 60·1 years *v*. 58·4 years), more likely to be retired (30 % *v*. 17 %) and less likely to be married (71 % *v*. 80 %, *P* = 0·02).

Willing participants at each time point provided fasting blood samples, underwent a series of physical measurements and completed a detailed general health and lifestyle questionnaire. Study methods and protocols have previously been described[33–35]. All procedures were carried out with reference to the detailed guidelines outlined in the 'Standard Operating Procedures Manual'[36,37]. The Clinical Research Ethics Committee of the Cork Teaching Hospitals granted ethical permission for the study.

Physical measurements

Anthropometric measurements

Physical measurements included height, weight and waist circumference. All physical measurements were taken by a qualified research nurse trained in the study protocols. In 1998 and 2008 respectively, height was measured using a Harpenden stadiometer and a portable Seca Leicester height measure to the nearest 0·1 cm and weight was measured using a Soehnle digital weighing scales and a portable electronic Tanita WB100MA weighing scale to the nearest 0·1 kg; at both time points waist circumference was measured using a Seca 200 measuring tape to the nearest 0·1 cm.

Participants were classified as overweight or obese based on a BMI of ≥25 kg/m^2 or ≥30 kg/m^2, respectively. Risk of central obesity was classified based on the WHO categorization of waist-to-hip ratio of ≥0·90 for men and ≥0·85 for women[38].

Other physical measurements

At both time points, blood pressure was measured according to the study's standard operating procedures. In summary, blood pressure was measured in a seated position.

Three readings were taken 1 min apart. The average of the second and third readings was noted. Clinical hypertension was defined as systolic blood pressure ≥140 mmHg and/or diastolic blood pressure ≥90 mmHg.

Biochemical and haematological measurements

At both time points participants provided fasting blood samples. Sample processing has been described previously[34]. Diabetes status based on glycated Hb (HbA1c) was defined as per the International Expert Committee Report[39]: diabetes, HbA1c ≥ 6·5 %; prediabetes, HbA1c = 6·0–6·5 %; normal, HbA1c < 6·0 %. Insulin resistance was defined on the basis of fasting glucose and insulin using the glucose homeostasis model assessment of insulin sensitivity (HOMA-IR)[40]. It derives an estimate of insulin sensitivity from the mathematical modelling of fasting plasma glucose and insulin concentrations. HOMA-IR scores were calculated by the formula (fasting insulin × fasting glucose)/22·5. Participants were defined as having insulin resistance if they had a HOMA-IR score >4·65.

Health and lifestyle questionnaire

A health and lifestyle questionnaire was distributed to participants with their examination appointment at baseline and follow-up. The themes of the questionnaire were past medical history (specifically CVD and diabetes), use of medication and lifestyle behaviours including smoking, alcohol use and physical activity. The socio-economic, behavioural and lifestyle characteristics categorized from the questionnaire have been described previously[34]. In summary, for the current analysis, based on questionnaire responses at baseline and follow-up, the following categories were defined.

1. Socio-economic status was classified according to education attainment, marital status and car ownership.
2. Physical activity was classified at baseline based on the British Regional Heart Study Physical Activity instrument[41] and grouped into three physical activity categories: inactive to occasional ('low'), light to moderate ('moderate') and moderate to high ('high')[33]. At follow-up physical activity was assessed using the International Physical Activity Questionnaire (http://www.ipaq.ki.se/scoring.pdf)[42]. Participant scores were classified as high (over 10 000 steps/d), moderate (5000–10 000 steps/d) or low (less than 5000 steps/d). For the current analysis a categorical variable was created: 'low', 'moderate' or 'high'.
3. Smoking status was categorized as 'never smoker', 'current smoker' or 'ex smoker'.
4. Alcohol status was classified into one of four categories based on participants' current average weekly intake: 'never' (no drink), 'occasional/light' (1–<14 units/week), 'moderate/heavy' (>14 units/week)' and 'ex drinker' (former drinker but not a current drinker).
5. Self-rated general health status was classified at baseline as 'excellent', 'good', 'fair' or 'poor' and at follow-up as 'excellent', 'very good', 'good', 'fair' or 'poor'.

Dietary assessment questionnaire

At both time points, participants completed a semi-quantitative FFQ[43] in which they were asked to indicate their average consumption of food items over the last year. In 1998 there were 159 items in the FFQ. Eight food items were added to the 2008 FFQ to reflect changes in food availability. These were vegetarian lasagne, cholesterol-lowering spreads, frozen fruit, sugar substitutes, homemade vegetable soup, homemade meat soup, curry sauce and alcopops. Frequency of consumption of a medium serving or common household unit was asked for each food and then converted into quantities using standard portion sizes. The frequency categories were 'never or less than once a month', '1–3 times per month' 'once a week', '2–4 per week', '5–6 per week', 'once a day', '2–3 per day', '4–5 per day' and '6+ per day'. Items in the FFQ were then expressed in terms of quantities of food/beverage consumed (g/d or ml/d). Nine hundred and twenty-five individuals (91 %) completed the FFQ in 1998, while 320 (89 %) completed it in 2008; 303 individuals completed it at both time points.

DASH scores

Based on validated work[44], we constructed a DASH (Dietary Approaches to Stop Hypertension) score for each completed FFQ[45]. For each food group, consumption was divided into quintiles and participants were classified according to their intake ranking. Consumption of healthy food components was rated on a scale of 1–5, with higher scores reflecting more frequent consumption. Less healthy dietary constituents, where low consumption is desired, were scored on a reverse scale with lower consumption receiving higher scores. Component scores were summed and an overall DASH score was calculated for each person. Overall DASH score was subsequently collapsed to quintiles for analysis – lower quintiles indicated a poorer dietary quality.

Latent class analysis

To identify data-driven dietary patterns using LCA, food items were first aggregated into twenty-three mutually exclusive food groups based previously published work from this cohort[16] (Table 1). For each food group, participants were categorized as those reporting zero intake, with the remainder divided into tertiles of consumption, based on the 1998 distributions (thus a change in category between 1998 and 2008 reflects a change in an individual's intake, not a change in the distribution of intake).

Information from completed FFQ from each time point was stacked into a single analytical data set, with a total of 1245 observations (*n* 925 from 1998, *n* 320 from 2008,

Table 1 Food groups

Food group
Alcoholic beverages
Drinks
Pasta
Potatoes
White bread
Brown breads
Fruit
Vegetables
Chips
Butter
Spreads
Salad dressings
High-fat dairy
Low-fat dairy
Red meats
Poultry
Fish
Meat products
Eggs
Desserts
Sweets
Snacks
Miscellaneous

n 303 at both time points). Using this stacked data set, the observed distribution of the categorized food groups was then modelled as a function of a single multinomial latent class variable, with the assumption that the food group distributions were independent, conditional on class membership. We estimated models ranging from two to eight classes using Mplus 6·11. All thresholds were freely estimated across classes for each food group (i.e. no constraints were made on estimated parameters). Models were not adjusted for energy, sex, age or any other covariates. Because some individuals contributed two observations to the analytical sample, the CLUSTER option was used to adjust standard errors of parameter estimates to account for within-person correlations.

The optimum number of classes was determined based on a combination of: (i) model fit (the Bayesian Information Criterion (BIC), the BIC adjusted for the number of parameters in the model (BICa), the Akaike Information Criterion (AIC) and the log likelihood (LL)); (ii) classification quality, based on the mean probability of class membership among individuals modally assigned to a given class; and (iii) substantive interpretation.

In assessing the dietary transitions, the *post hoc* analysis was confined to participants who had diet recorded at two time points (n 303). While the latent class model is probabilistic with respect to class membership, individuals were categorized by their modal class assignment for the *post hoc* analyses. This decision was supported by the high classification quality characterizing the models. All dietary transitions were relative the 1998 patterns.

All *post hoc* analysis was conducted using the statistical software package Stata 12 IC. ANOVA was used to describe associations between latent class group and continuous variables, while cross-tabulation with a χ^2 significance test was used to test associations with categorical variables. A P value <0·05 was considered statistically significant.

Results

Fit indices for latent class models including two to eight latent classes are given in Fig. 1. The AIC and LL indicated monotonically improving fit as the number of classes included in the model was increased; BICa indicated that model fit was best for the six-class model, but fit was not qualitatively different among the three- to eight-class solutions; BIC indicated that model fit was best for the three- and four-class solutions, with monotonically decreasing fit when more classes were added. For all models and respective classes, the mean probability of class membership given modal class assignment was >80%. For the three-class model, this probability approached 90% for all three classes. Based on this information, preference for model parsimony and the interpretation detailed below, we chose to report results from the three-class model.

Class 1, which we have labelled Western, was characterized by higher consumption of cereals, breads and potatoes, dairy products, meats (in particular red and processed meats) and foods from the top shelf of the food pyramid, especially sweet snacks. Participants in this class had higher dietary salt and overall energy intakes, and were the group with the lowest DASH dietary score (indicating a poorer quality diet compared with the other classes). Class 2, labelled Healthy, was characterized by significantly higher intakes of fruits and vegetables, higher intakes of low-fat dairy products and the highest DASH dietary score. Class 3, labelled Low-Energy, was distinguished from class 1 in having lower consumption of red meat, sweet snacks and overall energy.

Table 2 shows gives a summary of mean daily servings of the food groups which constitute the DASH dietary score for each latent class at baseline and follow-up, and distinguishes one class from another by highlighting the DASH food groups consumed at significantly higher or lower levels relative to the other classes. A clear distinction between dietary groups was evident particularly with respect to the consumption of fruit, vegetables, red meats and confectionery.

Lifestyle characteristics associated with each class at baseline and follow-up

Differences in food patterns were associated with differences in demographic, health and lifestyle characteristics at baseline and follow-up (Table 3). At baseline, differences were seen in sex, marital status, car ownership, alcohol status, self-rated health, central obesity and HDL-cholesterol. In relation to CVD risk factors, participants in the Healthy class were less likely to have diabetes, less
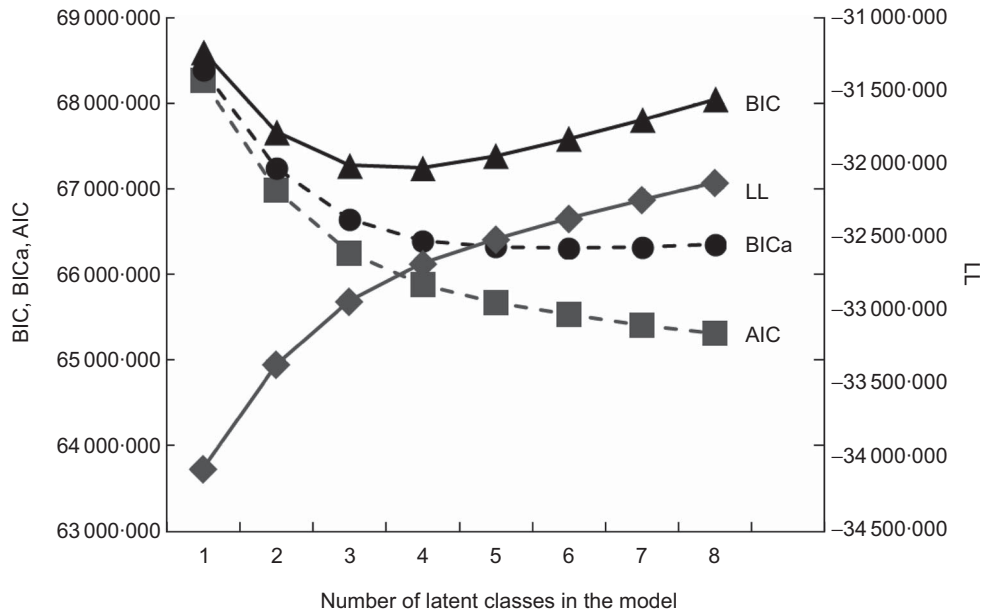
Fig 1 Model fit for latent class models including two to eight latent classes (BIC, Bayesian Information Criterion; BICa, BIC adjusted for the number of parameters in the model; AIC, Akaike Information Criterion; LL, log likelihood)

likely to be at risk of central obesity and more likely to have higher HDL-cholesterol concentration. Higher proportions of non-smokers, non-drinkers and high physical activity participation rates were seen in the Healthy class. At follow-up differences were seen in sex, age, marital status, education, hypertension status, HDL-cholesterol and homocysteine levels. Higher proportions of Healthy class participants (although not significant differences at $P < 0.05$) rated their health as 'excellent', were non-smokers, were occasional/light drinkers and had higher levels of physical activity. Although not significant at $P < 0.05$, the Low-Energy class had a disproportionate number of current smokers. Significant differences ($P < 0.004$) were seen in follow-up class membership in relation to hypertension status. A lower proportion of respondents in the Western class had hypertension compared with the Healthy and Low-Energy classes. Homocysteine levels were lower in the Healthy class than in the Western and Low-Energy classes.

Diet stability and transitions

Of the total sample (n 303), at follow-up 16 % remained in the Western class, 33 % remained in the Healthy class and 6 % remained in the Low-Energy class (Table 4). Within-class stability was highest in the Western class; of those in the Western class at baseline, 68 % remained in this class, while 57 % remained in the Healthy class and 35 % remained in the Low-Energy class. A further 25 % moved from either the Western (19 %) or Low-Energy class (6 %) to the Healthy class between 1998 and 2008.

Table 5 shows the health outcomes for participants whose dietary pattern remained stable over the 10-year period. In terms of stability, there were significant differences between men and women ($P < 0.001$), education group ($P < 0.05$), marital status ($P < 0.01$) and homocysteine levels ($P < 0.05$). More men were stable in the Western class while more women remained stable in the Healthy class. Those who remained stable in the Western class had significantly higher homocysteine levels at follow-up compared with those who remained stable in either the Healthy class or the Low-Energy class.

Analysis of CVD risk factors assessed at baseline and associations with stability or diet transitions (stable in Western, stable in Healthy and transition from Western/ Low-Energy to Healthy) at follow-up are shown in Table 6. Sociodemographic characteristics were significantly associated with class membership. Men were significantly more likely to be in the Western stable group, while women were significantly more likely to be in the Healthy stable group ($P < 0.001$). Participants with primary education were more likely to be in the Western stable group, while participants with tertiary education were more likely to be in the 'transition to healthy diet' group. Significant differences were seen in group membership by smoking status ($P = 0.05$). Current smokers (at baseline) were more likely to be in the 'transition to healthy diet' group. The Healthy stable group had fewer current smokers (at baseline) compared with the Western and Low-Energy groups. Participants in the Healthy stable group had lower baseline levels of homocysteine and central obesity (based on waist-to-hip ratio, $P = 0.01$) compared with the other two groups.

Analysis of health outcomes at follow-up by class pattern showed that the demographic variations seen

Table 2 Food group intakes by latent class group; men and women aged 50–69 years at baseline in 1998 and at 10-year follow up in 2008, Cork and Kerry Diabetes and Heart Disease Study, Republic of Ireland

1998	Western (n 125)		Healthy (n 135)		Low-Energy (n 43)		P value
	Mean	SD	Mean	SD	Mean	SD	
DASH score components (servings/d)							
Whole grains	2·4	2·7	3·2	2·6	2·2	2·6	0·036
Fruits	1·3	1·5	2·3	1·8	0·9	1·9	<0·001
Vegetables	1·9	1·9	2·6	2·3	1·1	1·3	<0·001
Legumes	0·3	0·9	0·3	0·5	0·08	0·1	0·192
Low-fat dairy	0·8	1·7	1·5	1·9	1·2	1·9	0·053
Salty snacks	0·3	0·3	0·3	0·3	0·1	1·2	<0·001
Sweet snacks	3·8	3·2	1·7	1·9	1·2	2·0	<0·001
Red meat	1·3	1·1	0·8	0·7	0·6	0·7	<0·001
Dietary salt (g/d)	8·8	3·4	7·4	2·8	6·3	3·5	<0·001
Mean DASH score	21·1	3·8	25·9	4·5	22·4	3·8	<0·001
kJ/d	9635	3372	8096	2503	6526	2510	<0·001
kcal/d	2302·8	806·0	1934·9	598·2	1559·8	599·8	<0·001

2008	Western (n 72)		Healthy (n 176)		Low-Energy (n 55)		P value
	Mean	SD	Mean	SD	Mean	SD	
DASH score components (servings/d)							
Whole grains	3·8	3·6	3·2	2·0	3·1	1·9	0·242
Fruits	1·8	1·3	2·5	2·0	1·8	1·2	0·0027
Vegetables	2·8	1·2	4·2	2·3	2·8	1·6	<0·001
Legumes	0·2	0·2	0·4	0·4	0·3	0·5	0·148
Low-fat dairy	0·3	0·3	0·3	0·3	0·3	0·3	0·972
Salty snacks	0·4	0·4	0·4	0·4	0·4	0·4	0·871
Sweet snacks	4·5	3·5	1·6	1·7	2·4	2·7	<0·001
Red meat	1·4	1·0	1·1	0·6	1·1	0·9	0·035
Dietary salt (g/d)	9·5	3·6	8·0	2·8	7·0	3·2	<0·001
Mean DASH score	18·2	5·1	22·4	5·8	17·6	5·7	<0·001
kJ/d	9251	2909	7832	2615	6306	2808	<0·001
kcal/d	2211·1	695·3	1871·9	624·9	1507·2	671·2	<0·001

DASH, Dietary Approaches to Stop Hypertension.

with baseline risk factors persisted at follow-up, with significant variations seen across sex, education and marital status (results not shown). Significant differences were seen in homocysteine and cholesterol levels between the classes. At follow-up, homocysteine levels for the Western stable group were higher than those for the Healthy stable and 'transition to healthy diet' groups. The Healthy stable group had higher total cholesterol and HDL-cholesterol concentrations than the other two groups.

Discussion

Principal findings

The current study was an exploratory one to evaluate the use of LCA as a method to assess dietary patterns and dietary transitions in a free-living population. Three primary results are evident from this research. First, three distinct dietary patterns emerged from the analysis: Western, Healthy and Low-Energy classes. These patterns were broadly similar to clusters derived in the original 1998 baseline study using k-means clustering[16], and were associated with significant differences in demographic and health outcomes at baseline and follow-up. Second, over half of all

participants remained 'stable' in their dietary class at follow-up. Third, stability in a Western dietary pattern was associated with increased homocysteine levels, while stability in a Healthy diet pattern was associated with lower rates of central obesity.

Study implications

Three previous studies have used LCA in modelling dietary patterns in cross-sectional analyses[28,30,31]. It is interesting to note that these three other studies found one group with considerably higher mean consumption of meat, similar to our class 1 group (Western). Many epidemiological studies investigating dietary patterns rely on one dietary assessment and assume that dietary patterns remain stable. To our knowledge, the present study is the first to use LCA to model longitudinal dietary transitions in older adults. Sotres-Alvarez et al.[29] documented dietary transition from pregnancy to postpartum. While some studies have investigated dietary stability, many of these use dietary scores based on a limited number of foods or food behaviours and results between studies are inconsistent. Mulder et al.[46] used a dietary score based on seven food behaviours and found great variability over a 4-year period, while Mishra et al.[47]

Table 3 Demographic and lifestyle characteristics by class group at baseline and follow-up; men and women aged 50–69 years at baseline in 1998 and at 10-year follow up in 2008, Cork and Kerry Diabetes and Heart Disease Study, Republic of Ireland

	Baseline							Follow-up						
	Western (n 125)		Healthy (n 135)		Low-Energy (n 43)			Western (n 125)		Healthy (n 135)		Low-Energy (n 43)		
	% or Mean	n or SD	% or Mean	n or SD	% or Mean	n or SD	P value	% or Mean	n or SD	% or Mean	n or SD	% or Mean	n or SD	P value
Sex							<0·001							<0·01
Male	61·6	77	33·6	45	53·5	23		61·3	76	34·1	44	53·3	23	
Female	28·4	48	66·4	89	40·5	20		38·7	48	65·9	85	46·5	20	
Age (years)	58·1	5·5	58·7	5·4	58·1	4·3	0·68	69·5	5·6	68·4	5·1	70·1	5·3	0·07
Marital status							0·03							0·007
Single	10·5	13	4·4	6	16·3	7		9·7	7	4·6	8	16·7	9	
Married	83·9	104	82·2	111	72·1	31		77·8	56	74·9	131	55·6	30	
Separated/divorced/other	5·6	7	13·3	18	11·6	5		12·5	9	20·6	36	27·8	15	
Education							0·53							0·003
Primary	24·8	31	21·5	29	25·6	11		40·8	29	24·6	42	31·5	17	
Secondary	44·8	56	40·7	55	32·6	14		42·2	30	40·4	69	53·7	29	
Tertiary	30·4	38	37·8	51	41·9	18		16·9	12	35·1	60	14·8	8	
Car ownership							0·06							0·76
Yes	93·6	117	89·6	120	81·4	35		88·9	64	91·4	160	88·9	48	
No	6·4	8	10·4	14	18·6	8		11·1	8	8·6	15	11·1	6	
Smoking status							0·16							0·29
Never smoker	48·4	60	52·6	70	53·7	22		47·2	34	53·8	92	51·9	28	
Ex smoker	30·6	38	37·6	50	31·7	13		44·4	32	38·0	65	31·5	17	
Current smoker	21·0	26	9·8	13	14·6	6		8·3	6	8·2	14	16·7	9	
Alcohol status							0·06							0·52
Never	28·1	34	26·9	36	38·5	15		60·3	38	57·2	87	60·8	31	
Occasional/light	54·6	66	54·5	73	30·8	12		23·8	15	33·6	51	25·5	13	
Moderate/heavy	15·7	19	11·9	10	23·1	9		11·1	7	7·9	12	11·8	6	
Ex drinker	1·6	2	6·7	9	7·7	3		4·8	3	1·3	2	2·0	1	
Physical activity							0·33							0·08
Low	43·6	48	31·7	38	30·0	12		38·6	17	32·4	24	28·6	8	
Medium	42·7	47	50·8	61	55·0	22		38·6	17	24·3	18	46·4	13	
High	13·6	15	17·5	21	15·0	6		22·7	10	43·2	32	25·0	7	
Self-rated health							0·02							0·18
Excellent	20·7	25	17·4	23	33·3	14		4·2	3	13·9	24	7·3	4	
Very good	–		–		–			31·9	23	34·1	59	27·3	15	
Good	56·2	68	69·7	92	57·1	24		48·6	35	39·3	68	54·6	30	
Fair/poor	23·1	28	12·9	17	9·5	4		15·3	11	12·7	22	10·9	6	
Blood pressure														
SBP (mmHg)	137·1	18·5	135·2	22·3	132·9	19·9	0·10	143·4	18·9	147·3	18·3	146·2	15·6	0·31
Hypertension	41·6	52	37·0	50	34·9	15	0·65	53·5	38	70·2	120	70·4	38	0·035
Measures of diabetes														
Fasting glucose (mmol/l)	5·0	0·8	4·9	0·6	5·2	1·6	0·14	5·2	0·9	5·3	0·9	5·6	1·2	0·14
Insulin (IU/l)	10·2	6·5	9·4	5·8	10·3	5·5	0·57	11·5	8·0	11·1	6·7	10·1	7·5	0·55
HOMA-IR	2·4	2·0	2·1	1·5	2·5	1·6	0·33	2·8	2·2	2·6	1·6	2·6	2·4	0·69
HOMA-IR >4·65	8·77	10	5·2	6	13·9	5	0·21	14·0	8	8·8	12	13·0	6	0·50
Diabetes status based on HbA1c >6·5%	3·2	4	1·5	2	6·9	3	0·18	45·7	16	40·2	37	32·0	8	0·56

Table 3 *Continued*

	Baseline							Follow-up						
	Western (*n* 125)		Healthy (*n* 135)		Low-Energy (*n* 43)			Western (*n* 125)		Healthy (*n* 135)		Low-Energy (*n* 43)		
	% or Mean	*n* or SD	% or Mean	*n* or SD	% or Mean	*n* or SD	*P* value	% or Mean	*n* or SD	% or Mean	*n* or SD	% or Mean	*n* or SD	*P* value
Measures of obesity														
Central obesity based on WHR	92·8	116	85·2	115	97·6	41	0·03	83·1	59	79·4	139	80·0	44	0·80
BMI (kg/m^2)	27·5	3·4	27·6	4·3	28·4	5·0	0·43	29·7	7·8	28·3	4·1	28·7	5·8	0·19
Overweight/obese	81·5	99	72·9	97	71·4	30	0·23	83·3	60	79·4	139	83·6	46	0·68
Lipid levels														
Total cholesterol (mmol/l)	5·8	0·9	5·9	0·9	5·6	0·7	0·11	1·2	0·5	1·4	0·9	1·3	0·9	0·23
HDL-cholesterol (mmol/l)	1·5	0·4	1·6	0·4	1·4	0·4	0·003	5·1	1·0	5·4	1·1	5·2	0·1	0·08
LDL-cholesterol (mmol/l)	3·6	0·9	3·6	0·8	3·5	0·7	0·57	1·5	0·4	1·6	0·4	1·5	0·4	0·40
TAG (mmol/l)	1·5	0·9	1·5	0·9	1·4	0·5	0·76	3·0	0·8	3·2	0·9	3·1	0·9	0·47
Homocysteine (units)	11·1	5·5	10·2	3·8	11·0	4·7	0·38	12·8	5·1	11·1	2·8	12·0	3·4	0·004

SBP, systolic blood pressure; HOMA-IR, homeostasis model assessment of insulin sensitivity; HbA1c, glycated Hb; WHR, waist-to-hip ratio.
Data are presented as % and *n* for categorical variables or as mean and SD for continuous variables.

Table 4 Breakdown of respondents remaining in baseline class (class stability) or transitioning to a different class between 1998 and 2008; men and women aged 50–69 years at baseline in 1998 and at 10-year follow up in 2008, Cork and Kerry Diabetes and Heart Disease Study, Republic of Ireland

	2008										
	Western			Healthy			Low-Energy				
1998 (*n*)	% within class	*n*	% of total	% within class	*n*	% of total	% within class	*n*	% of total	% stable*	*n*
Total (*n* 303)											
Western (125)	68·1	49	16·2	32·9	58	19·1	32·7	18	5·9	39·0	49
Healthy (135)	23·6	17	5·6	56·8	100	33·0	32·7	18	5·9	74·0	100
Low-Energy (43)	8·3	6	2·0	10·2	18	5·9	34·6	19	6·3	44·0	19
Male (*n* 145)											
Western (77)	70·4	31	21·4	47·8	33	22·8	40·6	13	8·9	40·2	31
Healthy (45)	18·2	8	5·5	44·9	31	21·4	18·7	6	4·1	68·9	31
Low-Energy (23)	11·4	5	3·4	7·3	5	3·4	40·6	13	8·9	56·5	13
Female (*n* 158)											
Western (48)	64·3	18	11·4	23·4	25	15·8	21·8	5	3·2	37·5	18
Healthy (89)	32·1	9	5·7	64·5	69	43·7	52·2	12	7·6	76·4	68
Low-Energy (6)	3·6	1	0·6	12·2	13	8·2	26·1	6	3·8	30·0	6

*Percentage stable in each class; denominator is class *n* at baseline.

Table 5 Lifestyle behaviours and health outcomes at follow-up for dietary stable participants*; men and women aged 50–69 years at baseline in 1998 and at 10-year follow up in 2008, Cork and Kerry Diabetes and Heart Disease Study, Republic of Ireland

Covariates at follow-up	Stable in Western (*n* 49)		Stable in Healthy (*n* 100)		Stable in Low-Energy (*n* 19)		P value
	% or Mean	*n* or SD	% or Mean	*n* or SD	% or Mean	*n* or SD	
Sex							
Male	63·3	31	31·3	31	68·4	13	<0·001
Female	36·7	18	68·7	68	31·6	6	
Age (years)	69·6	5·6	68·8	5·1	68·8	4·5	0·67
Education							
Primary	47·9	23	25·5	25	31·6	6	0·02
Secondary	35·4	17	37·8	37	52·6	10	
Tertiary	16·7	8	36·7	36	15·8	3	
Marital status							
Single	10·2	5	4·0	4	31·6	6	0·002
Married	77·6	38	76·0	76	47·4	9	
Separated/divorced/other	12·2	6	20·0	20	21·0	4	
Car ownership							
Yes	89·9	44	94·0	94	89·5	17	0·59
No	10·2	5	6·0	6	10·5	2	
Smoking status							
Never smoker	46·9	23	52·5	52	61·1	11	0·44
Ex smoker	42·9	21	40·4	40	22·2	4	
Current smoker	10·2	5	7·1	7	16·7	3	
Alcohol status							
Never	60·5	26	55·0	50	72·3	13	0·34
Occasional/light	25·6	11	34·1	31	11·1	2	
Moderate/heavy	9·3	4	9·9	9	16·7	3	
Ex drinker	4·6	2	1·1	1	0·0	0	
Physical activity							
Low	34·4	11	33·3	11	0·0	0	0·14
Medium	40·6	13	28·2	11	60·0	6	
High	25·0	8	38·5	15	40·0	4	
Self-rated health							
Excellent	6·1	3	17·5	17	10·5	2	0·26
Very good	34·7	17	32·0	31	26·3	5	
Good	38·8	19	36·1	35	57·9	11	
Fair/poor	20·4	10	14·4	14	5·3	1	
Blood pressure							
SBP (mmHg)	144·3	19·2	147·4	17·3	142·3	14·4	0·39
Hypertension	56·3	27	72·4	71	72·2	13	0·13
Measures of diabetes							
Fasting glucose (mmol/l)	5·1	0·5	5·2	0·6	5·5	1·2	0·13
Insulin (IU/l)	10·4	7·0	10·6	6·2	9·9	6·3	0·92
HOMA-IR	2·5	1·8	2·6	1·6	2·5	2·0	0·98
HOMA-IR >4·65	11·9	5	8·9	7	20·0	3·0	0·45
Diabetes status based on HbA1c >6·5 %	48·0	12	38·0	19	12·5	1	0·19
Measures of obesity							
Central obesity based on WHR	77·1	37	78·0	78	73·7	14	0·92
BMI (kg/m²)	28·9	8·7	28·0	4·2	28·8	7·2	0·69
Overweight/obese	79·6	39	79·0	79	78·9	15	0·99
Lipid levels							
Total cholesterol (mmol/l)	5·2	0·9	5·6	1·1	5·3	0·8	0·09
HDL-cholesterol (mmol/l)	1·5	0·4	1·6	0·4	1·5	0·4	0·14
LDL-cholesterol (mmol/l)	3·1	0·8	3·3	1·0	3·2	0·6	0·57
TAG (mmol/l)	1·2	0·5	1·4	0·7	1·3	0·5	0·19
Homocysteine (units)	13·0	5·5	11·1	2·8	11·2	2·6	0·03

SBP, systolic blood pressure; HOMA-IR, homeostasis model assessment of insulin sensitivity; HbA1c, glycated Hb; WHR, waist-to-hip ratio.
Data are presented as % and *n* for categorical variables or as mean and SD for continuous variables.
*This analysis relates only to participants who remained stable in their dietary class from baseline to follow-up.

used a dietary pattern score based on factor analysis and found that healthier eating behaviours were more stable than other behaviours. In contrast, van Dam *et al.*[48] found higher stability with the Western pattern in comparison to the 'Prudent' dietary pattern. The present study, similar to Mishra *et al.*[47], shows high dietary stability over time, particularly for the healthy diet pattern.

The sociodemographic determinants of diet are well documented[16,45,49–55], particularly the relationship between diet and educational status[56,57]. We found that educational status is also an important predictor of dietary stability and transition. Those with higher levels of education were more likely to have a healthy diet and to remain stable in the healthy diet group over time. Those who transitioned

Table 6 Univariate analysis of CVD factors at baseline as predictors of diet transitions/diet stability at follow-up; men and women aged 50–69 years at baseline in 1998 and at 10-year follow up in 2008, Cork and Kerry Diabetes and Heart Disease Study, Republic of Ireland

Covariates at baseline	Stable Western (n 68)		Stable Healthy (n 100)		Transition to healthy diet* (n 76)		
	% or Mean	n or SD	% or Mean	n or SD	% or Mean	n or SD	P value
Sex							
Male	64·7	44	31·3	31	50·0	38	<0·001
Female	35·3	24	68·7	68	50·0	38	
Age (years)	58·7	5·2	58·3	5·2	57·1	5·0	0·17
Education							
Primary	33·8	38	21·0	21	14·5	11	0·06
Secondary	39·7	27	42·0	42	42·1	32	
Tertiary	26·5	18	37·0	37	43·4	33	
Marital status							
Single	16·2	11	4·0	4	7·9	6	0·08
Married	77·9	53	86·0	86	82·9	63	
Separated/divorced/other	5·9	4	10·0	10	9·2	7	
Car ownership							
Yes	88·2	60	89·9	89	93·4	71	0·55
No	11·8	8	10·1	10	6·6	3	
Smoking status							
Never smoker	50·8	34	53·1	52	54·7	41	0·05
Ex smoker	32·8	22	38·8	38	22·7	17	
Current smoker	16·4	11	8·2	8	22·7	17	
Alcohol status							
Never	38·2	26	25·3	25	30·0	21	0·29
Occasional/light	42·6	29	56·6	56	48·6	34	
Moderate/heavy	13·2	9	13·1	13	20·0	14	
Ex drinker	5·9	4	5·1	5	1·4	1	
Physical activity							
Low	438·4	28	33·3	30	38·4	28	0·57
Medium	48·3	28	47·8	43	42·5	31	
High	10·3	6	18·9	17	19·2	14	
Self-rated health							
Excellent	26·2	17	18·4	18	20·0	15	0·13
Good	47·7	31	67·4	66	64·0	48	
Fair/poor	26·2	17	14·3	14	16·0	12	
Blood pressure							
SBP (mmHg)	134·6	17·9	135·8	23·4	136·9	19·8	0·79
Hypertension	36·8	25	36·0	36	42·1	32	0·68
Measures of diabetes							
Fasting glucose (mmol/l)	5·0	1·3	4·9	0·6	4·9	0·7	0·77
Insulin (IU/l)	9·6	5·6	8·7	4·1	9·3	5·1	0·52
HOMA-IR	2·2	1·5	1·9	1·1	2·1	1·4	0·42
HOMA-IR >4·65	10·0	6	2·4	2	4·4	3	0·12
Diabetes status (inc. prediabetes) based on HbA1c >6·0%	1·5	1	2·04	2	3·95	3	0·59
Measures of obesity							
Central obesity based on WHR	95·6	65	83·0	83	93·3	70	0·01
BMI (kg/m²)	27·1	4·2	27·2	4·0	27·8	3·4	0·51
Overweight/obese	72·7	48	70·0	70	81·6	62	0·20
Lipid levels							
Total cholesterol (mmol/l)	5·5	0·8	6·0	1·0	6·0	1·0	0·003
HDL-cholesterol (mmol/l)	1·5	0·4	1·7	0·4	1·5	0·3	0·007
LDL-cholesterol (mmol/l)	3·4	0·7	3·7	0·9	3·8	0·9	0·02
TAG (mmol/l)	1·4	0·8	1·5	1·1	1·5	0·8	0·58
Homocysteine (units)	11·8	7·3	9·9	3·8	10·4	3·5	0·08

SBP, systolic blood pressure; HOMA-IR, homeostasis model assessment of insulin sensitivity; HbA1c, glycated Hb; WHR, waist-to-hip ratio.
Data are presented as % and n for categorical variables or as mean and SD for continuous variables.
*Relates to participants who moved from either the Western or Low-Energy class to the Healthy class.

to the healthy diet class from the Western class were more likely to be participants with higher education, while those with lower education tended to remain in the Western class.

Many studies which use *a posteriori* methods to determine dietary patterns adjust for energy intake, age and sex in their model estimations. Our models did not adjust for these variables. We chose not to adjust for energy intake as total energy intake has been shown to be poorly estimated using FFQ. In relation to age, all participants were in a very narrow age band and few differences were seen by age category in univariate analysis. Further, we believed that adjustment for sex would limit the distinction between classes. Inclusion of these variables in the model estimation would reduce the opportunity to explore the determinants of dietary patterns and the stability/transitions over time.

Study strengths and limitations

We chose an LCA over algorithmic approaches to cluster detection such as *k*-means. Like most other cluster methods, the key challenge lies in choosing the 'correct' number of classes. However, because LCA is a model-based method, it is possible to use goodness-of-fit tests to help identify the ideal number of subgroups. This is especially important given that the external validity of the model cannot be assessed since the class labels are not known. Furthermore, once the model has been estimated, posterior probabilities of class membership for each observation can be calculated, providing additional information for model selection. Further, latent class models can be more easily extended under a structural equation modelling framework to explicitly incorporate longitudinal data, additional covariates, multiple outcomes and other latent variables in the model.

The manner in which we dealt with the longitudinal nature of our data bears some consideration. We chose not to use a latent class transition model[29], which would be composed of a separate latent class variable at each time point connected by a multinomial regression model. Our primary concern with this approach, given the exploratory nature of our investigation, was that small differences in parameter estimates between the two time points could risk being over-interpreted. Equality constraints could certainly be imposed on a latent class transition model to help address this, but lacking strong *a priori* knowledge, this can lead to a slippery slope of model testing and optimization that can lead to model overfit. It also requires one to specify the number of latent classes at each time point.

Instead we 'stacked' the observations from each time point into a single analytical data set and used this to estimate a single latent class variable. Viewing this single variable as a mix of classes from 1998 and 2008, the model is conservative in that minor differences in the parameters would be ignored, with the larger number of observations from 1998 appropriately having a greater influence on parameter estimates. Conversely, any substantial differences such as the emergence or disappearance of a class would still be evident in the model. Compared with the latent class transition approach, our model is considerably more parsimonious and requires fewer decisions to be made by the researcher, such as specifying the number of latent classes for two time points and setting parameter constraints. The key limitation to our approach, however, was that we were not able to explicitly estimate class transition probabilities between years. We instead had to assign observations to their modal class assignment based on posterior probabilities of class membership and analyse these assignments *post hoc* in individuals observed at each time point.

There are limitations to the study which need to be considered in interpreting the results. The Cork and Kerry Diabetes and Heart Disease Study (baseline and follow-up) was conducted in a general practice setting. In Ireland, there is no register of general practices, and individuals are free to move between practices if they so wish without notification. Coupled with the ageing cohort this can increase loss to follow-up for studies of this nature. One of the problems in surveying older people is the high proportion of non-responders, but also of drop-outs largely due to deaths[58]. However, due to the nature of the study, the healthier population also yields advantages. In less healthy elderly populations, the relatively rapid decline in functional status as a result of progressive illness is likely to be evident.

Further limitations are associated with measurement error in relation to the exposure (diet). The FFQ is associated with random error and there is likely some recall bias. However, the FFQ used in these studies has been validated for use in the Irish population[59] and has been used repeatedly in the National Health and Lifestyle Studies[43,60,61] to monitor trends in the population. Although the numbers are low, the current study is strengthened by the completion of the same dietary measurement instrument at both time points by participants. Our study is one of the few studies to track dietary patterns over a 10-year period in an older population. Retrospective dietary assessment is difficult, particularly in an ageing population where cognitive ability might be diminished; however, because of the method of recruitment, the participants in our study were probably more healthy, mobile and socially outgoing than would occur in a representative sampling of this age group in the community. Further, each participant was met by the study team as part of the physical examination component of the study and, where necessary, a member of the research team provided assistance with the questionnaires.

The current study was based on food consumption data at two time points. It did not take account of the possible changes in food supply over this time period, and changes in dietary patterns could potentially be due to food supply changes. The report from the European Heart Network[62] indicates that between 1998 and 2008, supplies in vegetable oils and fruits increased, animal fats decreased, sugar supplies fluctuated and the supply of vegetables remained stable. While there were changes in the food supply, the cross-sectional food consumption patterns reflect the dietary patterns of the Irish population at these two time periods[43,60,61].

Conclusions

The present study confirms that LCA is a useful method for estimating dietary patterns and their stability/transition over time. Using these methods, it was possible to confirm that healthy dietary patterns are associated with sociodemographic characteristics and that educational status is associated with both healthy diet stability and transition to healthy diets. Given the strong evidence for

the role of diet in the development of chronic diseases, in particular hypertension, diabetes and CVD, it is important to reliably identify the predictors of dietary transitions and stability. Robust methods are required to determine the predictors of these transitions and we propose that LCA is a useful complement to more commonly applied *a posteriori* methods.

Acknowledgements

References

1. Fawns HT (1975) Discovery of vitamin C: James Lind and the scurvy. *Nurs Times* **71**, 872–875.
2. Hu F (2002) Dietary pattern analysis a new direction in nutrition epidemiology. *Curr Opin Lipidol* **13**, 3–9.
3. Ioannidis JPA (2005) Why most published research findings are false. *PLoS Med* **2**, e124.
4. Anderson A, Harris T, Houston D *et al.* (2010) Relationships of dietary patterns with body composition in older adults differ by gender and PPAR-γ Pro12Ala genotype. *Eur J Nutr* **49**, 385–394.
5. Anderson AL, Harris TB, Tylavsky FA *et al.* (2011) Dietary patterns and survival of older adults. *J Am Diet Assoc* **111**, 84–91.
6. Berg CM, Lappas G, Strandhagen E *et al.* (2008) Food patterns and cardiovascular disease risk factors: the Swedish INTERGENE research program. *Am J Clin Nutr* **88**, 289–297.
7. Bhupathiraju SN & Tucker KL (2011) Coronary heart disease prevention: nutrients, foods, and dietary patterns. *Clin Chim Acta* **412**, 1493–1514.
8. Centritto F, Iacoviello L, di Giuseppe R *et al.* (2009) Dietary patterns, cardiovascular risk factors and C-reactive protein in a healthy Italian population. *Nutr Metab Cardiovasc Dis* **19**, 697–706.
9. Cho YA, Kim J, Cho ER *et al.* (2011) Dietary patterns and the prevalence of metabolic syndrome in Korean women. *Nutr Metab Cardiovasc Dis* **21**, 893–900.
10. Costacou T, Bamia C, Ferrari P *et al.* (2003) Tracing the Mediterranean diet through principal components and cluster analyses in the Greek population. *Eur J Clin Nutr* **57**, 1378–1385.
11. Davis MS, Miller CK & Mitchell DC (2004) More favorable dietary patterns are associated with lower glycemic load in older adults. *J Am Diet Assoc* **104**, 1828–1835.
12. Delavar M, Lye MS, Khor GL *et al.* (2009) Dietary patterns and the metabolic syndrome in middle aged women, Babol, Iran. *Asia Pac J Clin Nutr* **18**, 285–292.
13. Hydrie M, Basir A, Shera A *et al.* (2010) Dietary patterns associated with risk for metabolic syndrome in urban community of Karachi defined by cluster analysis. *Pakistan J Nutr* **9**, 93–99.
14. Kant AK (1996) Indexes of overall diet quality: a review. *J Am Diet Assoc* **96**, 785–791.
15. Kant AK (2010) Dietary patterns: biomarkers and chronic disease risk. *Appl Physiol Nutr Metab* **35**, 199–206.
16. Villegas R, Salim A, Collins MM *et al.* (2004) Dietary patterns in middle-aged Irish men and women defined by cluster analysis. *Public Health Nutr* **7**, 1017–1024.
17. Hu F (2008) Dietary assessment methods. In *Obesity Epidemiology*, 1st ed., pp. 108–111 [F Hu, editor]. New York: Oxford University Press.
18. Fahey MT, Ferrari P, Slimani N *et al.* (2012) Identifying dietary patterns using a normal mixture model: application to the EPIC study. *J Epidemiol Community Health* **66**, 89–94.
19. Schwerin H, Stanton J, Smith J *et al.* (1982) Food, eating habits, and health: a further examination of the relationship between food eating patterns and nutritional health. *Am J Clin Nutr* **35**, 1319–1325.
20. Slattery ML & Boucher KM (1998) The senior authors' response: factor analysis as a tool for evaluating eating patterns. *Am J Epidemiol* **148**, 20–21.
21. Slattery ML, Boucher KM, Caan BJ *et al.* (1998) Eating patterns and risk of colon cancer. *Am J Epidemiol* **148**, 4–16.
22. Bailey RL, Gutschall MD, Mitchell DC *et al.* (2006) Comparative strategies for using cluster analysis to assess dietary patterns. *J Am Diet Assoc* **106**, 1194–1200.
23. DiBello JR, Kraft P, McGarvey ST *et al.* (2008) Comparison of 3 methods for identifying dietary patterns associated with risk of disease. *Am J Epidemiol* **168**, 1433–1443.
24. Lo Siou G, Yasui Y, Csizmadi I *et al.* (2011) Exploring statistical approaches to diminish subjectivity of cluster analysis to derive dietary patterns. *Am J Epidemiol* **173**, 956–967.
25. Panagiotakos DB, Pitsavos C & Stefanadis C (2009) α-Priori and α-posterior dietary pattern analyses have similar estimating and discriminating ability in predicting 5-Y incidence of cardiovascular disease: methodological issues in nutrition assessment. *J Food Sci* **74**, H218–H224.
26. Villegas R, Salim A, Flynn A *et al.* (2004) Prudent diet and the risk of insulin resistance. *Nutr Metab Cardiovasc Dis* **14**, 334–343.
27. Villegas R, Yang G, Gao Y-T *et al.* (2010) Dietary patterns are associated with lower incidence of type 2 diabetes in middle-aged women: the Shanghai Women's Health Study. *Int J Epidemiol* **39**, 889–899.
28. Sotres-Alvarez D, Herring AH & Siega-Riz AM (2010) Latent class analysis is useful to classify pregnant women into dietary patterns. *J Nutr* **140**, 2253–2259.

29. Sotres-Alvarez D, Herring AH & Siega-Riz A-M (2013) Latent transition models to study women's changing of dietary patterns from pregnancy to 1 year postpartum. *Am J Epidemiol* **177**, 852–861.
30. Fahey MT, Thane CW, Bramwell GD et al. (2007) Conditional Gaussian mixture modelling for dietary pattern analysis. *J R Stat Soc Ser A (Stat Soc)* **170**, 149–166.
31. Padmadas SS, Dias JG & Willekens FJ (2006) Disentangling women's responses on complex dietary intake patterns from an Indian cross-sectional survey: a latent class analysis. *Public Health Nutr* **9**, 204–211.
32. Creagh D, Neilson S, Collins A et al. (2002) Established cardiovascular disease and CVD risk factors in a primary care population of middle-aged Irish men and women. *Irish Med J* **95**, 298–301.
33. Villegas R (2002) *The Role of Diet and Other Environmental Modifiable Risk Factors in the Development of Type 2 Diabetes*. Cork: University College Cork.
34. Kearney PM, Harrington JM, McCarthy VJ et al. (2013) Cohort profile: the Cork and Kerry Diabetes and Heart Disease Study. *Int J Epidemiol* **42**, 1253–1262.
35. Creagh D, Neilson S & Perry IJ (1998) *The Cork and Kerry Diabetes and Heart Disease Study*. Cork: University College Cork, Department of Epidemiology & Public Health.
36. Department of Epidemiology & Public Health (1998–2008 *Cork and Kerry Diabetes and Heart Disease Study: Standard Operating Procedures*. Cork: University College Cork.
37. Department of Epidemiology & Public Health (2000) *The Cork and Kerry Diabetes and Heart Disease Study: A Report to the Department of Health and the Southern Health Board*. Cork: University College Cork.
38. World Health Organization (2008) *Waist Circumference and Waist–Hip Ratio. Report of a WHO Expert Consultation, Geneva, 8–11 December 2008*. Geneva: WHO.
39. The International Expert Committee (2009) International expert committee report on the role of the A1C assay in the diagnosis of diabetes. *Diabetes Care* **32**, 1327–1334.
40. Matthews D, Hosker J, Rudenski A et al. (1985) Homeostasis model assessment: insulin resistance and β-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* **28**, 412–419.
41. Shaper AG, Wannamethee G & Weatherall R (1991) Physical activity and ischaemic heart disease in middle-aged British men. *Br Heart J* **66**, 384–394.
42. Morgan K, McGee H, Watson D et al. (2008) *SLÁN 2007: Survey of Lifestyles, Attitudes and Nutrition in Ireland. Main Report*. Dublin: Department of Health and Children.
43. Harrington J, Perry IJ, Lutomski J et al. (2008) *SLÁN 2007: Survey of Lifestyles, Attitudes and Nutrition in Ireland. Dietary Habits of the Irish Population*. Dublin: Department of Health and Children.
44. Fung TT, Chiuve SE, McCullough ML et al. (2008) Adherence to a DASH-style diet and risk of coronary heart disease and stroke in women. *Arch Intern Med* **168**, 713–720.
45. Harrington J, Fitzgerald AP, Layte R et al. (2011) Socio-demographic, health and lifestyle predictors of poor diets. *Public Health Nutr* **14**, 2166–2175.
46. Mulder M, Ranchor AV, Sanderman R et al. (1998) The stability of lifestyle behaviour. *Int J Epidemiol* **27**, 199–207.
47. Mishra GD, McNaughton SA, Bramwell GD et al. (2006) Longitudinal changes in dietary patterns during adult life. *Br J Nutr* **96**, 735–744.
48. van Damm RM, Rimm EB, Willett WC et al. (2002) Dietary patterns and risk for type 2 diabetes mellitus in US men. *Ann Intern Med* **136**, 201–209.
49. Layte R, Harrington J, Sexton E et al. (2011) Irish exceptionalism? Local food environments and dietary quality. *J Epidemiol Community Health* **65**, 881–888.
50. Deshmukh-Taskar PR, O'Neil CE, Nicklas TA et al. (2009) Dietary patterns associated with metabolic syndrome, sociodemographic and lifestyle factors in young adults: the Bogalusa Heart Study. *Public Health Nutr* **12**, 2493–2503.
51. Ervin RB (2008) Healthy Eating Index scores among adults, 60 years of age and over, by sociodemographic and health characteristics: United States, 1999–2002. *Adv Data* issue 395, 1–16.
52. Ferrer C, Garcia-Esteban R, Mendez M et al. (2009) Social determinants of dietary patterns during pregnancy. *Gac Sanit* **23**, 38–43.
53. Hulshof KF, Wedel M, Lowik MR et al. (1992) Clustering of dietary variables and other lifestyle factors (Dutch Nutritional Surveillance System). *J Epidemiol Community Health* **46**, 417–424.
54. Kesse-Guyot E, Bertrais S, Peneau S et al. (2009) Dietary patterns and their sociodemographic and behavioural correlates in French middle-aged adults from the SU.VI.-MAX cohort. *Eur J Clin Nutr* **63**, 521–528.
55. Lutomski J, Van den Broeck J, Harrington J et al. (2011) Sociodemographic, lifestyle, mental health and dietary factors associated with direction of misreporting of energy intake. *Public Health Nutr* **14**, 532–541.
56. Mishra GD, Ball K, Arkbuckle J et al. (2002) Dietary patterns of Australian adults and their association with socioeconomic status: results from the 1995 National Nutrition Survey. *Eur Clin Nutr* **56**, 687–693.
57. 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.
58. Van Staveren WA, de Groot LC & Haveman-Nies A (2002) The SENECA study: potentials and problems in relating diet to survival over 10 years. *Public Health Nutr* **5**, 901–905.
59. Harrington J (1997) *Validation of a Food Frequency Questionnaire as a Tool for Assessing Nutrient Intake*. Galway: NUI, Galway.
60. Friel S, Nic Gabhainn S & Kelleher C (1999) *The National Lifestyle Surveys: Survey of Lifestyle, Attitudes and Nutrition (SLÁN) and the Irish Health Behaviour in School-Aged Children Survey (HBSC)*. Dublin: Department of Health and Children.
61. Kelleher C, Nic Gabhainn S, Friel S et al. (2003) *The National Health and Lifestyle Surveys: Survey of Lifestyle, Attitudes and Nutrition (SLÁN 2002) and The Irish Health Behaviour in School-Aged Children (HBSC)*. Galway and Dublin: Centre for Health Promotion Studies, NUI, Galway and Department of Public Health Medicine and Epidemiology, UCD.
62. European Heart Network (2011) *Diet, Physical Activity and Cardiovascular Disease Prevention in Europe*. Brussels: European Heart Network.