British Journal of Nutrition (2006), **95**, 860–869 DOI: 10.1079/BJN20061731

Review Article

Methodological approaches to study dietary patterns in relation to risk of coronary heart disease and stroke

Matthias B. Schulze* and Kurt Hoffmann

Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Arthur-Scheunert-Allee 114-116, 14558 Nuthetal, Germany

(Received 24 August 2005 – Revised 4 November 2005 – Accepted 19 December 2005)

Dietary pattern analysis, which reflects the complexity of dietary intake, has received considerable attention by nutritional epidemiology. For a long time, two general approaches have been used to define these summary variables in observational studies. The exploratory approach is based only on the data of the study, whereas the hypothesis-oriented approach constructs pattern variables based on scientific evidence available before the study. Recently, a new statistical method, reduced rank regression, was applied to nutritional epidemiology that is exploratory by nature, but can use scientific evidence by focusing on disease-related dietary components or biomarkers. Several studies, both observational and clinical, suggest that dietary patterns may predict the risk of CHD and stroke. In the present review, we describe the results of these studies and the available evidence regarding the relationships between dietary patterns and risk of CVD and we discuss limitations and strengths of the statistical methods used to extract dietary patterns.

Dietary patterns: Food habits: Coronary heart disease: Stroke: Statistical methods

Diseases of the heart and circulatory system (CVD) are the main cause of death in Europe (Petersen *et al.* 2005) and worldwide (World Health Organization, 2005). Nearly half of all deaths in Europe are from CVD. The main forms of CVD are CHD and stroke. They account for about 3·23 million deaths in Europe each year (Petersen *et al.* 2005). In addition, the economic costs of these diseases are enormous. For example, in 2003, direct medical and treatment costs, informal care costs, and indirect costs due to disability and mortality have been estimated at about 79 billion € in the European Union (Petersen *et al.* 2005). Therefore, primary prevention is of paramount importance, and risk factors for these diseases, among them diet, have long been a focus of epidemiological research.

With regard to diet, however, the dominant approach in the past, that of examining single nutrients or foods, is fraught with problems due to the complexity of dietary intake in relation to diseases. Nutrients may interact with each other and influence their bioavailability and absorption. Because many nutrients are contained in the same foods and foods are eaten in combination, it is difficult to attribute effects to single dietary components. Furthermore, given the relative stability of energy intake at the individual level, changes in dietary habits are generally characterised by substitution effects for macronutrient intake. For example, high-carbohydrate diets will tend to be low-fat and/or low-protein diets as well. Similar substitution effects reflect food consumption habits where foods are chosen alternatively and a high consumption of one food (for example, potatoes) must be associated with lower intake of other foods (for example, rice and pasta). These make inferences about the relevance of individual foods and nutrients particularly troublesome. Therefore, to gain full understanding of the relationship between diet and the development of CVD, it is desirable to use several methodological approaches. The study of food patterns, besides studying individual foods and nutrients or food components, in reducing disease risk is seen as a complementary strategy here (Jacobs & Steffen, 2003). Dietary patterns account for cumulative and interactive effects and reflect real-world dietary preferences and may be particularly suitable for analyses in nutritional epidemiology when many dietary components are relevant for a disease. Dietary pattern analysis might therefore be particularly useful in the context of prevention of CVD. The objective of the present paper is to review published studies of dietary patterns for CHD and stroke and to discuss methodological issues relevant for epidemiological applications.

Abbreviations: CRP, C-reactive protein; EPIC, European Prospective Investigation into Cancer and Nutrition; RR, relative risk.
* **Corresponding author:** Dr Matthias B. Schulze, fax +49 33 200 88 721, email mschulze@mail.dife.de

Reference selection procedures

To identify dietary studies that evaluated dietary patterns with regard to CHD and stroke, we searched PubMed (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db = PubMed) using the terms 'factor analysis', 'principal component analysis', 'cluster analysis', 'index', and 'score' in combination with 'patterns or habits' and 'dietary, food, or nutrients' as well as 'cardiovascular disease, coronary heart disease, coronary artery disease, myocardial infarction, or stroke'. Reference lists from selected articles and from reviews on dietary pattern analysis were also reviewed to locate additional papers that were not retrieved in the PubMed search.

Methods of pattern analysis

Two general approaches have been used to define dietary patterns in observational studies (Jacques & Tucker, 2001; Schulze & Hu, 2002) (Fig. 1). On the one hand, patterns can be defined in a hypothesis-oriented way based on available scientific evidence for specific diseases. The evidence may here be drawn from studies on single dietary components, for example, single nutrients that may be beneficial or detrimental for CVD, or from studies on overall dietary habits, for example, Mediterranean or vegetarian diets. It is also common to use pattern scores as summary measures of the degree to which an individual's diet conforms to specific dietary recommendations. For example, the Healthy Eating Index (Kennedy *et al.* 1995) is a summary measure of the degree to which an individual's diet conforms to the recommendations set out by the previous United States Department of Agriculture Food Guide Pyramid (Kennedy *et al.* 1995) and to specific recommendations in the 1990 US Dietary Guidelines for Americans (Kennedy *et al.* 1995). Hypothesis-oriented patterns are defined as a composite score from various food items or nutrients. Individual components are summed as dichotomised variables (based on predefined cut-points), ordinal variables (for example, quintiles), or as continuous variables, where in the latter case some sort of standardisation allows to assign equal weights to each component (Schulze *et al.* 2003*a*). Such summary variables do not reflect overall diet, but rather selected aspects of individual nutrition.

In contrast to hypothesis-oriented patterns, exploratory methods rely totally on the data at hand and derive patterns purely empirically. Principal component analysis and exploratory factor analysis, and cluster analysis are the two predominant exploratory methods applied in nutritional epidemiology. Both might work well in identifying the major dietary patterns of a particular study population (Hu, 2002; Newby & Tucker, 2004). Principal component analysis identifies foods that are frequently consumed together. It aggregates food items or food groups on the basis of the degree to which they are correlated with one another. The goal is to identify linear composites of food items or food groups that account for the largest amount of variation in diet between individuals. While in principal component analysis the component score represents a mathematical transformation (a linear combination) of the observed variables, factor scores computed in factor analysis are considered only estimates of where individuals stand on the actual underlying and unobservable factor. Besides the theoretical differences between both methods, factor analysis based on the principal factor method (usually applied in dietary pattern analyses) gives generally similar results as principal component analysis. Initial patterns identified are usually rotated by an orthogonal transformation to achieve simple structure with greater interpretability. The factor score for each pattern can be computed by combining the standardised food variables with weights that are proportionate to their component loadings. However, a simpler approach is to combine with equal weight only those standardised food groups that showed high factor loadings (Schulze *et al.* 2003*a*). Food groups with high loadings can easily be identified if the pattern structure is simple (high or near-zero loadings), but this is frequently not the case for dietary patterns (Schulze & Hu, 2002). Chronbach's coefficient α (Bland & Altman, 1997) and confirmatory factor analysis (Bentler, 1980) allow to formally test which food items contribute significantly to the pattern, but these techniques have rarely been applied in dietary pattern analysis (Gittelsohn *et al.* 1998; Maskarinec *et al.* 2000; Schulze *et al.* 2003*b*).

Cluster analysis defines mutually exclusive clusters of individuals. Cluster analysis is based on distance measures between observations of individuals. Initial cluster seeds are followed by repeated comparisons between the means of initial clusters and subsequent updates of cluster groupings and means. Subjects are moved between clusters and new means are computed until the distances between the observations within clusters are small enough compared with the distances between cluster means. Applying cluster analysis to nutritional science, dietary differences between clusters need to be descriptively evaluated. Because variables with large variances tend to have a greater effect on the resulting clusters than those with small variances, food intake should be standardised in advance or the percentage of energy contributed by each food should be used in the cluster analysis (Newby & Tucker, 2004).

Just recently, we have introduced reduced rank regression into nutritional epidemiology as an additional tool for dietary pattern analysis (Hoffmann *et al.* 2004*a*). Reduced rank regression does not focus on explaining variance between foods like principal components analysis but identifies linear functions of predictors (for example, food groups) that explain as much variation as possible in a set of intermediate response variables, for example, nutrients (Hoffmann *et al.* 2004*a*) or biomarkers (Hoffmann *et al.* 2004*b*). Since it uses both available dietary data of the study and prior knowledge for choosing appropriate response variables, it is a mix of an exploratory and a hypothesis-oriented approach. Reduced rank regression is similar to factor analysis in its mathematical foundation and technique of deriving factors. It can be interpreted as a principal component analysis applied to responses

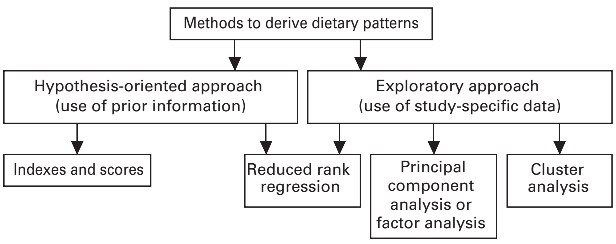

Fig. 1. Approaches to define dietary patterns in observational studies.

https://doi.org/10.1079/BJN20061731 Published online by Cambridge University Press

and a subsequent linear regression of principal components on predictors, although it is somewhat more sophisticated and efficient than this two-step procedure. As with principal component analysis, the score for each pattern can be computed by combining the standardised food variables, either with weights that are proportionate to their component loadings or with equal weights when restricting the food items to those that showed high loadings.

Dietary patterns and risk of coronary heart disease and stroke

Only one randomised trial has examined the effect of a dietary pattern on CVD, conducted among patients with existing myocardial infarction. In the Lyon Diet Heart study, de Lorgeril *et al.* (1999) found that a Mediterranean-type diet rich in α-linolenic acid reduced the rate of recurrence after a first myocardial infarction over a period of 46 months. In comparison with the control group, the Mediterranean-type diet group had a relative risk (RR) of cardiac deaths and non-fatal myocardial infarctions of 0·28 (95 % CI 0·15, 0·83).

Most evidence for a relationship between dietary patterns and CVD risk comes from observational studies. A growing number of studies, many carried out in large and well-described prospective cohorts, have been published in recent years (Table 1). Some studies used cluster analysis to derive dietary patterns. In the Italian centres of the Seven Countries study (Farchi *et al.* 1989), mortality rate was highest among those participants with a low-carbohydrate and high-alcohol, high-energy diet (high consumption of alcohol, low consumption of fruit, cakes, and meat compared with other clusters) but was lowest with a diet high in polyunsaturated fat (high consumption of seed oils, low consumption of olive oil, processed meat, fish, eggs, and vegetables compared with other clusters). However, mortality rates were adjusted for age and geographical area only, giving the possibility that differences in other characteristics between clusters may explain this finding. In the Cardiovascular Health study (Diehr & Beresford, 2003), participants with low energy, carbohydrate, fibre, and protein intake had a significantly lower risk of developing angina pectoris or myocardial infarction compared with participants who followed a ‘healthy’ diet (characterised by low fat and high fibre and carbohydrate intakes). Millen *et al.* (2004) compared women in a ‘heart healthy’ diet cluster, characterised by high intakes of fruit, vegetables, lean protein sources, whole grains, low-fat dairy products and lower fats, with the rest of women in the Framingham Offspring/Spouse study, but could not observe a significant difference in the incidence of subclinical heart disease.

Factor analysis was applied in the ‘Monitoring Trends and Determinants in Cardiovascular Disease’ (MONICA) Denmark study to empirically derive patterns (Osler *et al.* 2001, 2002). While a lower CVD mortality was observed among women with high scores at a ‘prudent’ pattern (RR for 1 SD increase: 0·63; 95 % CI 0·44, 0·90), which represented a diet rich in wholemeal bread, pasta, rice, oatmeal products, fruit, vegetables, fish, and low in white bread, it was not associated with CVD mortality among men or with CHD risk (fatal and non-fatal) among both men and women. In the US-based Nurses’ Health study and Health Professionals Follow-up study (HPFS), two dietary patterns derived with factor analysis have been evaluated with regard to CHD and stroke risk (Hu *et al.* 2000; Fung *et al.* 2001*b*, 2004*b*). In both studies, a ‘prudent’ pattern, rich in vegetables, fruit, legumes, whole grains, fish and poultry, significantly lowered CHD risk. In the Health Professionals Follow-up study, the RR comparing the highest with the lowest quintile of pattern score was 0·70 (95 % CI 0·56, 0·86), while the corresponding RR in the Nurses’ Health study was 0·76 (95 % CI 0·60, 0·98). In contrast, a ‘Western’ diet pattern (high in red meat, processed meat, refined grains, sweets and desserts, French fries, and high-fat dairy products) found in both studies was directly associated with CHD incidence. The RR comparing the highest with the lowest quintile of pattern score was 1·64 (95 % CI 1·24, 2·17) in the Health Professionals Follow-up study and 1·46 (95 % CI 1·07, 1·99) in the Nurses’ Health study. While the ‘prudent’ pattern was associated with a non-significantly decreased risk of stroke in the Nurses’ Health study (RR 0·78; 95 % CI 0·61, 1·01), the ‘Western’ pattern was related to a significantly increased risk (RR 1·58; 95 % CI 1·15, 2·15).

In the same study populations, several hypothesis-oriented patterns were evaluated with regard to CHD and stroke risk. Stampfer *et al.* (2000) demonstrated that a pattern score reflecting a diet low in *trans* fat and glycaemic load, and high in cereal fibre, marine *n*-3 fatty acids and folate and with a high polyunsaturated:saturated fat ratio strongly predicted the risk of CHD in the Nurses’ Health study. Women in the lowest quintile of the pattern score had a RR of CHD of 1·90 (95 % CI 1·55, 3·34). The Healthy Eating Index, a measure of how well diets conform to the US Department of Agriculture Dietary Guidelines for Americans (Kennedy *et al*. 1995), was associated with a modest reduction in CVD risk (McCullough *et al.* 2000*a*,*b*). The Healthy Eating Index consists of ten equally weighted components measuring adherence to serving recommendations for grains, vegetables, fruit, milk, and meat, as well as measuring intake of total fat, saturated fat, cholesterol, Na, and diet diversity. The RR of CVD comparing the highest with the lowest quintile of pattern score was 0·86 (95 % CI 0·72, 1·03) in the Nurses’ Health study and 0·72 (95 % CI 0·60, 0·88) in the Health Professionals Follow-up study. Modifying the index by incorporating specific goals for vegetables, fruit, nuts, red and white meat, cereal fibre, *trans* fat, PUFA:saturated fatty acid ratio, supplement use and alcohol intake resulted in a better prediction of CVD risk (McCullough *et al.* 2002). Participants in the highest quintile of the modified index called Alternate Healthy Eating Index had a RR of 0·72 (95 % CI 0·60, 0·86) in the Nurses’ Health study and 0·61 (95 % CI 0·49, 0·75) in the Health Professionals Follow-up study. In the same study, the Recommended Food Score, a summary measure of the consumption of vegetables, fruit, legumes, potatoes, poultry, fish, whole grains, and low-fat dairy products, was associated with reduced CVD risk among men in the Health Professionals Follow-up study (RR 0·77; 95 % CI 0·64, 0·93), but not among women in the Nurses’ Health study. The Recommended Food Score was evaluated in two further studies, where it was associated with decreased CHD and stroke morbidity (Michels & Wolk, 2002) and mortality (Kant *et al.* 2000). It is important to note here that the Recommended Food Score emphasises specific foods, not food groups, and is heavily contributed to by fruits and vegetables. Mediterranean-style diets have particularly drawn attention with

Table 1. Prospective cohort studies on dietary patterns and risk of coronary heart disease and stroke

Study	Study population	Dietary assessment method or pattern method	Main findings
Seven Countries study – Italy (Farchi *et al.* 1989)	1536 men; age 45–64 years; follow-up 20 years; events: 168 incident cases of CHD death and eighty-nine stroke deaths	Dietary history; four clusters from cluster analysis based on nutrient densities (carbohydrates, protein, alcohol, PUFA, MUFA, SFA) and total energy intake: cluster 1 (high alcohol and energy, low carbohydrate), cluster 2 (high PUFA), cluster 3 (high MUFA and SFA, low energy and alcohol) and cluster 4 (high carbohydrate and protein, low alcohol, SFA, and MUFA)	Highest mortality rates in cluster 1 and lowest in cluster 2. Age-adjusted mortality rates for clusters 1, 2, 3, and 4: CHD 10 years follow-up: 5·3, 0·6, 1·7, 3·5 ($P<0·05$) 15 years follow-up: 9·4, 2·0, 4·2, 7·5 ($P<0·05$) 20 years follow-up: 14·0, 10·2, 11·3, 13·4 (NS) Stroke 10 years follow-up: 3·3, 0·6, 0·9, 1·2 ($P<0·005$) 15 years follow-up: 5·0, 3·4, 5·9, 2·6 ($P<0·05$) 20 years follow-up: 8·7, 5·4, 7·4, 5·5 ($P<0·05$)
Health Professionals Follow-up study (Hu *et al.* 2000)	44 875 men; age 40–75 years; follow-up 8 years; events: 1089 incident CHD cases	FFQ with 131 items; two factors from factor analysis of forty food groups: 'prudent' (vegetables, fruit, legumes, whole grains, fish, and poultry) and 'Western' (red meat, processed meat, refined grains, sweets and desserts, French fries, and high-fat dairy products)	Decrease in CHD incidence with 'prudent' pattern, increased with 'Western' pattern. Multivariate-adjusted RR for extreme quintiles: 'Prudent': 0·70 (95 % CI 0·56, 0·86) 'Western': 1·64 (95 % CI 1·24, 2·17)
Breast Cancer Detection Demonstration project (Kant *et al.* 2000)	42 254 women; age 40–93 years; follow-up 5·6 years (median); events: 274 CHD deaths, 130 stroke deaths	FFQ with sixty-two items; recommended food score based on twenty-three food items (fruit, vegetables, whole grains, lean meats or meat alternates, and low-fat dairy)	Decrease in CHD and stroke mortality with higher recommended food score. RR for extreme quartiles: CHD: 0·67 (95 % CI 0·47, 0·95) Stroke: 0·58 (95 % CI 0·35, 0·96)
Health Professionals Follow-up study (McCullough *et al.* 2000*a*)	38 622 men; age 40–75 years; follow-up 8 years; events: 1092 incident cases of CVD (fatal and non-fatal MI and stroke)	FFQ with 131 items; healthy eating index based on ten dietary recommendations (specific goals for grains, vegetables, fruit, milk, meat, total fat, saturated fat, cholesterol, Na, variety)	Reduced CVD risk with higher healthy eating index. Multivariate-adjusted RR for extreme quintiles: 0·72 (95 % CI 0·60, 0·88); $P<0·001$ for trend
Nurses' Health study (McCullough *et al.* 2000*b*)	67 272 women; age 30–55 years; follow-up 12 years; events: 1365 incident cases of CVD (fatal and non-fatal MI and stroke)	FFQ with 116 items; healthy eating index based on ten dietary recommendations (specific goals for grains, vegetables, fruit, milk, meat, total fat, saturated fat, cholesterol, Na, variety)	Higher healthy eating index not significantly associated with risk. Multivariate-adjusted RR for extreme quintiles: 0·86 (95 % CI 0·72, 1·03); $P=0·085$ for trend
Nurses' Health study (Stampfer *et al.* 2000)	84 129 women; age 30–55 years; follow-up 14 years; events: 1128 incident cases of CHD (fatal CHD and non-fatal MI)	FFQ with 61–116 items; diet score based on six nutrients (low in *trans* fat and glycaemic load, high in cereal fibre, marine *n*-3 fatty acids, and folate, and with a high PUFA:SFA ratio)	Increased CHD risk with low diet score. Multivariate-adjusted RR for extreme quintiles: 1·90 (95 % CI 1·55, 2·34)
Nurses' Health study (Fung *et al.* 2001*b*)	69 017 women; age 38–63 years; follow-up 12 years; events: 821 incident cases of CHD (fatal CHD and non-fatal MI)	FFQ with 116 items; two factors from factor analysis of thirty-eight food groups: 'prudent' pattern (fruit, vegetables, whole grains, legumes, poultry, and fish) and 'Western' pattern (refined grains, red and processed meat, desserts, high-fat dairy products, and French fries)	Significantly lower risk of CHD with 'prudent' pattern, higher risk with 'Western' pattern. Multivariate-adjusted RR for extreme quintiles: 'Prudent': 0·76 (95 % CI 0·60, 0·98) 'Western': 1·46 (95 % CI 1·07, 1·99)
MONICA Denmark (Osler *et al.* 2001)	5872 men and women; age 30–60 years; follow-up 15 years (median); events: 147 CVD deaths	FFQ with twenty-eight items; two factors from factor analysis of twenty-eight food groups: 'prudent' pattern (directly associated with wholemeal bread, pasta, rice, oatmeal products, fruit, vegetables, and fish, inversely with other types of bread) and 'Western' (directly associated with meat, sausages, potatoes, butter, and white bread); healthy food index based on four items (specific intakes of butter, lard and margarine, vegetables, coarse bread, and fruit)	Lower cardiovascular mortality with the 'prudent' pattern in women. No significant associations between 'Western' pattern and healthy food index and CVD mortality in men and women and between 'prudent' pattern and CVD mortality in men. Multivariate-adjusted RR: Healthy food index (1 point) Men: 0·98 (95 % CI 0·82, 1·19) Women: 0·92 (95 % CI 0·68, 1·24) 'Prudent' (1 SD) Men: 0·87 (95 % CI 0·71, 1·06) Women: 0·63 (95 % CI 0·44, 0·90) 'Western' (1 SD) Men: 1·06 (95 % CI 0·86, 1·32) Women: 0·92 (95 % CI 0·69, 1·27)

Table 1. *Continued*

Study	Study population	Dietary assessment method or pattern method	Main findings
MONICA Denmark (Osler *et al.* 2002)	7316 men and women; age 30–60 years; follow-up up to 14 years; events: 280 incident cases of CHD (fatal and non-fatal)	FFQ with twenty-eight items; two factors from factor analysis of twenty-eight food groups: 'prudent' pattern (directly associated with wholemeal bread, pasta, rice, oatmeal products, fruit, vegetables, and fish, inversely with other types of bread) and 'Western' (directly associated with meat, sausages, potatoes, butter, and white bread); healthy food index based on four items (specific intakes of butter, lard and margarine, vegetables, coarse bread, and fruit)	No significant associations between patterns and CHD risk. Multivariate-adjusted RR: Healthy food index (3 or 4 points *v.* 0 points) 1·21 (95 % CI 0·80, 1·82) 'Prudent' (1 SD) 1·06 (95 % CI 0·93, 1·21) 'Western' (1 SD) 0·97 (95 % CI 0·85, 1·10)
Health Professionals Follow-up study and Nurses' Health study (McCullough *et al.* 2002)	38 615 men and 67 271 women; age 40–75 years (men) and 30–55 years (women); follow-up 8–12 years; events: 1092 incident cases of CVD (fatal CHD, non-fatal MI and stroke) in men, 1365 in women	FFQ with 131 and 116 items; alternate healthy eating index (specific goals for vegetables, fruit, nuts and soya protein, white meat:red meat ratio, cereal fibre, *trans* fat, PUFA:SFA ratio, duration of vitamin supplement use, and alcohol); recommended food score (vegetables, fruit, legumes, potatoes, poultry, fish, whole grains, low-fat dairy products)	Reduced risk of CVD among men and women with higher alternate healthy eating index and among men with higher recommended food score. Multivariate-adjusted RR comparing extreme quintiles: Alternate healthy eating index Men: 0·61 (95 % CI 0·49, 0·75) Women: 0·72 (95 % CI 0·60, 0·86) Recommended food score Men: 0·77 (95 % CI 0·64, 0·93) Women: 0·90 (95 % CI 0·75, 1·08)
Mammography Screening Cohort Sweden (Michels & Wolk, 2002)	59 038 women; age 40–76 years; follow-up 4 years; events: 779 incident cases of CHD and 342 cases of stroke	FFQ with sixty items; 'recommended food score' (fruit, vegetables, legumes, low-fat dairy products, whole grains, and fish) and 'not recommended food score' (meat, processed meat, fried potatoes, French fries, chips, butter, cheese, margarine, white bread, cookies, sugar, sweets)	Reduced risk of CHD and stroke mortality with high 'recommended food score'. No significant association between 'not recommended food score' and mortality. Multivariate-adjusted RR for extreme quintiles: Recommended food score CHD: 0·47 (95 % CI 0·33, 0·68) Stroke: 0·40 (95 % CI 0·22, 0·73) Not recommended food score CHD: 0·79 (95 % CI 0·47, 1·32) Stroke: 0·96 (95 % CI 0·47, 1·97)
Cardiovascular Health study (Diehr & Beresford 2003)	5888 men and women; age $>$ 65 years (mean 73 years); follow-up 10 years in cohort 1 (*n* 5201), 7 years in cohort 2 (*n* 687); events: 1398 incident cases of angina, 777 MI, 764 stroke and 1209 congestive heart failure	FFQ with ninety-nine items; cluster analysis based on carbohydrates, fat, protein, fibre, and energy; five dietary clusters: 'unhealthy' (high fat and protein, low fibre and carbohydrates), 'high cal' (high energy), 'low cal' (low energy), 'low' (low energy, fibre, protein, and carbohydrates, high fat), and 'healthy' (low fat, high fibre and carbohydrates)	'Low' cluster significantly ($P<0·05$) associated with lower risk of CHD events compared with 'healthy' cluster: Angina: 21·26 *v.* 24·85 % MI: 11·71 *v.* 13·64 %
EPIC Greece (Trichopoulou *et al.* 2003)	22 043 men and women; age 20–86 years; follow-up 3·7 years (mean); events: fifty-four incident cases of CHD death	FFQ with 150 items; Mediterranean-diet score (specific goals for vegetables, legumes, fruit and nuts, cereals, fish, meat, poultry, dairy products, alcohol, and MUFA:SFA ratio)	Lower risk of CHD mortality with higher degree of adherence to Mediterranean diet score. Multivariate-adjusted RR for a two-point increase in score: 0·67 (95 % CI 0·47, 0·94)
Nurses' Health study (Fung *et al.* 2004*b*)	71 768 women; age 38–63 years; follow-up 14 years; events: 791 cases of incident stroke	FFQ with 116 items; two factors from factor analysis of thirty-six to forty-two food groups: 'prudent' pattern (fruit, vegetables, whole grains, fish, and poultry) and 'Western' pattern (red and processed meat, refined grains, high-fat dairy products, and desserts and sweets)	No significant association between 'prudent' pattern and stroke risk, increased stroke risk with higher 'Western' pattern score. Multivariate-adjusted RR for extreme quintiles: 'Prudent': 0·78 (95 % CI 0·61, 1·01) 'Western': 1·58 (95 % CI 1·15, 2·15)
Healthy Ageing: a Longitudinal study in Europe (HALE) (Knoops *et al.* 2004)	1507 men and 832 women; age 70–90 years; follow-up 10 years; events: 371 deaths from CVD	Dietary history; modified Mediterranean-diet score (monounsaturated:saturated fat ratio, legumes, nuts, and seeds, grains, fruit, vegetables and potatoes, meat and meat products, dairy products, and fish)	Mediterranean diet associated with lower CHD and CVD mortality rates. RR comparing subjects with $\geq$ 4 points on eight-point scale with those with $<$ 4 points: CHD mortality: 0·61 (95 % CI 0·43, 0·88) CVD mortality: 0·71 (95 % CI 0·58, 0·88)

Table 1. *Continued*

Study	Study population	Dietary assessment method or pattern method	Main findings
Framingham Offspring/Spouse study (Millen *et al.* 2004)	1423 women; age 18–76 years; follow-up 12 years; events: 144 incident cases of subclinical heart disease assessed by carotid artery stenosis	FFQ; five clusters from cluster analysis based on food consumption: 'heart healthy' (fruit, vegetables, leaner protein sources, whole grains, lower-fat dairy products, fewer dietary fats) compared with all other clusters	No significant association between 'heart healthy' cluster and CHD risk. Multivariate adjusted RR: 0·74 (95 % CI 0·45, 1·17)
EPIC Potsdam (Weikert *et al.* 2005)	27 548 men and women; age 35–65 years; follow-up 4·6 years (mean); events: 157 incident cases of MI	FFQ with 148 items; dietary pattern from reduced rank regression based on food groups as predictors of plasma folate, B_{12}, homocysteine concentrations (pattern was initially derived in an unrelated case–control study and was related to high plasma folate, high B_{12}, and low homocysteine): dietary pattern with high intake of mushrooms, olive oil, fresh fruit, wine, cruciferous vegetables, wholegrain bread, nuts, low intake of fried potatoes	Reduced CHD risk with higher pattern score. Multivariate-adjusted RR for extreme quintiles: 0·72 (95 % CI 0·43, 1·20); P=0·041 for trend

SFA, saturated fatty acids; FFQ, food-frequency questionnaire; RR, relative risk; MI, myocardial infarction; MONICA, Monitoring Trends and Determinants in Cardiovascular Disease; EPIC, European Prospective Investigation into Cancer and Nutrition.

regard to the prevention of CVD. Trichopoulou *et al.* (2003) evaluated adherence to a Mediterranean diet score with regard to overall as well as CVD mortality in the Greek part of the European Prospective Investigation into Cancer and Nutrition (EPIC) study. A high diet score reflected a diet rich in vegetables, legumes, fruit and nuts, cereals, and fish and low in meat, poultry, and dairy products, with a moderate alcohol intake, and a high MUFA:saturated fatty acid ratio. A 2-point increase in diet score was associated with a RR of CVD death of 0·67 (95 % CI 0·47, 0·94). Evaluating the same pattern score in the Healthy Ageing: a Longitudinal study in Europe (HALE) project, Knoops *et al.* (2004) observed also lower CHD and CVD mortality among participants who adhered more strongly to the Mediterranean diet.

Just recently, we derived a dietary pattern with reduced rank regression using plasma folate, vitamin B_{12} and homocysteine concentrations as responses (Weikert *et al.* 2005). The pattern reflected a diet rich in mushrooms, olive oil, fresh fruit, wine, vegetables, wholegrain bread, and nuts and low in fried potatoes and was associated with higher plasma folate and vitamin B_{12} concentrations but lower homocysteine. We observed a moderately reduced risk of myocardial infarction with higher pattern score in the prospective EPIC-Potsdam study (RR for extreme quintiles: 0·72; 95 % CI 0·43, 1·20; P=0·041 for trend).

Biological mechanisms linking dietary patterns to coronary heart disease and stroke

Several potential biological mechanisms may explain how overall dietary patterns are related to the risk of CHD and stroke, among them effects on blood pressure, blood lipid levels, blood homocysteine concentrations, oxidative stress, endothelial function, inflammation, and insulin sensitivity (Hu & Willett, 2002). Evidence for these relationships has been accumulated over the last years from several lines: clinical trials, cohort studies, as well as cross-sectional studies. That plant-based dietary patterns have favourable effects on blood pressure has been convincingly demonstrated by the Dietary Approaches to Stop Hypertension (DASH) trial (Appel *et al.* 1997), trials on vegetarian diets (Rouse *et al.* 1983; Margetts *et al.* 1986), and a prospective cohort (Schulze *et al.* 2003*b*). Furthermore, in a recent trial over a 2-year period among men and women with the metabolic syndrome, increased consumption of fruit, vegetables, walnuts, whole grains, and olive oil significantly reduced concentrations of C-reactive protein (CRP), IL-6, IL-7, and IL-18, reduced insulin resistance, as well as improved endothelial function compared with the control group which consumed an otherwise healthy diet (< 30 % fat, < 10 % saturated fat; Esposito *et al.* 2004). These effects were attenuated but not eliminated by additional adjustment for weight change over the course of the study. That overall dietary patterns derived using factor analysis predict risk of diabetes, a major risk factor for CHD and stroke, has been reported by several cohort studies, where specifically 'Western'-type diets were associated with an increased risk (van Dam *et al.* 2002; Fung *et al.* 2004*a*; Montonen *et al.* 2005). We have recently identified a dietary pattern with reduced rank regression in the prospective EPIC-Potsdam cohort which was characterised by a high intake of fresh fruit and a low intake of sugar-sweetened beverages, beer, meat, poultry, processed meat, legumes, and bread excluding wholegrain bread (Heidemann *et al.* 2005). Subjects who scored high had high plasma concentrations of HDL-cholesterol and adiponectin and low plasma concentrations of HbA_{1c} and had a lower risk to develop type 2 diabetes. Also using reduced rank regression, a dietary pattern high in sugar-sweetened soft drinks, refined grains, diet soft drinks, and processed meat, and low in wine, coffee, cruciferous vegetables, and yellow vegetables was related to higher levels of markers of inflammation and endothelial dysfunction and to a higher risk of type 2 diabetes in the Nurses' Health study (Schulze *et al.* 2005).

There is further evidence from cross-sectional studies on dietary patterns and risk markers of CHD and stroke. We were able to identify a dietary pattern with reduced rank regression among women in the German 'Coronary Risk Factors for Atherosclerosis in Women' (CORA) study that was associated with lower HDL-cholesterol levels, but higher C-peptide and CRP levels. The dietary pattern was characterised by high intakes of meat, margarine, poultry, and sauce and low intakes of vegetarian dishes, wine, vegetables, and wholegrain cereals (Hoffmann *et al.* 2004*b*). The US-based 'prudent' pattern was inversely associated with plasma concentrations of CRP and E-selectin and the 'Western' pattern showed a positive relationship with CRP, E-selectin, soluble intercellular adhesion molecule-1 and vascular cell adhesion molecule-1 after adjustment for age, BMI, physical activity, smoking status, and alcohol consumption in the Nurses' Health study (Lopez-Garcia *et al.* 2004). In addition, the 'Western' pattern was significantly correlated with fasting insulin, C-peptide, and homocysteine and inversely with folate levels in the Health Professionals Follow-up study after adjustment for a variety of risk factors including BMI (Fung *et al.* 2001*a*). In contrast, the 'prudent' pattern was associated with higher folate, but lower fasting insulin and homocysteine concentrations in this study. Somewhat similar patterns were observed in the third National Health and Nutrition Examination Survey (NHANES-III), where the 'Western' pattern (high intakes of processed meat, eggs, red meat, and high-fat dairy products) was directly associated with C-peptide, insulin and HbA_{1c} and inversely with folate concentrations after adjustment for confounding variables, while the 'American-healthy' pattern (high intakes of vegetables, salad dressings, and tea) was not associated with any of the CVD risk markers examined (Kerver *et al.* 2003). Van Dam *et al.* (2003) identified a 'cosmopolitan' pattern (greater intakes of fried vegetables, salad, rice, chicken, fish, and wine), a 'traditional' pattern (greater intakes of red meat and potatoes and lesser intakes of low-fat dairy and fruit), and a 'refined-foods' pattern (greater intakes of French fries, high-sugar beverages, and white bread and lesser intakes of wholegrain bread and boiled vegetables) in the 'Monitoring Project on Risk Factors and Chronic Diseases in the Netherlands' (MORGEN) study. Independent of other lifestyle factors and BMI, a high 'cosmopolitan' pattern score was significantly associated with lower blood pressure and higher HDL-cholesterol concentrations while a high 'traditional' pattern score was associated with higher blood pressure and higher concentrations of HDL-cholesterol, total cholesterol, and glucose. The 'refined-foods' pattern was directly associated with total cholesterol concentrations. The Alternate Healthy Eating Index, the Recommended Food Score, and the alternate Mediterranean Diet Index (all described earlier) were inversely associated with levels of inflammatory markers independent of lifestyle characteristics and BMI in the Nurses' Health study (Fung *et al.* 2005*b*), while the Healthy Eating Index and the Diet Quality Index Revised were not. In the Italian Risk Factors and Life Expectancy (RIFLE) project, the Food Pyramid Index reflecting the ratio of fatty foods (oils, fats, sweets, dairy products, meat, poultry, and fish) and low-fat foods (pasta, rice, bread, and legumes) was directly related to BMI, blood pressure, total cholesterol and glucose levels (Massari *et al.* 2004).

Discussion

We have identified several prospective studies that evaluated dietary patterns with regard to CHD or stroke risk. Overall, plant-based dietary patterns which are rich in fruit, vegetables, and whole grains, but low in meat and refined grains and that focus on healthy sources of fat appear to be useful for preventing CHD and stroke. This evidence is facilitated by one randomised secondary prevention trial as well as studies relating dietary patterns to CHD risk markers. Some of the studies reviewed in the present paper have been reviewed elsewhere (Hu, 2002; Schulze & Hu, 2002; Kant, 2004; Newby & Tucker, 2004). However, earlier reviews by Hu (2002) and Schulze & Hu (2002) were limited to studies published before 2002, missing a large number of studies published within the last few years (McCullough *et al.* 2002; Michels & Wolk, 2002; Osler *et al.* 2002; Diehr & Beresford, 2003; Trichopoulou *et al.* 2003; Fung *et al.* 2004*b*; Knoops *et al.* 2004; Millen *et al.* 2004; Weikert *et al.* 2005). The more recent reviews by Kant (2004) and Newby & Tucker (2004) were not focused on CHD and stroke, and were in part limited to exploratory patterns (Newby & Tucker, 2004). From the seventeen prospective observational studies which we identified, eight studies (Farchi *et al.* 1989; Stampfer *et al.* 2000; Diehr & Beresford, 2003; Trichopoulou *et al.* 2003; Fung *et al.* 2004*b*; Knoops *et al.* 2004; Millen *et al.* 2004; Weikert *et al.* 2005) have not been reviewed by Kant (2004) and nine studies (McCullough *et al.* 2000*a*,*b*, 2002; Stampfer *et al.* 2000; Michels & Wolk, 2002; Trichopoulou *et al.* 2003; Fung *et al.* 2004*b*; Knoops *et al.* 2004; Weikert *et al.* 2005) not by Newby & Tucker (2004).

Several limitations apply to pattern approaches in general. Correlated measurement error in assessing foods may distort definition of a pattern. This would lead to an overestimation of correlation between foods, and phenomena such as over-reporting of 'healthy' foods may selectively misclassify study individuals. Furthermore, there is not much information yet on whether the applied food grouping influences retained patterns and their subsequent risk estimates (McCann *et al.* 2001). Another question that has received little attention in pattern analysis is whether patterns are reproducible within the same study population. Hu *et al.* (1999) evaluated whether factor analysis-based patterns are reproducible in the Health Professionals Follow-up study and found reliability correlations for the factor scores between two food-frequency questionnaires of 0·70 for a 'prudent' pattern and of 0·67 for a 'Western' pattern. Similar reliability correlations were observed for the Diet Quality Index Revised in this cohort (Newby *et al.* 2003) and for factor analysis-based patterns in the Swedish Mammography Cohort (Khani *et al.* 2004). In randomly generated split-half samples, Newby *et al.* (2004) found clusters to be more accurately reproducible than factors. Similarly, few studies have evaluated the validity of dietary patterns based on dietary data from food-frequency questionnaires compared with other dietary assessment methods. Correlations between food-frequency questionnaire-based patterns and those based on dietary records ranged from 0·45 to 0·85, suggesting a reasonable validity (Hu *et al.* 1999; Newby *et al.* 2003; Khani *et al.* 2004).

Exploratory approaches to define dietary patterns involve arbitrary choices, such as the number of food groups to be included and the number of factors or clusters to be retained (Martinez *et al.* 1998). Although exploratory dietary patterns represent real-world dietary behaviours that may be of interest by themselves, they often seem to account only for a small or moderate proportion of total variance of foods (Schulze *et al.* 2001) and for a tiny proportion of disease-relevant variation of dietary intake. The derived patterns are therefore unlikely to be appropriate to predict a specific disease. Even if a pattern has been shown to be associated with a particular disease, it does not suggest that the observed pattern is the most beneficial or most detrimental. Another concern is that exploratory patterns are by nature population-specific, and results of aetiological studies may not be reproducible in other populations (Schulze & Hu, 2002). Moreover, it is possible that either only certain components trigger the beneficial or adverse health effect of the pattern or that only the full combination of food components does. Although dietary patterns might share important components across different populations, for example red meat for similarly labelled ‘Western’ patterns in the USA (Hu *et al.* 2000; Kerver *et al.* 2003) and Europe (Osler *et al.* 2002), it is therefore not straightforward to assume similar health effects. Application of exploratory pattern analysis might also not be appropriate in situations where the effect is caused by one specific nutrient or food, since their effect will most probably be diluted. For example, only fruit and vegetables were significantly associated with risk of postmenopausal breast cancer in the Nurses’ Health study, but this association was diluted evaluating dietary patterns: the ‘prudent’ pattern, which contained fruit and vegetables as components, was not significantly associated with risk (Fung *et al.* 2005*a*). Still, exploratory patterns can provide information for setting priorities for changing dietary habits in a population by public health initiatives (van Dam, 2005).

On the other hand, the hypothesis-oriented approach offers the possibility of constructing pattern variables based on existing scientific evidence. Whether hypothesis-driven patterns have greater effects on disease risk than exploratory patterns will largely depend on the strength of evidence the hypothesis-driven patterns are based on. Such dietary scores also require quite subjective decisions, for example, the identification of foods considered ‘recommended’ or the frequency of consumption deemed important; and their definition may vary substantially across investigators. In addition, hypothesis-oriented patterns may still depend on the data at hand. For example, although individual components of the Mediterranean pattern score are defined *a priori*, whether individuals score ‘0’ or ‘1’ at single components depend on their relative intake with sex-specific medians as cut-points for most components (Trichopoulou *et al.* 2003). Thus, the quantitative food intake of individuals considered to adhere to the Mediterranean pattern based on the scoring system may be quite different across different studies. Scores based on dietary recommendations are most often not disease-specific and can only be as good as the dietary guidelines. Food intake is the primarily recorded data in observational studies and is subsequently converted into nutrient intake by using weights for nutrient concentrations obtained from food tables. Thus, nutrient-based dietary patterns might be hampered by uncertainties and errors inherent to nutrient intake data (Leclercq *et al.* 2001).

The reduced rank regression approach requires response (for example, biomarker) information. This information may not be available in many studies otherwise suitable to evaluate diet–disease associations. Also, the biomarker information available may not reflect the current status of knowledge about pathways which may be relevant in the development of disease, hereof CHD and stroke. It is also unlikely that reduced rank regression is able to identify a dietary pattern that is linked to most or all potential pathways by which diet may influence CHD or stroke risk. For example, we were able to identify a dietary pattern with reduced rank regression that was related to lower HDL-cholesterol and higher C-peptide and CRP concentrations. But this pattern did not explain variance of LDL-cholesterol and lipoprotein(a) (Hoffmann *et al.* 2004*b*). Another concern is that associations between reduced rank regression patterns and CHD and stroke risk would need to be confirmed in independent studies, if the response biomarkers have been evaluated with regard to the endpoint in the same population (van Dam, 2005; Weikert *et al.* 2005). In addition, confounding may already distort the results of a study at the step of reduced rank regression, since food intake and CHD risk markers may both be related to lifestyle factors, anthropometric characteristics or existence and treatment of medical conditions (Schulze *et al.* 2005). Still, the use of biomarkers of disease as response variables in reduced rank regression is of scientific interest because it may not only help to identify dietary patterns with relevance for CHD and stroke, but can simultaneously aid the pathophysiological interpretation of observed associations.

Conclusions

Dietary pattern analysis is an analytical approach that allows for and reflects the complexity of dietary exposure. This approach might be useful, especially if many dietary components are relevant for a disease, as it accounts for latent interactions and cumulative effects. This may be particularly important for CHD and stroke for which multiple dietary components are established or putative risk factors. There is a large body of evidence, mainly from prospective observational studies, that plant-based dietary patterns which are rich in fruit, vegetables, and whole grains, but low in meat and refined grains and that focus on healthy sources of fat are useful for preventing CHD and stroke. This evidence is facilitated by studies relating dietary patterns to CHD risk markers. Reduced rank regression as a new tool in dietary pattern analysis may be useful in this context to identify dietary patterns relevant to CHD and stroke and also to aid their pathophysiological interpretation.

References

Appel LJ, Moore TJ, Obarzanek E, *et al.* (1997) A clinical trial of the effects of dietary patterns on blood pressure. DASH Collaborative Research Group. *N Engl J Med* **336**, 1117–1124.

Bentler PM (1980) Mutivariate analysis with latent variables: causal modeling. *Annu Rev Psychol* **31**, 11–21.

Bland JM & Altman DG (1997) Cronbach's alpha. *BMJ* **314**, 572.

de Lorgeril M, Salen P, Martin JL, Monjaud I, Delaye J & Mamelle N (1999) Mediterranean diet, traditional risk factors, and the rate of cardiovascular complications after myocardial infarction: final report of the Lyon Diet Heart study. *Circulation* **99**, 779–785.

Diehr P & Beresford SA (2003) The relation of dietary patterns to future survival, health, and cardiovascular events in older adults. *J Clin Epidemiol* **56**, 1224–1235.

Esposito K, Marfella R, Ciotola M, Di Palo C, Giugliano F, Giugliano G, D'Armiento M, D'Andrea F & Giugliano D (2004) Effect of a Mediterranean-style diet on endothelial dysfunction and markers of vascular inflammation in the metabolic syndrome: a randomized trial. *JAMA* **292**, 1440–1446.

Farchi G, Mariotti S, Menotti A, Seccareccia F, Torsello S & Fidanza F (1989) Diet and 20-y mortality in two rural population groups of middle-aged men in Italy. *Am J Clin Nutr* **50**, 1095–1103.

Fung TT, Hu FB, Holmes MD, Rosner BA, Hunter DJ, Colditz GA & Willett WC (2005*a*) Dietary patterns and the risk of postmenopausal breast cancer. *Int J Cancer* **116**, 116–121.

Fung TT, McCullough ML, Newby P, Manson JE, Meigs JB, Rifai N, Willett WC & Hu FB (2005*b*) Diet-quality scores and plasma concentrations of markers of inflammation and endothelial dysfunction. *Am J Clin Nutr* **82**, 163–173.

Fung TT, Rimm EB, Spiegelman D, Rifai N, Tofler GH, Willett WC & Hu FB (2001*a*) Association between dietary patterns and plasma biomarkers of obesity and cardiovascular disease risk. *Am J Clin Nutr* **73**, 61–67.

Fung TT, Schulze M, Manson JE, Willett WC & Hu FB (2004*a*) Dietary patterns, meat intake, and the risk of type 2 diabetes in women. *Arch Intern Med* **164**, 2235–2240.

Fung TT, Stampfer MJ, Manson JE, Rexrode KM, Willett WC & Hu FB (2004*b*) Prospective study of major dietary patterns and stroke risk in women. *Stroke* **35**, 2014–2019.

Fung TT, Willett WC, Stampfer MJ, Manson JE & Hu FB (2001*b*) Dietary patterns and the risk of coronary heart disease in women. *Arch Intern Med* **161**, 1857–1862.

Gittelsohn J, Wolever TM, Harris SB, Harris-Giraldo R, Hanley AJ & Zinman B (1998) Specific patterns of food consumption and preparation are associated with diabetes and obesity in a Native Canadian community. *J Nutr* **128**, 541–547.

Heidemann C, Hoffmann K, Spranger J, Klipstein-Grobusch K, Mohlig M, Pfeiffer AF & Boeing H (2005) A dietary pattern protective against type 2 diabetes in the European Prospective Investigation into Cancer and Nutrition (EPIC) – Potsdam study cohort. *Diabetologia* **48**, 1126–1134.

Hoffmann K, Schulze MB, Schienkiewitz A, Nothlings U & Boeing H (2004*a*) Application of a new statistical method to derive dietary patterns in nutritional epidemiology. *Am J Epidemiol* **159**, 935–944.

Hoffmann K, Zyriax BC, Boeing H & Windler E (2004*b*) A dietary pattern derived to explain biomarker variation is strongly associated with the risk of coronary artery disease. *Am J Clin Nutr* **80**, 633–640.

Hu FB (2002) Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* **13**, 3–9.

Hu FB, Rimm E, Smith-Warner SA, Feskanich D, Stampfer MJ, Ascherio A, Sampson L & Willett WC (1999) Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *Am J Clin Nutr* **69**, 243–249.

Hu FB, Rimm EB, Stampfer MJ, Ascherio A, Spiegelman D & Willett WC (2000) Prospective study of major dietary patterns and risk of coronary heart disease in men. *Am J Clin Nutr* **72**, 912–921.

Hu FB & Willett WC (2002) Optimal diets for prevention of coronary heart disease. *JAMA* **288**, 2569–2578.

Jacobs DR Jr & Steffen LM (2003) Nutrients, foods, and dietary patterns as exposures in research: a framework for food synergy. *Am J Clin Nutr* Suppl. **3**, 78, 508S–513S.

Jacques PF & Tucker KL (2001) Are dietary patterns useful for understanding the role of diet in chronic disease? *Am J Clin Nutr* **73**, 1–2.

Kant AK (2004) Dietary patterns and health outcomes. *J Am Diet Assoc* **104**, 615–635.

Kant AK, Schatzkin A, Graubard BI & Schairer C (2000) A prospective study of diet quality and mortality in women. *JAMA* **283**, 2109–2115.

Kennedy ET, Ohls J, Carlson S & Fleming K (1995) The Healthy Eating Index: design and applications. *J Am Diet Assoc* **95**, 1103–1108.

Kerver JM, Yang EJ, Bianchi L & Song WO (2003) Dietary patterns associated with risk factors for cardiovascular disease in healthy US adults. *Am J Clin Nutr* **78**, 1103–1110.

Khani BR, Ye W, Terry P & Wolk A (2004) Reproducibility and validity of major dietary patterns among Swedish women assessed with a food-frequency questionnaire. *J Nutr* **134**, 1541–1545.

Knoops KT, de Groot LC, Kromhout D, Perrin AE, Moreiras-Varela O, Menotti A & van Staveren WA (2004) Mediterranean diet, lifestyle factors, and 10-year mortality in elderly European men and women: the HALE project. *JAMA* **292**, 1433–1439.

Leclercq C, Valsta LM & Turrini A (2001) Food composition issues – implications for the development of food-based dietary guidelines. *Public Health Nutr* **4**, 677–682.

Lopez-Garcia E, Schulze MB, Fung TT, Meigs JB, Rifai N, Manson JE & Hu FB (2004) Major dietary patterns are related to plasma concentrations of markers of inflammation and endothelial dysfunction. *Am J Clin Nutr* **80**, 1029–1035.

McCann SE, Marshall JR, Brasure JR, Graham S & Freudenheim JL (2001) Analysis of patterns of food intake in nutritional epidemiology: food classification in principal components analysis and the subsequent impact on estimates for endometrial cancer. *Public Health Nutr* **4**, 989–997.

McCullough ML, Feskanich D, Rimm EB, Giovannucci EL, Ascherio A, Variyam JN, Spiegelman D, Stampfer MJ & Willett WC (2000*a*) Adherence to the Dietary Guidelines for Americans and risk of major chronic disease in men. *Am J Clin Nutr* **72**, 1223–1231.

McCullough ML, Feskanich D, Stampfer MJ, Giovannucci EL, Rimm EB, Hu FB, Spiegelman D, Hunter DJ, Colditz GA & Willett WC (2002) Diet quality and major chronic disease risk in men and women: moving toward improved dietary guidance. *Am J Clin Nutr* **76**, 1261–1271.

McCullough ML, Feskanich D, Stampfer MJ, Rosner BA, Hu FB, Hunter DJ, Variyam JN, Colditz GA & Willett WC (2000*b*) Adherence to the Dietary Guidelines for Americans and risk of major chronic disease in women. *Am J Clin Nutr* **72**, 1214–1222.

Margetts BM, Beilin LJ, Vandongen R & Armstrong BK (1986) Vegetarian diet in mild hypertension: a randomised controlled trial. *Br Med J (Clin Res Ed)* **293**, 1468–1471.

Martinez ME, Marshall JR & Sechrest L (1998) Invited commentary: Factor analysis and the search for objectivity. *Am J Epidemiol* **148**, 17–19.

Maskarinec G, Novotny R & Tasaki K (2000) Dietary patterns are associated with body mass index in multiethnic women. *J Nutr* **130**, 3068–3072.

Massari M, Freeman KM, Seccareccia F, Menotti A & Farchi G (2004) An index to measure the association between dietary patterns and coronary heart disease risk factors: findings from two Italian studies. *Prev Med* **39**, 841–847.

Michels KB & Wolk A (2002) A prospective study of variety of healthy foods and mortality in women. *Int J Epidemiol* **31**, 847–854.

Millen BE, Quatromoni PA, Nam BH, O'Horo CE, Polak JF, Wolf PA & D'Agostino RB (2004) Dietary patterns, smoking, and subclinical heart disease in women: opportunities for primary prevention from the Framingham Nutrition Studies. *J Am Diet Assoc* **104**, 208–214.

Montonen J, Knekt P, Harkanen T, Jarvinen R, Heliovaara M, Aromaa A & Reunanen A (2005) Dietary patterns and the incidence of type 2 diabetes. *Am J Epidemiol* **161**, 219–227.

Newby PK, Hu FB, Rimm EB, Smith-Warner SA, Feskanich D, Sampson L & Willett WC (2003) Reproducibility and validity of the Diet Quality Index Revised as assessed by use of a food-frequency questionnaire. *Am J Clin Nutr* **78**, 941–949.

Newby PK, Muller D & Tucker KL (2004) Associations of empirically derived eating patterns with plasma lipid biomarkers: a comparison of factor and cluster analysis methods. *Am J Clin Nutr* **80**, 759–767.

Newby PK & Tucker KL (2004) Empirically derived eating patterns using factor or cluster analysis: a review. *Nutr Rev* **62**, 177–203.

Osler M, Heitmann BL, Gerdes LU, Jorgensen LM & Schroll M (2001) Dietary patterns and mortality in Danish men and women: a prospective observational study. *Br J Nutr* **85**, 219–225.

Osler M, Helms Andreasen A, Heitmann B, Hoidrup S, Gerdes U, Morch Jorgensen L & Schroll M (2002) Food intake patterns and risk of coronary heart disease: a prospective cohort study examining the use of traditional scoring techniques. *Eur J Clin Nutr* **56**, 568–574.

Petersen S, Peto V, Rayner M, Leal J, Luengo-Fernandez R & Gray A (2005) *European Cardiovascular Disease Statistics*. London: British Heart Foundation.

Rouse IL, Beilin LJ, Armstrong BK & Vandongen R (1983) Blood-pressure-lowering effect of a vegetarian diet: controlled trial in normotensive subjects. *Lancet* **i**, 5–10.

Schulze MB, Hoffmann K, Kroke A & Boeing H (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.

Schulze MB, Hoffmann K, Kroke A & Boeing H (2003*a*) An approach to construct simplified measures of dietary patterns from exploratory factor analysis. *Br J Nutr* **89**, 409–419.

Schulze MB, Hoffmann K, Kroke A & Boeing H (2003*b*) Risk of hypertension among women in the EPIC-Potsdam study: comparison of relative risk estimates for exploratory and hypothesis-oriented dietary patterns. *Am J Epidemiol* **158**, 365–373.

Schulze MB, Hoffmann K, Manson JE, Willett WC, Meigs JB, Weikert C, Heidemann C, Colditz GA & Hu FB (2005) Dietary pattern, inflammation, and incidence of type 2 diabetes in women. *Am J Clin Nutr* **82**, 675–684.

Schulze MB & Hu FB (2002) Dietary patterns and risk of hypertension, type 2 diabetes mellitus, and coronary heart disease. *Curr Atheroscler Rep* **4**, 462–467.

Stampfer MJ, Hu FB, Manson JE, Rimm EB & Willett WC (2000) Primary prevention of coronary heart disease in women through diet and lifestyle. *N Engl J Med* **343**, 16–22.

Trichopoulou A, Costacou T, Bamia C & Trichopoulos D (2003) Adherence to a Mediterranean diet and survival in a Greek population. *N Engl J Med* **348**, 2599–2608.

van Dam RM (2005) New approaches to the study of dietary patterns. *Br J Nutr* **93**, 573–574.

van Dam RM, Grievink L, Ocke MC & Feskens EJ (2003) Patterns of food consumption and risk factors for cardiovascular disease in the general Dutch population. *Am J Clin Nutr* **77**, 1156–1163.

van Dam RM, Rimm EB, Willett WC, Stampfer MJ & Hu FB (2002) Dietary patterns and risk for type 2 diabetes mellitus in U.S. men. *Ann Intern Med* **136**, 201–209.

Weikert C, Hoffmann K, Dierkes J, Zyriax BC, Klipstein-Grobusch K, Schulze MB, Jung R, Windler E & Boeing H (2005) A homocysteine metabolism-related dietary pattern and the risk of coronary heart disease in two independent German study populations. *J Nutr* **135**, 1981–1988.

World Health Organization (2005) *World Health Report: 2005, Make Every Mother and Child Count*. Geneva: World Health Organization.