May 2004: 177–203

Empirically Derived Eating Patterns Using Factor or Cluster Analysis: A Review

P. K. Newby, Sc.D., M.P.H., M.S. and Katherine L. Tucker, Ph.D.

This paper reviews studies performed to date that have employed cluster or factor analysis to empirically derive eating patterns. Since 1980, at least 93 studies were published that used cluster or factor analysis to define dietary exposures, of which 65 were used to test hypotheses or examine associations between patterns and disease outcomes or biomarkers. Studies were conducted in diverse populations across many countries and continents and suggest that patterns are associated with many different biomarkers and disease outcomes, whether measured by cluster or factor analysis. Despite clear differences in approaches and interpretations, there is some evidence that underlying eating patterns are revealed by either method. Although the research considered herein has created a meaningful body of literature, refining both the factor and cluster analysis methods will help to further establish eating patterns as a sound dietary assessment method.

Key words: eating patterns, dietary patterns, food patterns, factor analysis, cluster analysis, review

doi: 10.1301/nr.2004.may.177–203

Introduction

More than 20 years ago, some of the first papers to examine food patterns and health were published,[1,2] which defined eating patterns as "foods as they are actually consumed in various characteristic combinations."[2] Eating patterns (also referred to as *food* or *dietary* patterns) may be defined theoretically, in which nutritional variables (e.g., foods and nutrients) are grouped according to some a priori criteria of nutritional health, or empirically, in which variables are reduced into a smaller number of variables through statistical manipulation and evaluated a posteriori. Theoretically derived dietary patterns generally fall under the realm of a dietary index that an agency or research group has created in order to rank more and less healthy dietary behavior; examples include the Healthy Eating Index,[3] the Diet Quality Index,[4] and the Alternative Healthy Eating Index.[5] Such constructs are built upon current nutrition knowledge or theory and generally include variables that represent current nutrition guidelines, recommendations, and/or a specific dietary composition that is thought to be most healthful, much of which is generated from empirical research. The index variables are usually quantified and, when summed, provide an overall measure of dietary quality. Conflicts can arise, however, when guidelines or recommendations do not have scientific consensus; theoretically defined measures of dietary patterns often include different dietary variables, or different weightings of variables, resulting in indexes that measure different definitions of "healthful" behavior.

By contrast, empirically derived eating patterns are not defined a priori and do not depend on how the authors define a healthful pattern. Rather, statistical methods are used to generate patterns from collected dietary data. In nutritional epidemiology, factor and cluster analysis are two commonly used methods to derive eating patterns. Factor analysis reduces data into patterns based upon intercorrelations between dietary items, whereas cluster analysis reduces data into patterns based upon individual differences in mean intakes.

Whether derived theoretically or empirically, the assessment of eating patterns has emerged as an alternative method of measuring dietary exposures in nutritional epidemiology because of the limitations of the single nutrient or food approach. Limitations of the traditional approach are methodological, including failing to account for interactions between nutrients, intercorrelations between nutrients (collinearity), and the inability to de-

Drs. Newby and Tucker are with the Jean Mayer USDA Human Nutrition Research Center on Aging at Tufts University, Dietary Assessment and Epidemiology Research Program, 711 Washington St., Boston, MA, 02111. Reprints are not available. Address correspondence to PK Newby, Jean Mayer USDA Human Nutrition Research Center on Aging at Tufts University, 711 Washington St., 9th Floor, Boston, MA 02111. E-mail: pknewby@post.harvard.edu.

tect small effects from single nutrients. Limitations are also conceptual in that people eat meals that include many foods and nutrients, hence studies that are reduced to just one nutrient or food are not only unable to account for how these elements may act in combination, but also are less poised to provide tangible dietary advice. These limitations are well appreciated and often noted.

As a result, studies using patterning methods have increased substantially in the past two decades and many studies using these alternative methods (dietary index, cluster, or factor analysis) have been published. A review of diet quality indexes was published several years ago,[6] and a recent paper on dietary pattern analysis provides a brief overview of methods but focused on a limited group of studies.[7] We are aware of no comprehensive review of empirically derived eating patterns.

The objective of this paper is to review published studies performed to date that have employed cluster or factor analysis to empirically derive eating patterns. Although disease associations are also briefly discussed, this review is not intended to provide an extensive assessment of specific pattern-disease associations. Rather, our specific goals are (1) to examine papers that have used factor analysis, (2) to examine papers that have used cluster analysis, (3) to present and summarize major findings from studies that have tested hypotheses or examined associations using these methods, (4) to discuss methodological issues concerning empirically derived eating patterns in general, and (5) to compare and contrast the cluster and factor analysis methods. The paper concludes by contemplating whether eating pattern analysis is a sound dietary assessment method and providing suggestions for future research.

It is hoped that this review will be a resource to both investigators in the field, who may find this summary useful to their own research projects, and to investigators outside of the field, who may want to better acquaint themselves with the literature using these patterning methods. In addition, is it hoped that this review will stimulate new research ideas using empirically derived eating patterns, which should help to move the field forward.

Literature Review Selection Procedure and Organization

To identify dietary studies that have used pattern analysis, Medline was searched using the terms "factor analysis and diet," "cluster analysis and diet," "eating patterns," "dietary patterns," or "food patterns." The titles and/or abstracts from all articles retrieved from these searches were reviewed to determine whether they should be included. Reference lists from selected articles were also reviewed to locate additional papers that were not retrieved in the Medline search. Only papers published since 1980 were included, although very few papers existed in this field before that time.

Several types of articles are not included in this review, as follows: articles that used dietary data to measure very specific dietary behaviors, such as changes in fat intake[8–10]; articles that included non-dietary variables (e.g., physical activity, use of aspirin, disease states)[11–21]; and articles that measured dietary preferences and behaviors (e.g., preference for meat, eating location).[22–26] Abstracts were not included in this review. Furthermore, the use of factor and cluster analysis is very common in the social science literature and many studies have employed these methods to examine eating behavior with regards to disordered eating and eating disorders,[27,28] but these are not included here. Each of the aforementioned sets of research articles represent additional ways to employ patterning methods to understand eating behavior and are omitted only due to space limitations.

After exclusions, 58 articles using factor analysis and 35 articles using cluster analysis were included in the review and are presented in three separate tables. In Table 1, articles using factor analysis to derive food patterns are described. In Table 2, articles using cluster analysis to derive dietary patterns are described. Table 1 and Table 2 can be found online at: http://www.ilsi.org/info/infolist.cfm?pubentityid=3&infoid=26. One study performed both a factor and cluster analysis in the same article[29] and the derived patterns are separately described in the first two tables, respectively. In both Table 1 and Table 2, studies in which the same patterns were derived within the same or similar datasets from the same population are presented only once, at the first time of publication, with additional reports cited in the table row. Sample details are therefore presented for the first publication only.

In the final table, Table 3, articles testing hypotheses or examining associations between eating patterns derived using either cluster or factor analysis and disease outcomes or biomarkers are presented, along with a brief summary of the major findings. Studies are organized by study design and year of publication. Study results from similar datasets are presented separately in Table 3 if different hypotheses were tested. Descriptive studies appearing in either of the first two tables in which patterns are described according to food or nutrient content and/or demographic characteristics but hypotheses are not tested are not presented in Table 3 but are summarized in the Discussion. There are several articles in which the outcomes as stated in Table 3 were not the primary stated aims of the research article. Three validation studies are also included in Table 3; studies were called "validation" only if they compared patterns with nutrients or patterns derived from a gold standard method (e.g., diet records) and plasma biomarkers (e.g.,

Table 3. Relations between Empirically Derived Eating Patterns and Biomarkers of Health or Disease Outcomes, Sorted by Study Design and Year of Publication*†

First Author (year)	Study Design	Pattern Exposures	Outcome(s)	Results‡
Millen (2004)[101]	Prospective cohort	5 clusters 1. Heart healthy 2. Light eating 3. Wine and moderate eating 4. High-fat 5. Empty calorie (Clusters 2–5 collapsed into 1 "Less heart healthy" cluster)	Subclinical heart disease	Heart healthy cluster/never smoker vs. less heart healthy/current smoker associated with lower risk Subclinical heart disease: OR = 0.17, CI: (0.07, 0.36)
Newby (2004)[30]	Prospective cohort	3 factors 1. Reduced-fat dairy, fruit, and fiber 2. Protein and alcohol 3. Sweets (4–6 not analyzed)	BMI and waist circumference	Reduced-fat dairy factor, highest vs. lowest quintile, associated with smaller change BMI/y (W): $\beta = -0.51$ kg/m^2, CI: (-0.82, -0.20), trend test, $P < 0.01$ Waist circumference/y: $\beta = -1.06$ cm, CI: (-1.88, -0.24), trend test, $P =$ 0.03
Togo (2004)[40]	Prospective cohort	3 factors (M) 1. Green 2. Sweet 3. Traditional 2 factors (W) 1. Green 2. Sweet–Traditional	BMI and obesity	Traditional factor (M) associated with smaller change BMI/11-y: $\beta = -0.40$ kg/m^2, CI: (-0.78, -0.01) Sweet-Traditional factor (W) associated with smaller change BMI/5-y: $\beta = -0.33$ kg/m^2, CI: (-0.54, -0.13) All factors NS (M and W) associated with change in BMI (5-y and 11-y) and obesity when models were adjusted for change in factor scores during follow-up
Tseng (2004)[32]	Prospective cohort	3 factors 1. Vegetable–fruit 2. Red meat–starch 3. Southern	Prostate cancer	All factors NS
Campain (2003)[85]	Prospective cohort	6 clusters 1. High-sugar–low-starch 2. Medium-sugar–medium-starch 3. Medium-sugar–low-starch 4. Low-sugar–low-starch 5. Low-sugar–medium-starch 6. Low-sugar–high-starch	Dental caries	Low sugar–high starch cluster vs. other clusters associated with higher risk Caries–all surfaces: OR = 1.23, CI: (1.06, 0.43) Caries–pit and fissure surfaces: OR = 1.27, CI: (1.09, 1.47) Other clusters NS
Diehr (2003)[87]	Prospective cohort	5 clusters 1. Unhealthy 2. High-calorie 3. Low-calorie 4. Low 4 5. Healthy	Years of healthy life (YHL), Years of life (YOL), myocardial infarction (MI), angina, stroke, congestive heart failure (CHF)	Healthy cluster vs. Unhealthy cluster Most YHL, 10 y and lifetime Most YOL, 10 y and lifetime Low 4 cluster vs. Healthy cluster associated with Lower angina Lower MI

Table 3. Relations between Empirically Derived Eating Patterns and Biomarkers of Health or Disease Outcomes, Sorted by Study Design and Year of Publication*† (Cont'd)

First Author (year)	Study Design	Pattern Exposures	Outcome(s)	Results‡
Masaki (2003)[36]	Prospective cohort	4 factors 1. Vegetable and fruit 2. Western breakfast 3. Meat 4. Rice/snack	Stomach cancer	All factors NS
Newby (2003)[90]	Prospective cohort	5 clusters 1. Healthy 2. White bread 3. Alcohol 4. Sweets 5. Meat and potatoes	BMI and waist circumference	Healthy cluster associated with smallest changes BMI/y: 0.05 ± 0.06 kg/m^2 Waist circumference/y: 0.43 ± 0.27 cm Meat/potatoes cluster vs. Healthy cluster associated with largest change BMI/y: 0.30 ± 0.06 kg/m^2 White bread cluster vs. Healthy cluster associated with largest change Waist circumference: 1.32 ± 0.29 cm
Schulze (2003)[38]	Prospective cohort	2 factors 1. Traditional cooking 2. Fruits and vegetables	Hypertension	All factors NS
Shimizu (2003)[91]	Prospective cohort	4 clusters 1. Vegetables 2. Dairy products 3. Beverages 4. Cereals	Survival, Activities of Daily Living (ADL), Clinical Dementia Rating (CDR), presence/absence of illness	Dairy products cluster vs. Beverage cluster associated with increased Survival: $P = 0.048$ All clusters NS associated with ADL, CDR, or presence/absence of illness at baseline
Millen (2002)[99]	Prospective cohort	5 clusters 1. Heart health 2. Light eating 3. Wine and moderate eating 4. High-fat 5. Empty calorie	Carotid atherosclerosis	Empty calorie cluster vs. Heart Health cluster associated with higher risk Carotid atherosclerosis (W): OR = 2.28, CI: (1.12, 4.62) Other clusters NS vs. Healthy cluster
Osler (2002)[53]	Prospective cohort	2 factors 1. Prudent 2. Western	Coronary Heart Disease (CHD)	Effect of Prudent and Western factors modified by BMI ($P = 0.0106$ for interaction) CHD: Inverse association if low BMI (both factors) CHD: Direct association if high BMI (both factors)
Quatromoni (2002)[100]	Prospective cohort	5 clusters 1. Heart healthy 2. Light eating 3. Wine and moderate eating 4. High-fat 5. Empty calorie	Overweight	All clusters NS vs. Heart healthy cluster
van Dam (2002)[67]	Prospective cohort	2 factors 1. Prudent 2. Western	Type 2 diabetes	Western factor associated with higher risk Type 2 diabetes: RR = 1.59, CI: (1.32, 1.93), trend test, $P < 0.001$ Prudent factor NS

Table 3. Relations between Empirically Derived Eating Patterns and Biomarkers of Health or Disease Outcomes, Sorted by Study Design and Year of Publication*† (Cont'd)

First Author (year)	Study Design	Pattern Exposures	Outcome(s)	Results‡
Fung (2001)[46]	Prospective cohort	2 factors 1. Prudent 2. Western	CHD	Prudent factor, highest vs. lowest quintile, associated with lower risk CHD (W): RR = 0.76, CI (0.60, 0.98) Western factor, highest vs. lowest quintile, associated with higher risk CHD (W): RR = 1.46, CI: (1.07, 1.99) Prudent factor, highest quintile, together with Western factor, lowest quintile associated with lowest risk CHD (W): RR = 0.64, CI: (0.44, 0.92)
Osler (2001)[52]	Prospective cohort	2 factors 1. Prudent 2. Western	All-cause mortality and cardiovascular disease (CVD) mortality	Prudent factor associated with lower risk All-cause mortality (M): HR = 0.84, CI: (0.75, 0.93) All-cause mortality (W): HR = 0.74, CI: (0.64, 0.85) CVD mortality (W): HR = 0.63, CI: (0.44, 0.90) Western pattern NS
Terry (2001)[57]	Prospective cohort	3 factors 1. Healthy 2. Western 3. Drinker	Colorectal cancer	Healthy factor, highest vs. lowest quartile, associated with lower risk Colorectal cancer (W): RR = 0.45, CI: (0.23, 0.88), trend test, $P = 0.03$, for <50 y Other factors NS
Terry (2001)[58]	Prospective cohort	3 factors 1. Healthy 2. Western 3. Drinker	Breast cancer	Healthy and Western factors NS Drinker pattern associated with higher risk Breast cancer (W): RR = 1.27, CI: (1.06, 1.52), trend test, $P = 0.002$, attenuated for W <50 y
Hu (2000)[65]	Prospective cohort	2 factors 1. Prudent 2. Western	CHD	Prudent factor, from lowest to highest quintile, associated with decreasing risk CHD (M): RR = 1.0, 0.87, 0.79, 0.75, 0.70, trend test, $P = 0.0009$ Western factor, from lowest to highest quintile, associated with increasing risk CHD (M): RR = 1.0, 1.21, 1.36, 1.40, 1.64, trend test, $P = 0.0001$)
Kumagai (1999)[68]	Prospective cohort	4 factors§ 1. Plant foods 2. Meat, fat, and oil 3. Bread (rice inverse) 4. Pickled vegetables, fish, and shellfish	All-cause mortality	Plant foods factor associated with lower risk All-cause mortality, last 4 y: OR = 0.75, CI: (0.60, 0.95) All-cause mortality, last 7 y: OR = 0.77, CI: (0.63, 0.95) Other factors NS
Whichelow (1996)[74]	Prospective cohort	4 factors§ 1. Fruit, salads, fiber 2. Starchy and meat 3. High-fat and pasta/rice 4. Sweets	All-cause mortality	Fruit, salads, fiber factor associated with lower risk All-cause mortality: OR = 0.76, CI: (0.65, 0.89) (W only) High-fat and pasta/rice factor associated with higher risk All-cause mortality: OR = 1.2, CI: (1.02, 1.42) (W only) Other factors NS

Table 3. Relations between Empirically Derived Eating Patterns and Biomarkers of Health or Disease Outcomes, Sorted by Study Design and Year of Publication*† (Cont'd)

First Author (year)	Study Design	Pattern Exposures	Outcome(s)	Results‡
Farchi (1989)[116]	Prospective cohort	4 clusters§ 1. Alcohol 2. PUFA 3. MUFA, SFA 4. Carbohydrate	All-cause mortality and CVD mortality	Alcohol cluster associated with highest rate 20-y mortality from stroke/CVD: 14.4 deaths/100 PUFA cluster associated with lowest rate 20-y mortality from stroke/CVD: 5.4 deaths/100
Markaki (2003)[35]	Case-control	4 factors 1. Fruits 2. Raw vegetables 3. Mixed vegetables and fruits 4. Fish and cooked vegetables	Thyroid cancer (papillary and follicular)	Fruit factor, cases vs. controls, associated with lower risk Papillary thyroid cancer: OR = 0.68 Raw vegetables factor, cases vs. controls, associated with lower risk Papillary thyroid cancer: OR = 0.67 Fish and cooked vegetables factor, cases vs. controls, associated with higher risk Follicular thyroid cancer: OR = 2.79
Tsai (2003)[92]	Case-control	2 clusters 1. Healthy high-fiber, low-fat 2. Unhealthy high-fat, low-fiber	Lung cancer	Healthy cluster/not homozygous GSTP1 allele vs. Unhealthy cluster/homozygous GSTP1 allele associated with lower risk Lung cancer: OR = 0.16, CI: (0.04, 0.57) Healthy cluster/GSTM1 null vs. unhealthy cluster/GSTM1 null associated with lower risk Lung cancer: OR = 0.46, CI: (0.21, 1.01)
Chen (2002)[93]	Case-control	6 clusters 1. Healthy 2. High-meat 3. High-salty snacks 4. High-dessert 5. High-milk 6. High-white bread	Adenocarcinoma of the esophagus and distal stomach	All clusters, risk difference, for Distal stomach cancer: $P = 0.04$ All clusters NS vs. Healthy cluster
Handa (2002)[43]	Case-control	13 factors (M) 1. Fruits/vegetables 2. Dessert 3. Meat 4. Rice, tofu 5. Juice 6. Healthful 7. Beef 8. Pies, cake 9. No fast food 10. Miscellaneous 11. Fruit 12. Butter 13. Bottled water 8 factors (W) 1. Fruits/vegetables 2. Dessert 3. Miscellaneous + 4. Meat 5. Miscellaneous − 6. High protein 7. Beverage 8. Unhealthy	Renal cell carcinoma	Unhealthy factor, third vs. first quartile, associated with higher risk Renal cell carcinoma (W): OR = 2.1, CI: (1.3, 3.4) Dessert factor, highest vs. lowest quartile, associated with higher risk Renal cell carcinoma: OR = 3.7, CI: (2.0, 6.8)

Table 3. Relations between Empirically Derived Eating Patterns and Biomarkers of Health or Disease Outcomes, Sorted by Study Design and Year of Publication*† (Cont'd)

First Author (year)	Study Design	Pattern Exposures	Outcome(s)	Results‡
McCann (2001)[49]	Case-control	2 factors 1. Healthy 2. High-fat	Endometrial cancer	Healthy factor, highest vs. lowest score, associated with lower risk Endometrial cancer (W): RR = 0.55, CI: (0.35, 0.84) High-fat factor NS
Ogura (2001)[51]	Case-control	2 factors§ 1. High in Japanese radishes, seaweed, spinach, and low in soft drinks, liquors, sake 2. Hi in beef, pork, chicken, low in cinnamaldehyde and benzoic acid.	Recurrent aphthous stomatitis (RAS)	Japanese vegetables factor, cases vs. controls, associated with RAS: Median factor score = -0.37 (cases) vs. -0.07 (controls), $P < 0.05$ Animal foods factor NS, cases vs. controls
Palli (2001)[54]	Case-control	4 factors 1. Vitamin-rich 2. Traditional 3. Refined 4. Fat-rich	Gastric cancer	Traditional factor, highest vs. lowest tertile, associated with higher risk Gastric cancer: OR = 3.0, CI: (1.8, 4.8), trend test, $P = 0.0001$ 39% cases attributable to Traditional pattern Vitamin-rich factor, highest vs. lowest tertile, associated with lower risk Gastric cancer: OR = 0.5, CI: (0.4, 0.7), trend test, $P = 0.0003$ Factors 3 and 4 NS
Armstrong (1998)[70]	Case-control	4 factors§ 1. Fresh fruit and vegetable 2. Salted, preserved foods 3. Organ meats 4. Alcohol	Nasopharyngeal carcinoma (NPC)	Fresh fruit and vegetable factor associated with lower risk NPC: OR = 0.55 Salted, preserved foods factor associated with higher risk NPC: OR = 1.71 Organ meats factor associated with higher risk NPC: OR = 1.57 Alcohol factor associated with higher risk NPC: OR = 1.48
Slattery (1998)[73]	Case-control	5 factors (M) 1. Western 2. Prudent 3. High-fat/sugar-dairy 4. Drinker 5. Substituter 6. Fruit juice 5 factors (W) 1. Western 2. Prudent 3. High-fat/sugar-dairy 4. Substituter 5. Coffee and roll 6. Drinker	Colon cancer	Western factor associated with higher risk Colon cancer (M): OR = 1.96, CI: (1.22, 3.15) Colon cancer (W): OR = 2.02, CI: (1.21, 3.36) Prudent factor associated with lower risk Colon cancer (M): OR = 0.63, CI: (0.43, 0.92) Colon cancer (W): OR = 0.58, CI: (0.38, 0.87) Both patterns strongest effect if diagnosed <67 y (M and W)

Table 3. Relations between Empirically Derived Eating Patterns and Biomarkers of Health or Disease Outcomes, Sorted by Study Design and Year of Publication*† (Cont'd)

First Author (year)	Study Design	Pattern Exposures	Outcome(s)	Results‡
Newby (2004)[31]	Cross-sectional	5 clusters 1. Healthy 2. White bread 3. Alcohol 4. Sweets 5. Meat-and-potatoes 3 factors 1. Reduced-fat dairy, fruit, and fiber 2. Protein and alcohol 3. Sweets (4–6 not analyzed)	Plasma lipids	Healthy cluster and reduced-fat dairy factor associated respectively with lower Plasma TAG: $\beta = -15.97$, CI: (-29.51, -2.43) Plasma TAG: $\beta = -7.02$, CI: (-12.92, -1.12) Alcohol cluster associated with higher Total cholesterol: $\beta = 12.81$, CI: (2.74, 22.88) Protein and alcohol factor associated with higher Total cholesterol: $\beta = 1.59$, CI: (0.55, 2.63)
Kerver (2003)[34]	Cross-sectional	6 factors 1. Western 2. American-healthy 3. Californian 4. Breakfast 5. Southwestern 6. Convenience (3–6 not analyzed)	Plasma biomarkers	Western factor associated with Higher C-peptide Higher insulin Higher glycated hemoglobin Lower RBC folate American-healthy factor NS
Lin (2003)[88]	Cross-sectional	5 clusters 1. Fruit and breakfast cereal 2. Starchy vegetables 3. Rice 4. Whole milk 5. Sweets	BMI and waist circumference	Rice cluster vs. other clusters associated with larger BMI: OR = 1.05, CI: (1.02, 1.09) Waist circumference: OR = 1.03, CI: (1.01, 1.04) Whole milk cluster vs. other clusters associated with smaller BMI: OR = 0.95, CI: (0.91, 0.99)
Martikainen (2003)[89]	Cross-sectional	6 clusters 1. Very healthy 2. Moderately healthy 3. French 4. Sweet unhealthy 5. Unhealthy 6. Very unhealthy	BMI, waist-to-hip circumference ratio, HDL, and TAG	Unhealthy cluster vs. other clusters associated with Highest BMI (M and W) French cluster vs. other clusters associated with Highest HDL (M and W) Very unhealthy cluster vs. other clusters associated with Highest TAG (M) Sweet cluster vs. other clusters associated with Highest TAG (W) (*P* for heterogeneity = 0.0001 for all associations)
Sanchez-Villegas (2003)[37]	Cross-sectional	2 factors 1. Western 2. Spanish–Mediterranean	BMI and obesity	Western factor (highest quintile) vs. Spanish-Mediterranean factor (highest quintile) associated with higher prevalence Obesity: 8% vs. 5% (*P* not reported)

Table 3. Relations between Empirically Derived Eating Patterns and Biomarkers of Health or Disease Outcomes, Sorted by Study Design and Year of Publication*† (Cont'd)

First Author (year)	Study Design	Pattern Exposures	Outcome(s)	Results‡
van Dam (2003)[41]	Cross-sectional	3 factors 1. Cosmopolitan 2. Traditional 3. Refined foods	Systolic blood pressure, total cholesterol, HDL, and glucose	Cosmopolitan factor associated with Lower blood pressure Higher HDL-C Traditional factor associated with Higher blood pressure Higher HDL-C Higher total cholesterol Higher glucose Refined foods pattern associated Higher total cholesterol
Rasanen (2002)[94]	Cross-sectional	4 clusters 1. Fat milk and butter 2. Sugar and sweets 3. Cereal, rice, and pasta 4. Bread, skim milk, and margarine	Total cholesterol, LDL, HDL, TAG	All clusters NS, comparing all clusters
Sichieri (2002)[45]	Cross-sectional	3 factors 1. Mixed 2. Traditional 3. Western	Overweight (BMI 25–29.99) and obesity (BMI ≥30)	Traditional factor vs. Mixed and Western factor associated with lower risk Overweight (M): OR = 0.87, CI: (0.77, 0.99) Overweight (W): OR = 0.86, CI: (0.75, 0.99) Western and Mixed factors NS
Tucker (2002)[95]	Cross-sectional	6 clusters 1. Meat, dairy, and bread 2. Meat and sweet baked products 3. Sweet baked products 4. Alcohol 5. Candy 6. Fruit, vegetables, and cereal	Bone mineral density (BMD)	Fruit, vegetables, cereal cluster associated with higher Femoral BMD vs. clusters 1, 2, 4, and 5 (M) Trochanter BMD vs. clusters 2 or 4 (M) Ward's area BMD vs. cluster 1 or 5 (M) Radius BMD vs. cluster 5 (M) Candy cluster associated with lower Radius BMD vs. clusters 1, 2, 4, and 6 (W)
Fung (2001)[66]	Cross-sectional	2 factors 1. Prudent 2. Western	Nutrient and plasma biomarkers	Prudent factor associated with Insulin (M): $r = -0.25$ Homocysteine (M): $r = -0.19$ Folate (M): $r = 0.31$ Western factor associated with HDL-C (M): $r = 0.17$ Folate (M): $r = -0.36$ Insulin (M): $r = 0.28$ C-peptide (M): $r = 0.28$
Pryer (2001)[103]	Cross-sectional	3 clusters (M) 1. Mixed 2. Healthy 3. Traditional, high in alcohol 3 clusters (W) 1. Sweet-Traditional 2. Healthy 3. Mixed	Nutrient biomarkers and BMI	Healthy cluster associated with highest α- and β-carotene, folate, and vitamin C Traditional cluster associated with lowest BMI (W): 25.7 kg/m^2 All clusters NS for BMI (M)

Table 3. Relations between Empirically Derived Eating Patterns and Biomarkers of Health or Disease Outcomes, Sorted by Study Design and Year of Publication*† (Cont'd)

First Author (year)	Study Design	Pattern Exposures	Outcome(s)	Results‡
Schulze (2001)[55]	Cross-sectional	7 factors (M) 1. Plain cooking 2. Sweets 3. Cereals 4. Fruit and vegetables 5. Alcohol 6. High-fat dairy 7. Bread and sausage 7 factors (W) 1. Plain cooking 2. Bread and sausage 3. Sweets 4. Fruit and vegetables 5. Low-fat dairy 6. Alcohol 7. Cereals and meat	BMI	Alcohol factor associated with BMI (M): $\beta = 0.02$ BMI (W): $\beta = 0.003$ Sweets factors associated with BMI (M): $\beta = 0.05$ BMI (W): $\beta = -0.01$ Fruit and vegetables factor associated with BMI (M): $\beta = 0.012$ BMI (W): $\beta = 0.004$
Wirfalt (2001)[106]	Cross-sectional	6 clusters 1. Many foods and drinks 2. Fiber bread 3. Low-fat and high-fiber 4. White bread 5. Milk fat 6. Sweets and cakes	Metabolic syndrome	Fibre bread cluster associated with lower risk Central obesity (M): OR = 0.61, CI: (0.42, 0.89) Milk fat cluster associated with lower risk Hyperinsulinemia (W): OR = 0.58, CI: (0.40, 0.84)
Greenwood (2000)[104]	Cross-sectional	7 clusters 1. Monotonous low-quantity omnivores 2. Health-conscious 3. Traditional meat, chips, and pudding 4. Higher-diversity traditional omnivores 5. Conservative omnivores 6. Low-diversity vegetarians 7. High-diversity vegetarians	BMI and obesity	Vegetarian clusters 6 and 7 associated with smallest BMI (W): 23 kg/m^2 Monotonous low-quantity omnivores associated with higher Obesity (W): 12% Traditional meat, chips, and pudding associated with highest Obesity (W): 12%
Maskarinec (2000)[60]	Cross-sectional	4 factors 1. Meat 2. Vegetable 3. Bean 4. Cold foods	BMI	Meat factor associated with BMI (W): $r = 0.17$ Beans factor associated with BMI (W): $r = -0.13$ Cold foods associated with BMI (W): $r = -0.13$ Vegetable factor NS Associations similar across ethnic groups

Table 3. Relations between Empirically Derived Eating Patterns and Biomarkers of Health or Disease Outcomes, Sorted by Study Design and Year of Publication*† (Cont'd)

First Author (year)	Study Design	Pattern Exposures	Outcome(s)	Results‡
Tseng (2000)[62]	Cross-sectional	3 factors (M) 1. Vegetables 2. High-calorie 3. Traditional 4 factors (W) 1. Vegetables 2. High-calorie 3. Traditional 4. Fruit	Gallbladder disease	All factors NS (W) Traditional factor associated with Gallbladder disease: OR = 0.42, CI: (0.21, 0.85), comparing 3rd to 1st quartile
Williams (2000)[63]	Cross-sectional	4 factors§ 1. Vegetables, fruit, fish, pasta, rice 2. Sweets, root vegetables, biscuits, potatoes 3. Chocolate, sweets, crisps, cheese 4. Eggs, cheese, processed meats, fried food	Metabolic syndrome	Vegetables factor associated with Waist-to-hip ratio: $r = -0.29$ TAG: $r = -0.18$ HDL-C: $r = 0.19$ Eggs factor associated with Waist/hip ratio: $r = 0.15$ Sweets factor associated with 120 min plasma insulin: $r = -0.09$ Plasma non-esterified fatty acids: $r = -0.09$ Vegetables pattern associated with lower risk Undiagnosed type 2 diabetes: OR = 0.54, CI: (0.32, 0.91), strongest ≥50 y
Wirfalt (2000)[105]	Cross-sectional	6 clusters (unstandardized) 1. Many foods and drinks 2. Fiber bread 3. Low-fat and high-fiber 4. White bread 5. Milk fat 6. Sweets and cakes 6 clusters (standardized) 1. Drinks and fries 2. Ice cream and cake 3. Dieters 4. Healthy 5. Traditional 6. Mediterranean	BMI and waist-hip ratio	Unstandardized solution, lowest vs. highest BMI group: Many foods and drinks cluster associated with lower risk BMI <23.9: OR = 0.84, CI: (0.71, 0.99) Low-fat and high-fiber cluster associated with lower risk BMI <23.9: OR = 0.70, CI: (0.57, 0.87) White bread cluster associated with lower risk BMI <23.9: OR = 0.79, CI: (0.63, 0.98) Standardized solution, lowest vs. highest BMI group: Ice-cream and cake cluster associated with lower risk BMI <23.9: OR = 0.57, CI: (0.40, 0.81) BMI 23.9–26.9: OR = 0.63, CI: (0.50, 0.95) All clusters NS (waist-hip ratio)
Gittelsohn (1998)[72]	Cross-sectional	7 factors 1. Vegetables 2. Junk foods 3. Bush foods 4. Breakfast foods 5. Hot meal foods 6. Tea foods 7. Bread and butter	Diabetes, impaired glucose tolerance (IGT), and obesity	Bush foods factor, highest vs. lowest quartile, associated with higher risk Obesity: OR = 1.94, CI: (1.22, 3.08) Vegetables factor, highest vs. lowest quintile, associated with lower risk IGT: OR = 0.41, CI: (0.18, 0.91) Breakfast foods factor, highest vs. lowest quintile, associated with lower risk IGT: OR = 0.41, CI: (0.18, 0.93) Junk foods factor, highest vs. lowest quintile, associated with higher risk Diabetes: OR = 2.4, CI: (1.13, 5.10) Bread and butter factor, highest vs. lowest quintile, associated with higher risk Diabetes: OR = 2.2, CI: (1.22, 4.41)

Table 3. Relations between Empirically Derived Eating Patterns and Biomarkers of Health or Disease Outcomes, Sorted by Study Design and Year of Publication*† (Cont'd)

First Author (year)	Study Design	Pattern Exposures	Outcome(s)	Results‡
Haveman-Nies (1998)[107]	Cross-sectional	5 clusters 1. Light snackers 2. Fruit and vegetable snackers 3. Sweet drinkers 4. Dairy snackers 5. Alcohol drinkers	Biomarkers, BMI, ADL	Sweet drinkers cluster associated with highest Total carotenes Vitamin B_6 Vitamin B_{12} Fruit and vegetable cluster with highest Vitamin B_{12} Dairy and alcohol clusters with highest BMI Dairy snackers cluster vs. other clusters highest percent Reporting no chronic disease: All clusters NS associated with ADLs
Wirfalt (1997)[109]	Cross-sectional	6 clusters 1. Soft drinks 2. Pastry 3. Skim milk 4. Meat 5. Meat-cheese 6. White bread	BMI	Soft drink cluster vs. Skim milk and Meat-cheese clusters associated with higher BMI (M): 30.5 ± 0.76 (vs. 28.3 ± 0.73 and 28.0 ± 0.59) All clusters NS (W)
Schroll (1996)[111]	Cross-sectional	4 clusters 1. Lean and green eaters 2. Gourmands 3. Milk drinkers 4. Small eaters 5. Modest eaters (W only)	Activities of Daily Living (ADL), BMI, presence of chronic disease, self-perceived health	Gourmands cluster vs. Small eaters cluster associated with higher percentage Reporting good health (M): 68% vs. 51% Small eaters cluster associated with higher percentage Reporting chronic disease (M): 82% Modest eaters cluster associated with higher percentage Reporting chronic disease (W): 84% All clusters NS associated with ADLs or BMI (M) All clusters NS associated with self-perceived health, ADL, or BMI (W)
Huijbregts (1995)[112]	Cross-sectional	4 clusters 1. Alcohol 2. Meat 3. Healthy 4. Refined sugars	Cardiovascular risk factors	Alcohol cluster vs. other clusters associated with higher HDL-C (M) All clusters NS (M) associated with BMI, total cholesterol, non-HDL cholesterol, blood pressure, or hypertension, across clusters
Wolff (1995)[76]	Cross-sectional	7 factors 1. Nutrient dense 2. Traditional 3. Transitional 4. Nutrient-dilute 5. Protein-rich 6. High-fat dairy 7. Mixed dishes	Birth weight	Nutrient dense factor positively associated with Birth weight: $\beta = 20.4 \pm 4.6$, $P = 0.0001$ Protein-rich factor positively associated with Birth weight: $\beta = 36.1 \pm 14.1$, $P = 0.05$ Nutrient-dilute factor negatively associated with Birth weight: $\beta = -22.2 \pm 10.0$, $P = 0.05$

Table 3. Relations between Empirically Derived Eating Patterns and Biomarkers of Health or Disease Outcomes, Sorted by Study Design and Year of Publication*† (Cont'd)

First Author (year)	Study Design	Pattern Exposures	Outcome(s)	Results‡
Barker (1992)[78]	Cross-sectional	4 factors 1. Traditional 2. Cosmopolitan 3. Convenience 4. Meat and two vegetables	Biochemical and hematologic variables	Meat and two vegetable factor associated with HDL (M): $r = 0.19$ Cosmopolitan factor associated with Total cholesterol (W): $r = -0.11$ White cell count (W): $r = -0.15$ HDL (M): $r = 0.12$ HDL (W): $r = 0.21$ Convenience factor associated with HDL (M): $r = 0.16$ HDL (W): $r = 0.20$
Tucker (1992)[114]	Cross-sectional	4 clusters 1. Alcohol 2. Milk/cereal/fruit 3. Bread/poultry 4. Meat/potatoes	Biochemical and hematologic variables, BMI	Milk/cereal/fruit cluster associated with highest Plasma B_{12} Alcohol cluster associated with highest Lowest BMI Highest HDL Plasma folate RBC riboflavin Meat/potatoes clusters associated with highest Vitamin B_6
Barker (1990)[79]	Cross-sectional	4 factors 1. Traditional 2. Cosmopolitan 3. Convenience 4. Meat and two vegetables	BMI	Convenience factor associated with lower BMI: $\beta = -0.14$ Factors 2–4 NS
Iizumi (1986)[83]	Cross-sectional	8 factors 1. Staple foods 2. Sweets 3. Alcoholic beverages 4. Animal foods 5. Prepared foods 6. Traditional Japanese foods 7. Vegetable oil 8. Nutrition-consciousness	Total cholesterol	Staple foods, Prepared foods, Traditional Japanese foods, Vegetable oils, and Nutrition-consciousness factors together in one model predicted Total cholesterol ($R = 0.56$, $R^2 = 0.31$)
Schwerin (1982)[2]	Cross-sectional	7 factors (Ten State) 1. Dairy and soups 2. Nonsugary beverages 3. Eggs, legumes, and cereals 4. Meats, vegetables, fruits, and desserts 5. Poultry 6. Mixed protein and shellfish 7. Fish and fats	Clinical symptoms (e.g., of nutrient deficiencies) and biochemical deficiencies (e.g., hemoglobin, albumin, iron)	Dairy and soups factor associated with highest percentage Free from clinical or biochemical symptoms: 50% (Ten State) Soup and shellfish factor associated with highest percentage Free from clinical or biochemical symptoms: 41% (NHANES I) Data not presented for NCFS

Table 3. Relations between Empirically Derived Eating Patterns and Biomarkers of Health or Disease Outcomes, Sorted by Study Design and Year of Publication*† (Cont'd)

First Author (year)	Study Design	Pattern Exposures	Outcome(s)	Results‡
		6 factors (NFCS) 1. Soups and dairy 2. Nonsugary beverages 3. Cereals, eggs, and legumes 4. Vegetables, fruits, fats, and dessert 5. Poultry and fish 6. Mixed protein and sugary 8 factors (HANES I) 1. Soups and shellfish 2. Nonsugary beverages 3. Cereals, eggs, and legumes 4. Vegetables, fruits, desserts, and meats 5. Poultry 6. Fish 7. Mixed protein 8. Sugary		
Schwerin (1981)[1]	Cross-sectional	7 factors 1. Dairy and soups 2. Nonsugary beverages 3. Eggs, legumes, nuts, cereal, and grains 4. Meats, vegetables/fruit/juices, and desserts 5. Poultry 6. Mixed protein, shellfish 7. Fish, fats, and oils	Clinical symptoms and biochemical deficiencies	Dairy and soups factor associated with highest percentage Free from clinical symptoms: 58% (Ten State) and 52% (NHANES I) Free from biochemical deficiencies: 80% (Ten State) and 66% (NHANES I) Non-sugary beverages factor associated with lowest percentage Free from clinical symptoms: 35% (Ten State) and 26% (NHANES I)
Ishimoto (1994)[77]	Ecologic	2 factors§ 1. High-total fat, saturated fat, MUFA, PUFA, n-6, and low-cholesterol 2. Low-total fat, high-cholesterol, low-MUFA, PUFA, n-6, and n-3	Breast cancer mortality	Low-total fat factor associated with Breast cancer mortality: $r = 0.72$ High-total fat factor NS All factors NS associated with rates
Quatromoni (2002)[98]	Validation	5 clusters 1. Heart Healthy 2. Light eating 3. Wine and moderate eating 4. High-fat 5. Empty calorie	Nutrients from 3-d diet records and CHD risk factors	Heart Healthy cluster vs. other clusters associated with Most favorable nutrient profiles (W) Lowest overall dietary risk (W) Lowest overall CHD risk (W) Lowest total cholesterol (W) Lowest TAG (W)

Table 3. Relations between Empirically Derived Eating Patterns and Biomarkers of Health or Disease Outcomes, Sorted by Study Design and Year of Publication*† (Cont'd)

First Author (year)	Study Design	Pattern Exposures	Outcome(s)	Results‡
Millen (2001)[97]	Validation	5 clusters 1. Heart healthy 2. Light eating 3. Wine and moderate eating 4. High-fat 5. Empty calorie	Nutrients from diet records and heart disease risk factors	Heart healthy cluster vs. other clusters had Lowest percentage fat (W) Lowest cholesterol intake (W) Highest intakes of fiber, calcium, vitamin C, vitamin B_6, vitamin E, β-carotene, and folate (W) Wine and moderate eating cluster vs. other clusters highest percentage Normal weight (W) Desirable LDL (W) Normotensive (W) Empty calorie cluster vs. other clusters had lowest percentage Normal weight (W) Desirable LDL (W)
Hu (1999)[64]	Validation	2 factors 1. Prudent 2. Western	Nutrients from diet records and plasma biomarkers Reproducibility and validity correlations	Prudent factor positively associated with Many nutrients and biomarkers (M) Reproducibility (M): $r = 0.70$ (2 FFQs) Validity (M): $r = 0.52$ (FFQ and diet record) Western factor inversely associated with Many micronutrients and carotenoids (M) Reproducibility (M): $r = 0.67$ (2 FFQs) Validity (M): $r = 0.74$ (FFQ and diet record)

*Only studies associating patterns with disease outcomes or biomarkers are presented in this table. Longitudinal studies are included under the "Prospective" study design category.

†As much as possible, only major, multivariate, adjusted results concerning the main outcome variable(s) are presented. In studies where multiple analyses are performed (e.g., cross sectional and prospective), prospective results are presented. All results presented are significant with 95% confidence intervals (CI, $P < 0.05$) unless otherwise stated. Results presented are for both men and women unless otherwise noted as (M) or (W).

‡Abbreviations: BMI = body mass index, M = men, W = women, OR = odds ratio, RR = risk ratio, HR = hazard rate, NS = not significant ($P > 0.05$), HDL = HDL cholesterol, LDL = LDL cholesterol, TAG = triacylglycerols, RBC = red blood cell, PUFA = polyunsaturated fatty acids, MUFA = monounsaturated fatty acids, SFA = saturated fatty acids, *n*-6 = omega-6 fatty acids, *n*-3 = omega-3 fatty acids.

§Articles in which patterns were not named are labeled here by following descriptions provided by the author(s), where applicable, or by listing the food (groups) with the highest mean intakes for each cluster or highest factor loadings for each factor.

plasma nutrients). Analyses performed that were not related to eating patterns are not presented or discussed here.

Summary of Findings

Studies from each table are summarized in their own section below. Given that the focus of this review is methodological, the summaries for tables 1 and 2 focus on the statistical methods used in the analyses and the reproducibility of patterns across studies. Likewise, the summary for Table 3 is limited to describing study designs and outcomes. A broader discussion of study findings appears in the following section.

Table 1: Factor Analysis Studies

Table 1 ($n = 58$) presents studies that have used factor analysis to derive eating patterns.[1,2,29–84] Factor analysis was performed in diverse populations and settings, including many American states, several countries in Europe and Asia (Japan, Malaysia, and India), Australia, Canada, and Brazil. In most studies, food-frequency questionnaires (FFQs) or dietary records were used as the primary dietary assessment method. Some studies used a 24-hour recall, dietary history, dietary interview, or another type of dietary questionnaire or survey to assess diet. Individual food and nutrient intakes were derived from these methods and most studies collapsed the original measured dietary items into a smaller number of input variables, usually food groups, for entry into the factor analysis. Although input variables were most often food groups, some studies entered all individual food items from the primary method into a factor analysis. Several studies used macronutrient and/or micronutrient intakes in the analysis rather than foods or food groups.[54,77,84]

In preparing the data for factor analysis, intakes from the primary dietary assessment method may be measured in several ways, including frequency (servings), weight (grams), or daily percent energy contribution. Input variables may be further adjusted for total energy intake, transformed to a standard deviation score (Z-score), or log-transformed. Several studies ordinally ranked[42,68,69] and two studies dichotomized[33,47] their input variables. Some analyses used different units of measure for the different types of input variables. For example, van Dam[41] converted 178 food items from an FFQ to 46 food groups, which included energy-adjusted micronutrients (g/d), macronutrients (percent energy contribution), and alcohol consumption (drinks/d).

After the input variables are prepared, they are entered into a factor analysis procedure using statistical software. In the majority of studies, a principal components analysis (PCA) using varimax (orthogonal) rotation and eigenvalues >1 (eigenvalues >1 indicate that the factor explains more of the variance in the correlations than is explained by a single variable), was performed, although not all studies provided information on whether and how factors were rotated and/or the cutpoints for the eigenvalues. Factor scores were often categorized for further analysis.

The number of derived factors ranged from 2 to 25 and percent variance explained ranged from 15%[62] to 93%.[60] Many authors did not report the percent variance explained by the factors.

Patterns were labeled either quantitatively or qualitatively. Many patterns were named according to the input variable with the highest factor loading (i.e., quantitatively), such as fruits, vegetables, cereals, or meat. Other patterns were named according to quantitative descriptions of dietary composition (e.g., High-fat, Vitamin-rich, High-energy density). The majority of patterns were named qualitatively, according to specific combinations of foods and/or descriptions of nutritional composition. Patterns that contained a variety of different foods or food groups that combined together in "more" and "less" healthy combinations were often given qualitative labels to denote healthfulness. For example, a pattern with high factor loadings for fruits, vegetables, grains, and low-fat dairy was named "Healthy."[49] Other qualitative labels referred to the overall nutritional quality of the pattern (e.g., Refined, Satiating capacity) or to cultural or geographic descriptions of dietary intake, such as "Traditional," "Cosmopolitan," "Southern," and "Spanish-Mediterranean."

Recalling that factor analysis methods and input variables have varied across studies, the extent to which similar patterns are seen in diverse populations may be an indicator of reproducibility. For example, a pattern with high factor loadings for several healthy foods, including fruit, vegetables, legumes, and fish, was called "Prudent," and a pattern with high factor loadings for red and processed meats, eggs, refined grains, and sugar was called "Western"; these patterns were first documented in 1998.[73] Since that time, Western and Prudent patterns were observed in three additional study populations.[46,52,64] Five additional populations reported a Western pattern but not a Prudent pattern per se.[29,34,37,45,57] In what could be considered a variant of the "Prudent" pattern, 12 populations reported a so-called healthful diet, named as "Healthy,"[49,57] "Healthful,"[43,69,80] "Healthfulness,"[81] "Health foods,"[80] "Health conscious,"[71] "Healthy/Prudent,"[75] "Southern European/Healthful,"[80] "American-healthy,"[34] and "Nutrition consciousness."[83]

A pattern high in desserts or sweets appeared the most reproducible and was observed in 15 separate populations.[29,30,33,39,42–44,55,56,61,63,74,80,83] In addition, seven populations observed a pattern high in alcohol, named "Alcohol,"[33,55,70] "Alcoholic beverages,"[83] or "Drinker."[57,73] (Note: all counts refer to separate study populations and not to separate reports from the same study population.)

Table 2: Cluster Analysis Studies

Studies (n = 35) that have used cluster analysis to derive eating patterns are presented in Table 2.[29,31,85–117] Like the factor analysis studies, studies employing cluster analysis have been conducted in many countries and continents using varying types of dietary assessment methods (e.g., FFQs, dietary records). Likewise, measurement of input variables (e.g., daily servings, percent energy) and treatment of input variables (e.g., adjusted for energy or standardized) also varied across studies. Most studies formed a reduced set of food groups from the primary dietary data. A greater percentage of cluster studies treated the input variables as percent energy contribution and/or Z-scores compared with the factor studies.

In most cluster analysis studies, the input variables were analyzed using either the K-Means or Ward's Method. One study also used the partitioning around medoids (PAM) method.[85]

The number of derived clusters ranged from 2 to 8, with the majority of studies reporting a 5- or 6-cluster solution. The method by which the clusters were named was similar to that used in the factor analysis studies. In addition, similar variations of a healthy pattern were observed, including "Healthy,"[87,90,92,93,102,105,112] "More healthy,"[108] "Healthy cosmopolitan,"[102] "Very healthy,"[89] "Healthier,"[102,103] "Health conscious,"[104] and "Heart Healthy."[97,110] Notable similarities with the factor analysis studies were the reproducibility of a pattern high in sweets[86,88,90,93–96,105,106,109,117] and a pattern high in alcohol.[86,90,95,96,107,110,112,114–117] Unlike the factors, none of the clusters were named "Western" or "Prudent."

Table 3: Patterns in Relation to Disease Outcomes or Biomarkers

Table 3 presents results from 65 studies that have examined the relation between eating patterns derived from either cluster or factor analysis and a biomarker or disease outcome.[1,2,30–32,34–38,40,41,43,45,46,49,51–55,57,58,60,62–68,70,72,74,76–79,83,85,87–95,97–99,101,103–107,109,111,112,114,116,118] In total, 22 prospective studies, 9 case control studies, 30 cross-sectional studies, 1 ecologic study, and 3 validation studies were included.

Derived patterns were examined in relation to many different outcomes, including indicators of cardiovascular or coronary heart disease; anthropometric measures; overweight and obesity; many different cancers; symptoms of the metabolic syndrome, including hypertension, blood pressure, cholesterol, blood glucose, and blood insulin measures; type 2 diabetes; and all-cause mortality. Three studies examined the association between patterns and Activities of Daily Living (ADL).[91,107,111] One study each looked at the relation with bone mineral density (BMD),[95] dental caries,[85] recurrent aphthous stomatitis,[51] and birthweight.[76]

Very few validation studies of eating patterns have been performed.[64,97,98] Only one of the validation studies[64] considered the stability of patterns over time, an indicator of reproducibility, by comparing factor solutions from an FFQ at two time points. Correlations were 0.70 for the Prudent pattern and 0.67 for the Western pattern among male health professionals.[64] Although not validation studies per se, some studies did include a reproducibility or validation component in their analyses. A longitudinal study[40] found that correlation coefficients ranged from 0.88 to 0.95 for factors derived 5 years apart, although this study involved very few dietary items in its analysis. An early reproducibility study in Japan[82] found that many of the derived patterns were stable after an 11-year follow-up, although stability differed among men and women. Confirmatory factor analysis for factors[32,39,40,60] and discriminant analysis for clusters[60,87,92,98,105] may be used to test the internal validity of a pattern solution. Internal validation of both methods can also be performed by splitting the study sample and repeating the analysis.[33,38,44,60,62–64,89]

Discussion

The discussion includes three sections. Because the summaries of the tables above were focused on methodological issues, the first section of the Discussion will examine in greater detail the findings from the articles reviewed in the tables, elaborating on demographic associations with patterns and pattern-disease relations. The second section will discuss several general areas of methodological considerations that pertain to patterning analysis (both cluster and factor analysis). The third section will compare and contrast the factor and cluster analysis methods.

Major Findings

Since 1980, at least 93 studies were published that used cluster or factor analysis to define dietary exposures and 65 of these were used to test hypotheses or examine associations between patterns and disease outcomes or biomarkers. These studies were conducted in diverse populations across many countries and continents.

Many reports in the literature have described eating patterns and suggest that patterns do differ in nutrient composition. Likewise, many studies show that eating patterns are associated with other characteristics, including sex, age, socioeconomic status, and general health habits (e.g., smoking, drinking). For example, more women than men have been shown to have a healthier eating pattern[29,59,63,71,75,86,90,106,108] and socioeconomic associations were in the expected direction (e.g., income and education were highest in the Healthy cluster).[59,86,103] Age has been directly[29,79,82] and inversely[63,86] associated with a healthier pattern.

Eating patterns are also associated with other health behaviors. A healthy eating pattern group usually has the highest percentage of nonsmokers,[66,75,97,98,112] exercisers,[55,66,73] and vitamin users,[66] for example. Patterns high in alcohol, which are observed in many studies using either cluster and factor analysis, usually contain a greater percentage of men than women[107] and more smokers.[55,97,112,116] In fact, it is because diet is just one aspect of a healthy lifestyle that many investigators included other health behavior variables (e.g., exercise behavior) as input variables in the pattern analysis, although these articles[11–20] were not included in this review.

While descriptive studies are useful, research efforts today have expanded into analytic epidemiology, in which hypotheses concerning patterns are tested. Ultimately, whether eating patterns can reliably predict disease is an important indicator of their validity, hence their utility in epidemiologic and clinical research. Studies that examined the associations between eating patterns and multiple diseases are presented in Table 3. There is not room to adequately cover each of these disease areas in this review, nor is it the goal of this review to do so. Furthermore, any discussion of diet and disease should include evidence from many different types of studies and should not focus on only one assessment method of dietary exposure. Therefore, the relation of eating patterns to body mass index (BMI, kg/m^2) and obesity, to diabetes and the metabolic syndrome, and to cardiovascular diseases will be discussed only briefly as examples of some of the research in this area.

Several studies examined the relation between eating patterns and anthropometry, including BMI, overweight, obesity, and waist circumference. Findings were inconsistent, as Togo et al.[119] also noted in their review of the literature on food patterns and BMI and obesity. While one prospective study showed no significant relation between any eating pattern and development of overweight,[100] two prospective studies in the same dataset[30,90] found an inverse association between a Healthy cluster[90] and a Reduced-fat dairy, cereals, fruit, and fiber factor[30] and annual change in BMI. A Rice pattern was directly and a Whole milk pattern inversely related to BMI in one study.[88] A Meat pattern[60] and a Meat-and-potatoes pattern[90] have been positively related to BMI, as has a Soft drinks patterns (in men only).[109] In addition, while Greenwood[104] found a protective effect of two vegetarian patterns on BMI, Huijbregts and colleagues[112] saw no differences in BMI across patterns.

Inconsistencies may be explained in part by differences in sex or age. For example, the Prudent and Western patterns were significantly related to BMI in the expected directions in a study of men[65] but not among women[46] and another study found that a Western pattern was positively related to BMI while a Prudent pattern was unrelated to BMI in both sexes.[73] Other differences in BMI by sex have been observed.[55,60,79] Likewise, a Traditional pattern high in rice and beans was inversely related to risk of overweight, but only among adults 20 to 40 years old.[45] More research is needed to better understand whether the effects of diet are truly modified by sex or age. In addition, as with other obesity research, prospective studies are of more use than cross-sectional studies, which pose a greater risk of reverse causality.

Several studies considered the relation between eating patterns and metabolic disorders and diseases, including type 2 diabetes. A Western pattern was significantly related to increased risk of type 2 diabetes,[67] as were Junk food and Bread-and-butter patterns.[72] On the other hand, a pattern high in vegetables, fruit, fish, and pasta was inversely associated with type 2 diabetes in two studies,[63,72] as were two patterns high in breakfast foods and hot foods.[72]

Significant relations were also observed for associations between eating patterns and symptoms of the metabolic syndrome. A Cosmopolitan pattern was associated with lower blood pressure,[41] higher HDL cholesterol (HDL),[41,78] and lower total cholesterol (T-C),[78] while a Refined foods pattern was positively associated with T-C.[41] A Vegetable pattern was negatively correlated with waist-to-hip ratio and triacylglycerols and directly correlated with HDL.[63] A pattern high in reduced-fat dairy, cereals, fruits, and fruit juice was inversely related to triacylglycerols.[31] In two studies, patterns high in alcohol were significantly directly related to HDL.[112,114] Two studies showed no relation between patterns and hyptertension[38] or serum lipids.[94]

As with anthropometric outcomes, relations between eating patterns and lipids differed between men and women. A Cosmopolitan pattern was inversely associated with total cholesterol in women but not in men.[78] In that study, several patterns were directly related to HDL cholesterol in both men and women, although the associations were stronger for women. In another study, a Fiber bread cluster was inversely related to central obesity among men but not women.[106] Several studies in humans[41,120] and animals[121,122] have also observed different effects of diet on plasma lipids between men and women, possibly due to hormonal and sex differences in cholesterol metabolism.[120,122]

Risk of heart disease and other cardiovascular outcomes have been investigated using eating pattern exposures. In one study, a Heart Healthy cluster had the lowest overall CHD risk,[98] and an Empty Calorie cluster in another study had an increased risk of carotid atherosclerosis compared with a Heart Healthy cluster.[99] Factor analysis studies found that a Prudent pattern is protective against CHD[46,65] and CVD.[52] There is an increased risk for CHD for a Western pattern in two of these three studies.[46,65] The effect of eating pattern on CHD may be modified by BMI.[53]

The majority of the remaining studies included in Table 3 examined the relation between eating patterns and a variety of cancer outcomes, including cancer of the breast, colon, and stomach, among others. For example, a Western pattern was directly related to colon cancer in one study[118] but not another,[57] while a Healthy pattern was inversely related to colorectal cancers in both studies.[57,118] One notable case-control study[92] examined the relation between eating patterns and lung cancer among smokers of differing genotype, and found that smokers who were not homozygous for the most common GSTP1 allele and had a Healthy pattern had a lower risk of lung cancer than smokers who were homozygous for the GSTP1 allele with an Unhealthy pattern. For many of the individual cancer outcomes, very few (often only 1) patterning studies have been performed.

In summary, studies to date suggest that eating patterns are significantly associated with many different disease outcomes and plasma biomarkers, whether measured by cluster or factor analysis. However, many inconsistencies in findings also remain.

Methodological Considerations of Pattern Analysis

There are several methodological issues involved in pattern analysis. A major consideration is the inherent subjectivity of both factor and cluster analysis. Another question is whether patterns should be derived separately for women and men.

There are many opportunities for subjectivity that occur throughout the pattern analysis and decisions made by the investigators may have an impact on the number and type of patterns that are derived, reported, and analyzed. Several investigators in this field have provided useful commentary on the issue of subjectivity and other methodological considerations.[7,118,123,124] Therefore, discussion points of methodological issues and subjectivity in particular are reiterated only as a framework from which the literature reviewed here is discussed. Specifically, the investigator must first decide whether or not to further collapse the primary dietary data into a smaller number of items for entry into the analysis. If the data are collapsed, s/he must decide how to group the data. Next, the investigator must decide how the input variables should be treated. After the input variables have been entered into the procedure, the author must decide how many patterns (the output variables) s/he should retain in the final solution, which patterns should be reported or analyzed, and how the patterns should be named. Each of these steps will be briefly discussed.

First, this review indicates that the majority of authors did further collapse the dietary data into food groups before beginning the pattern analysis, or the dietary data were of a sufficiently limited number (e.g., <25 items in total) such that further collapsing was not necessary. Food-grouping schemes share similarities and differences across studies. For example, vegetables may appear as one group or may be further expanded to a larger number of specific vegetable groups (e.g., according to nutrient or phytochemical content). Studies with more vegetable or fruit groups may affect the factor solution and study results, since types of vegetables and fruits may in fact belong to different overall eating patterns. Nonetheless, how vegetables are grouped often depends not only how other studies have previously defined food groups but also on the individual study question. For example, it may be more important to consider different types of vegetables, hence nutrient and carotenoid content, in studies examining the relation between eating patterns and certain cancers. It is not, therefore, clear that all studies should use the same food groupings. Rather, it should be noted that differences in the food group input variables may affect the patterns derived. Specifically regarding factor analysis, a smaller number of input variables included in the procedure explains a greater percentage of the variance in intake compared with a larger number of input variables.[29,49] A study investigating the relation between patterns and endometrial cancer found that, while changing the number of input variables from 168 to 56 to 36 did not affect either the number or composition of derived factors, input variables with reduced detail tended to attenuate the odds ratios for the Healthy pattern.[49]

The author next must decide how the input variables should be quantified. Schwerin et al.[1] pointed out that the input variables could be treated as weight, frequency, or percent energy contribution. The first two tables indicate that most studies use average daily servings from each variable and some studies used daily percent energy contribution variables. Each approach has advantages and disadvantages. Inputting variables as a percent energy contribution, for example, does not allow for noncaloric foods to be included in the analysis, notably, diet sodas, water, coffee, and tea,[2] and these variables do help to define eating patterns. On the other hand, treating variables as percent energy contribution allows patterns to be derived that are in proportion to daily energy intake.

Another question surrounding treatment of input variables is whether variables should be adjusted for total energy. A few studies energy adjusted the nutrients included in the pattern analysis using the nutrient residual approach.[33,41] We are aware of two studies that have specifically addressed whether energy adjustment affects the pattern solution. In one study,[48] 5 factors (Vegetable, Fish, Contaminant, Offal, and Anti-fat/high-carbohydrate factor) were obtained when nutrient input variables were energy adjusted, whereas only 2 major factors (Low-fat/low-energy and Grain-nut) were obtained when variables were not energy adjusted. The authors concluded that nutrients entered into a factor analysis should be energy adjusted.[48] In another report, Balder et al.[33] found that the factor solution was not dramatically affected by energy adjustment, although they did note that energy adjustment forces "substitution" patterns because it requires that food groups are not correlated with total energy.

Authors must also decide whether or not to transform input variables. Only two studies log-transformed some of their food group input variables to improve normality.[34,60] However, many cluster studies and some factor studies transformed input variables into Z-scores, a transformation that standardizes the contribution of all the variables to the distance measured.[105] Because the cluster analysis procedure is very sensitive to outliers, untransformed dietary variables with different ranges, scales, or units can affect the cluster solution in that variables with larger values will outweigh those with smaller values. There may be cases where Z-score transformation is either appropriate or undesirable. (Recall that factor scores are always Z-scores. This is different from transforming the input variables into Z-scores prior to the analysis.) Greenwood and colleagues[104] conducted their cluster analysis using both non-standardized and standardized input variables, hypothesizing that standardizing the variables would dilute differences between clusters and would also ignore correlations between variables. When they repeated their analysis using standardized variables, however, the derived clusters were similar. In another study, Wirfalt et al.[105] performed a cluster analysis in which they converted the dietary variables to percent energy contribution and then treated them as both unstandardized and standardized. While both solutions led to "more healthy" and "less healthy" patterns, the nutrient intake differences across clusters were greater and the contribution of the number of individuals more even across clusters with the unstandardized variables. The authors concluded that treating the variables as unstandardized provided greater discrimination.[105] An elegant study recently published by Balder et al.[33] found that derived patterns using factor analysis were robust in four European cohort studies after testing whether the number of factors extracted, treatment of input variables, or energy adjustment affected the factor solution. More research of this nature is clearly needed to better understand whether and how treatment of input variables affects empirically derived pattern solutions.

Final subjective decisions made by the investigator include determining which pattern solution to report, name, and analyze. It was previously noted that the investigator must pre-specify the number of derived patterns before running the factor or cluster analysis. While this is true, most investigators run the procedures for a range of solutions (e.g., from 2 to 20) to see which solution makes the most sense, both internally and externally in light of other research reports. One also must make additional decisions, including choosing cutpoints for eigenvalues and factor loadings and method of rotation, if any, in factor analysis and choosing the clustering method (Ward's, K-Means, or another method) in cluster analysis. Scree plots are also often used, especially in factor analysis, to guide the selection process. Ultimately a decision must be made as to which solution is most "meaningful," and that solution will be the one on which the report is based. Occasionally, not all patterns are reported or discussed even though they were part of the original solution. In several studies, more factors are derived than reported.[29,52,64,65] It is important to remember that, whereas the first factor is usually quite similar in many solutions since it first explains the greatest amount of variance in the data, factors 2 and 3 from a 6-factor solution may be different from factors 2 and 3 from a 3-factor solution after rotation. Therefore, authors should report both the solution selected as well as the patterns presented, especially if these numbers differ.

Lastly, naming the derived patterns also involves subjectivity. Whether derived from factor or cluster analysis, both methods named patterns using either quantitative or qualitative criteria. Of note are the many different labels that were used to denote a healthy eating pattern, ranging from Prudent to Healthful to Nutrition consciousness. Each of these so-named healthy patterns differed from each other to greater or lesser degrees in their nutrient and food composition. For example, a Healthy pattern in one study was high in vegetables, fruits, fish, whole grains, and low-fat dairy,[57] while a Healthy pattern in another study was high in brown bread, low-fat spreads, low-fat milk, fruit, and alcohol.[103] Likewise, patterns high in cereals[30,55,88] may have healthful qualities but are not always named as such. It is likely that a number of different patterns exist that may be considered healthy. In another example, many patterns were named "Traditional," and clearly "traditional" foods vary by ethnicity and culture. Since qualitatively naming a pattern does not clearly provide information about food or nutrient composition of the patterns, nor does leaving them unnamed, it may be more useful if patterns are named quantitatively. Regardless of how patterns are named, reporting factor loadings for most if not all of the input variables in factor analysis studies and presenting summary data for each cluster (e.g., means of food groups for each pattern) in cluster studies is important so that patterns may be easily compared across studies.

Another important area of consideration is whether patterns should be separately derived for men and women or whether it is sufficient to derive patterns in a mixed population and then explore relations with sex through the analysis. Food patterns may differ by sex[17,80,109,115] due to gender-influenced eating behaviors. For example, perhaps women eat a more healthful diet or are more likely to be restrained eaters and to reduce calories due to concern about body shape and body image. In this review, the majority of studies containing both men and women included both together in one pattern analysis. In studies that derived patterns separately for men and women,[43,73,80,102,103,110,111,117] many of the patterns were similar across sex groups. In one study,[102] the clusters derived were exactly the same for men and women. Some studies first performed analyses stratified by sex but then collapsed men and women together when patterns appeared similar.[29,90] Comparisons across studies and populations indicate that certain patterns do seem to be present in both men and women, including those that are healthful, less healthful, high in alcohol, and high in sweets, as previously discussed. Reproducibility between men and women within a study and between studies suggest similar patterns among men and women, but additional research is still needed in this area.

Comparison of Factor and Cluster Analysis Methods

This section of the Discussion specifically compares and contrasts the factor and cluster analysis methods. Each of the methods will be separately explained, as will implications for analysis. One author has previously described the similarities and differences of factor and cluster analysis.[7] The discussion of the factor and cluster analysis methods here is thus abbreviated and relegated to the purposes of reviewing the literature cited in this article.

Factor analysis groups input variables according to the degree to which they are correlated with each other, thus aggregating dietary data into distinct patterns (factors). Each individual has a factor score for every derived pattern (factor) and this score is a continuous variable. Table 1 indicates that the vast majority of factor analysis studies employed principal components analysis (PCA), orthogonal rotation, and eigenvalues >1. Eigenvalues >1.5[43,80] and >2.0[52] were used in few studies to derive factors, particularly when large numbers of factors have eigenvalues >1.0. Factors were not rotated in one study[37] and were rotated using promax (oblique) rotation in two other studies.[56,72] Not all authors report eigenvalues or whether and how factors were rotated. More research is needed to examine how eigenvalue cutpoints and rotation affects factor solutions. In addition, several studies performed confirmatory factor analysis, in which the plausibility of the derived factor model can be empirically tested.[125]

Because factors are not mutually exclusive, many investigators categorize individuals by their factor scores into mutually exclusive groups. Most often, factors will be divided using quantiles. In one study, a group was created in which individuals in the highest quintile of a Prudent factor and the lowest quintile of a Western factor were defined as the healthiest group for analysis.[65] In an earlier report, Schwerin[2] grouped individuals into mutually exclusive groups by assigning an individual to the factor for which s/he had the highest factor score.

Because factor scores are continuous variables and individuals have scores for each factor, there are several analytical challenges that arise. First, relationships do not hold true for individuals but rather for a food pattern, such that extrapolating from a pattern derived from factor analysis to overall individual behavior is not clear. An individual with a high score on one factor may have high or low scores on others. Their overall dietary pattern is therefore a combination of factors that is not readily interpretable. (We previously refer to factors as food patterns and not dietary patterns because a person's overall dietary pattern is really represented by his or her scores across all derived factors.[30]) Because an individual has a score for each factor and often scores highly on more than one factor, investigators using factor analysis should be cautious about assigning individuals to non-overlapping groups in light of the way factor scoring works within a subject.

Cluster analysis groups individuals into mutually exclusive categories (clusters) and may use several different methods to do so. Table 2 indicates that most studies used either the K-Means or Ward's methods to derive clusters. One study used the partitioning around medoids (PAM) method.[85] Ward's method is hierarchical and is designed to optimize the minimum variance within clusters (i.e., the within-groups sum of squares or the error sum of squares), while K-means is non-hierarchical and iterative and is designed to create the most distance between clusters.[126] The K-means method derives clusters based upon the mean intakes, or centroids, of the input variables, whereas the PAM algorithm uses the median, or medoid, to create clusters. Thus, the PAM method is less sensitive to outliers than the K-means method.[85] Few studies have examined which of these methods is better, although it is more common to use non-hierarchical (K-means) clustering when there is a large number of input variables being entered into the analysis. Campain et al.[85] compared solutions from both the K-Means and PAM methods and concluded that the PAM algorithm produced clusters that were more consistent with literature published in their field (of dental epidemiology). Few direct comparisons of clustering methods have been conducted and it is not clear from looking at the studies in Table 2 which method is preferable for use with dietary variables.

Compared to factors, clusters are arguably easier to handle in the analysis since they are mutually exclusive and continuous. That individuals have scores for all of the derived factors makes the concept of factor scores less intuitive than an individual belonging to a specific dietary pattern. It is worthy of reiteration that an individual's dietary pattern is really represented by their scores on all factors, not simply their score on one particular factor. Analyses using factors are true only for particular food combinations and cannot be translated to an individual without further analysis. On the other hand, if factors were rotated orthogonally (and are therefore independent from each other), all factors can be included together in regression analyses without affecting the results of any one factor, whereas clusters can only be analyzed using a reference group.

Despite clear differences in approaches and interpretations, there is some evidence that underlying eating patterns are revealed by either method. For example, the healthy pattern and its variants were derived in studies using both factor and cluster analysis and many other similar patterns were also observed. Only two studies of which we are aware have specifically compared the cluster and factor analysis methods in the same dataset.[29,31] Costacou et al.[29] derived four factors (Mediterranean-type, Vegetarian-type, Sweets-focused, and Western-type) and three clusters (clusters were unnamed, but cluster A was similar to factor 1, while clusters B and C were low in Mediterranean components and high in sweets and alcohol, respectively) among Greeks. The mean factor score for the Mediterranean-type pattern was higher in Cluster A than in Clusters B and C (combined into one cluster, BC).[29] Another set of studies suggests that patterns derived from either factor analysis[30] or cluster analysis[90] are comparable and are similarly associated with plasma lipids.[31]

Conclusion and Future Research Directions

In response to a commentary on factor analysis and the search for objectivity,[123] Slattery et al. concluded that the soundness of using factor analysis to identify eating patterns will be better understood when more epidemiologists have begun to use the method.[118] An unstated objective, if not hope, was that this review of the literature may allow us to make inferences about the soundness of patterning methods. Even in the 5 years since Slattery's comment, many more epidemiologists have used patterning methods to further analyze dietary intake. Essentially, whether a method of dietary assessment is useful, or "sound," is generally dependent on validity and reproducibility. Ultimately important is whether the methods used to empirically derive eating patterns can reliably predict disease in diverse populations. How sound is eating pattern methodology? Is the concept of eating patterns of sound use to nutritional epidemiology?

Subjectivity compromises soundness. That subjectivity exists, however, is not a reason to abandon empirical, data-driven methods. (All scientific research involves subjectivity, to greater and lesser degrees, conscious or unconscious.) Both improved reporting of statistical methods and additional methodological studies will help to decrease the impact of subjectivity using eating patterns. Some suggestions for future research directions are discussed below to help move this field forward.

First, investigators must report in as much detail as possible how all decisions were made, beginning with the grouping of the dietary data and ending with how patterns are presented and analyzed. The more these procedures are clearly explained, the better we are able to understand if and how these decisions affect study findings. Second, more validation studies that include a reproducibility component are warranted for both methods. Third, methodological work is needed in several areas. Some outstanding research questions include: Does treatment of the input variable (e.g., servings per day, percent energy contribution, standardized, energy adjusted) affect the pattern solution? Which is the best method (Ward's, K-means, PAM, or another method) to derive clusters? What should be the direction of rotation and eigenvalue cutpoint in the factor analysis procedure; are rotation and eigenvalues important? Is overlap of factor scores within an individual meaningful? When should confirmatory or exploratory factor analysis be used? Are methodological differences between cluster and factor analysis meaningful? Is one method more useful than the other? Would combining factor and cluster analysis procedures provide more meaningful patterns than using one procedure alone? What is the impact of measurement error on empirically derived patterns? Some of these questions were also identified at an international workshop on eating patterns,[127] and a few studies discussed herein have begun to address these questions. Finally, more studies are clearly needed to examine gene-diet interactions, and the eating pattern approach can play a role in this research.[92] In time, greater consensus may exist to guide decision-making in patterning methods, which should help to standardize subjective decisions, thus decreasing the subjectivity and increasing the reproducibility of these methods.

Whether a method is sound has practical meaning beyond research laboratories only if the concept measured provides useful information about diet. This review paper describes the food and nutrient composition of patterns, demographic and socioeconomic characteristics, and pattern-disease relations. Additional data on reproducibility and validity and associations with disease and biomarkers suggest that eating patterns are a valid measure of dietary intake and are biologically meaningful. Whereas some patterns are reproduced across populations, other patterns are culture specific due to ethnic and geographic differences in food habits, preferences, and availability, thereby explaining natural variation. Likewise, reproducibility of patterns over time may either represent instability of the measurements or actual changes in dietary intakes. In other words, there is likely natural and expected variation in eating patterns both within and between individuals that create irreproducibility but do not compromise the validity of the methods. Importantly, empirically derived eating patterns provide a unique look at dietary data, which may enhance our understanding of actual eating behavior and provide a better basis from which to provide dietary advice.

One major finding of this review is that many variations of a "Healthy" or "Prudent" pattern have been seen across different populations, which, although having somewhat different food and nutrient composition, have been associated with less disease, smaller BMI, less mortality, and fewer cancers. Furthermore, this finding supports results using single nutrients or foods, which suggests that, in general, a healthy diet is one that is heavily based in plant foods such as vegetables, fruit, whole grains, nuts, and legumes. There is some evidence that healthy eating patterns may also include fish, reduced-fat dairy products, and alcohol. Individual food preferences (and other factors affecting food choices) likely create the variation we see in healthful eating patterns. Thus, an individual may select from a variety of different healthful foods and food groups to create a healthy diet.

As we improve our understanding of the basic composition of a healthy eating pattern, it is also useful to define that pattern a priori to measure dietary behavior against an appropriate reference. Indeed, empirically derived patterns can be used to further elucidate the food and nutrient composition of a healthy eating pattern, which can be further defined quantitatively and proportionally for use in a theoretically defined pattern, that is, a diet index. Patterns that have shown reproducibility, whether Healthy, Western, Alcohol/Drinker, and Sweets/Desserts, can help in generating or refining hypotheses concerning diet and disease that can be tested using other methods in nutritional epidemiology. Jacobs and Steffen recently proposed that a "top down" method starting with foods or food patterns can provide clues on where to look for specific dietary components, thereby feeding into the "bottom up" method that identifies and characterizes individual foods constituents.[128] In addition, this review suggests that there are many variants of eating patterns that may be considered healthful, which is helpful in providing dietary advice to individuals. Therefore, theoretically defined patterns, empirically derived patterns, and traditional epidemiology are complementary approaches that can be used together to better understand diet-disease relations and improve the public health.

In conclusion, reviewing the literature on empirically derived patterns and thinking specifically about methodological issues remind us that similar issues arose when the food-frequency questionnaire was establishing itself as an "alternative" dietary assessment method to dietary records and 24-hour recalls. In his text on nutritional epidemiology, Willett notes that maximal information is obtained in nutritional epidemiology when different methods are used.[129] We believe that, as was true for the FFQ, more attention to methodological issues, including reproducibility and validity of eating patterns, energy adjustment, and associations with biomarkers and disease, will clarify the utility of eating pattern methods in nutritional epidemiology. Although the research considered herein has created a meaningful body of literature, refining both the factor and cluster analysis methods will help to further establish eating patterns as a sound dietary assessment method.

Acknowledgments

This material is based upon work supported by the U.S. Department of Agriculture, under agreement No. 58-1950-4-401. Any opinions, findings, conclusion, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Agriculture. The authors would like to thank peer reviewers who took the time to provide helpful comments and suggestions on all of our manuscripts involving eating patterns along with the many investigators with whom conversations about this topic have stimulated our thinking. Despite our best attempts to include all relevant papers, we sincerely apologize to any investigators whose studies were inadvertently omitted from this review.

1. Schwerin HS, Stanton JL, Riley AM Jr, et al. Food eating patterns and health: a reexamination of the Ten-State and HANES I surveys. *Am J Clin Nutr.* 1981;34:568–580.
2. Schwerin HS, Stanton JL, Smith JL, Riley AM Jr, Brett BE. Food, eating habits, and health: a further examination of the relationship between food eating patterns and nutritional health. *Am J Clin Nutr.* 1982;35:1319–1325.
3. Kennedy ET, Ohls J, Carlson S, Fleming K. The Healthy Eating Index: design and applications. *J Am Diet Assoc.* 1995;95:1103–1108.
4. Haines PS, Siega-Riz AM, Popkin BM. The Diet Quality Index revised: a measurement instrument for populations. *J Am Diet Assoc.* 1999;99:697–704.
5. McCullough ML, Feskanich D, Stampfer MJ, et al. Diet quality and major chronic disease risk in men and women: moving toward improved dietary guidance. *Am J Clin Nutr.* 2002;76:1261–1271.
6. Kant AK. Indexes of overall diet quality: a review. *J Am Diet Assoc.* 1996;96:785–791.
7. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol.* 2002; 13:3–9.
8. Abusabha R, Hsieh KH, Achterberg C. Dietary fat reduction strategies used by a group of adults aged 50 years and older. *J Am Diet Assoc.* 2001; 101:1024–1030.
9. Keenan DP, Achterberg C, Kris-Etherton PM, Abusabha R, von Eye A. Use of qualitative and quantitative methods to define behavioral fat-reduction strategies and their relationship to dietary fat reduction in the Patterns of Dietary Change Study. *J Am Diet Assoc.* 1996;96:1245–1250, 1253, 1251–1252.
10. Kristal AR, Shattuck AL, Henry HJ. Patterns of dietary behavior associated with selecting diets low in fat: reliability and validity of a behavioral

approach to dietary assessment. *J Am Diet Assoc.* 1990;90:214–220.
11. Slattery ML, Edwards SL, Boucher KM, Anderson K, Caan BJ. Lifestyle and colon cancer: an assessment of factors associated with risk. *Am J Epidemiol.* 1999;150:869–877.
12. Llopis A, Morales M, Rodriguez R. Digestive cancer in relation to diet in Spain. *J Environ Pathol Toxicol Oncol.* 1992;11:169–175.
13. Wright L, Murcer S, Adams K, Welch S, Paris D. The factor analytic structure of seven physical CHD risk factors: a replication study. *J Clin Psychol.* 1994;50:216–219.
14. Kulbok PA, Carter KF, Baldwin JH, Gilmartin MJ, Kirkwood B. The Multidimensional Health Behavior Inventory. *J Nurs Meas.* 1999;7:177–195.
15. Wright L, Carbonari J, Voyles W. A factor analytic study of physical risk variables for CHD. *J Clin Psychol.* 1992;48:165–170.
16. Koifman S, Koifman RJ. Stomach cancer incidence in Brazil: an ecologic study with selected risk factors. *Cad Saude Publica.* 1997;13:85–92.
17. Patterson RE, Haines PS, Popkin BM. Health lifestyle patterns of U.S. adults. *Prev Med.* 1994;23: 453–460.
18. Burke V, Milligan RA, Beilin LJ, et al. Clustering of health-related behaviors among 18-year-old Australians. *Prev Med.* 1997;26:724–733.
19. Hagoel L, Ore L, Neter E, sSilman Z, Rennert G. Clustering women's health behaviors. *Health Educ Behav.* 2002;29:170–182.
20. Milligan RA, Thompson C, Vandongen R, Beilin LJ, Burke V. Clustering of cardiovascular risk factors in Australian adolescents: association with dietary excesses and deficiencies. *J Cardiovasc Risk.* 1995;2:515–523.
21. Zhuo XG, Watanabe S. Factor analysis of digestive cancer mortality and food consumption in 65 Chinese counties. *J Epidemiol.* 1999;9:275–284.
22. Hollis JF, Carmody TP, Connor SL, Fey SG, Matarazzo JD. The Nutrition Attitude Survey: associations with dietary habits, psychological and physical well-being, and coronary risk factors. *Health Psychol.* 1986;5:359–374.
23. Satia JA, Kristal AR, Patterson RE, Neuhouser ML, Trudeau E. Psychosocial factors and dietary habits associated with vegetable consumption. *Nutrition.* 2002;18:247–254.
24. Haines P, Popkin BM, Guilkey DK. Methods of patterning eating behaviors of American women. *J Nutr Educ.* 1990;22:124–132.
25. Schlundt DG, Hargreaves MK, Buchowski MS. The Eating Behavior Patterns Questionnaire predicts dietary fat intake in African American women. *J Am Diet Assoc.* 2003;103:338–345.
26. Haines PS, Hungerford DW, Popkin BM, Guilkey DK. Eating patterns and energy and nutrient intakes of US women. *J Am Diet Assoc.* 1992;92: 698–704, 707.
27. Schlundt DG, Taylor D, Hill JO, et al. A behavioral taxonomy of obese female participants in a weight-loss program. *Am J Clin Nutr.* 1991;53: 1151–1158.
28. Kristeller JL, Rodin J. Identifying eating patterns in male and female undergraduates using cluster analysis. *Addict Behav.* 1989;14:631–642.
29. Costacou T, Bamia C, Ferrari P, Riboli E, Trichopoulos D, Trichopoulou A. Tracing the Mediterranean diet through principal components and cluster analyses in the Greek population. *Eur J Clin Nutr.* 2003;57:1378–1385.
30. Newby PK, Muller D, Hallfrisch J, Andres R, Tucker KL. Food patterns measured by factor analysis and anthropometric changes in adults. *Am J Clin Nutr.* 2004. [in press]
31. Newby PK, Muller D, Tucker K. Associations of empirically derived eating patterns with plasma lipid biomarkers: comparison of factor and cluster analysis methods. *Am J Clin Nutr.* 2004. [in press]
32. Tseng M, Breslow RA, DeVellis RF, Ziegler RG. Dietary patterns and prostate cancer risk in the National Health and Nutrition Examination Survey Epidemiological Follow-up Study cohort. *Cancer Epidemiol Biomarkers Prev.* 2004;13:71–77.
33. Balder HF, Virtanen M, Brants HA, et al. Common and country-specific dietary patterns in four European cohort studies. *J Nutr.* 2003;133:4246–4251.
34. Kerver JM, Yang EJ, Bianchi L, Song WO. Dietary patterns associated with risk factors for cardiovascular disease in healthy US adults. *Am J Clin Nutr.* 2003;78:1103–1110.
35. Markaki I, Linos D, Linos A. The influence of dietary patterns on the development of thyroid cancer. *Eur J Cancer.* 2003;39:1912–1919.
36. Masaki M, Sugimori H, Nakamura K, Tadera M. Dietary patterns and stomach cancer among middle-aged male workers in Tokyo. *Asia Pac J Cancer Prev.* 2003;4:61–66.
37. Sanchez-Villegas A, Delgado-Rodriguez M, Martinez-Gonzalez MA, De Irala-Estevez J. Gender, age, socio-demographic and lifestyle factors associated with major dietary patterns in the Spanish Project SUN (Seguimiento Universidad de Navarra). *Eur J Clin Nutr.* 2003;57:285–292.
38. Schulze MB, Hoffmann K, Kroke A, Boeing H. 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.* 2003;158:365–373.
39. Togo P, Heitmann BL, Sorensen TI, Osler M. Consistency of food intake factors by different dietary assessment methods and population groups. *Br J Nutr.* 2003;90:667–678.
40. Togo P, Osler M, Sorensen TI, Heitmann BL. A longitudinal study of food intake patterns and obesity in adult Danish men and women. *Int J Obes Relat Metab Disord.* 2004;28:583–593.
41. van Dam RM, Grievink L, Ocke MC, Feskens EJM. Patterns of food consumption and risk factors for cardiovascular disease in the general Dutch population. *Am J Clin Nutr.* 2003;77:1156–1163.
42. Gerdes LU, Bronnum-Hansen H, Osler M, Madsen M, Jorgensen T, Schroll M. Trends in lifestyle coronary risk factors in the Danish MONICA population 1982–1992. *Public Health.* 2002;116:81–88.
43. Handa K, Kreiger N. Diet patterns and the risk of renal cell carcinoma. *Public Health Nutr.* 2002;5: 757–767.
44. Mishra G, Ball K, Arbuckle J, Crawford D. Dietary

patterns of Australian adults and their association with socioeconomic status: results from the 1995 National Nutrition Survey. *Eur J Clin Nutr.* 2002;56: 687–693.
45. Sichieri R. Dietary patterns and their associations with obesity in the Brazilian city of Rio de Janeiro. *Obes Res.* 2002;10:42–48.
46. Fung TT, Willett WC, Stampfer MJ, Manson JE, Hu FB. Dietary patterns and the risk of coronary heart disease in women. *Arch Intern Med.* 2001;161: 1857–1862.
47. Guinot C, Latreille J, Malvy D, et al. Use of multiple correspondence analysis and cluster analysis to study dietary behaviour: food consumption questionnaire in the SU.VI.MAX. cohort. *Eur J Epidemiol.* 2001;17:505–516.
48. Hulshof KF, Valsta LM, Welten DC, Lowik MR. Analytical approaches to food-based dietary guidelines in the European setting. *Public Health Nutr.* 2001;4:667–672.
49. McCann SE, Marshall JR, Brasure JR, Graham S, Freudenheim JL. 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.* 2001;4:989–997.
50. McCann SE, Weiner J, Graham S, Freudenheim JL. Is principal components analysis necessary to characterise dietary behaviour in studies of diet and disease? *Public Health Nutr.* 2001;4:903–908.
51. Ogura M, Yamamoto T, Morita M, Watanabe T. A case-control study on food intake of patients with recurrent aphthous stomatitis. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2001;91:45–49.
52. Osler M, Heitmann BL, Gerdes LU, Jorgensen LM, Schroll M. Dietary patterns and mortality in Danish men and women: a prospective observational study. *Br J Nutr.* 2001;85:219–225.
53. Osler M, Helms Andreasen A, Heitmann B, et al. Food intake patterns and risk of coronary heart disease: a prospective cohort study examining the use of traditional scoring techniques. *Eur J Clin Nutr.* 2002;56:568–574.
54. Palli D, Russo A, Decarli A. Dietary patterns, nutrient intake and gastric cancer in a high-risk area of Italy. *Cancer Causes Control.* 2001;12:163–172.
55. Schulze MB, Hoffmann K, Kroke A, Boeing H. 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.* 2001;85:363–373.
56. Speck BJ, Bradley CB, Harrell JS, Belyea MJ. A food frequency questionnaire for youth: psychometric analysis and summary of eating habits in adolescents. *J Adolesc Health.* 2001;28:16–25.
57. Terry P, Hu FB, Hansen H, Wolk A. Prospective study of major dietary patterns and colorectal cancer risk in women. *Am J Epidemiol.* 2001;154: 1143–1149.
58. Terry P, Suzuki R, Hu FB, Wolk A. A prospective study of major dietary patterns and the risk of breast cancer. *Cancer Epidemiol Biomarkers Prev.* 2001;10:1281–1285.
59. Tseng M, DeVellis RF. Fundamental dietary patterns and their correlates among US whites. *J Am Diet Assoc.* 2001;101:929–932.
60. Maskarinec G, Novotny R, Tasaki K. Dietary patterns are associated with body mass index in multiethnic women. *J Nutr.* 2000;130:3068–3072.
61. Mullen K, Williams R, Hunt K. Irish descent, religion and food consumption in the west of Scotland. *Appetite.* 2000;34:47–54.
62. Tseng M, DeVellis RF, Maurer KR, et al. Food intake patterns and gallbladder disease in Mexican Americans. *Public Health Nutr.* 2000;3:233–243.
63. Williams DE, Prevost AT, Whichelow MJ, Cox BD, Day NE, Wareham NJ. A cross-sectional study of dietary patterns with glucose intolerance and other features of the metabolic syndrome. *Br J Nutr.* 2000;83:257–266.
64. Hu FB, Rimm E, Smith-Warner SA, et al. Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *Am J Clin Nutr.* 1999;69:243–249.
65. Hu FB, Rimm EB, Stampfer MJ, Ascherio A, Spiegelman D, Willett WC. Prospective study of major dietary patterns and risk of coronary heart disease in men. *Am J Clin Nutr.* 2000;72:912–921.
66. Fung TT, Rimm EB, Spiegelman D, et al. Association between dietary patterns and plasma biomarkers of obesity and cardiovascular disease risk. *Am J Clin Nutr.* 2001;73:61–67.
67. van Dam RM, Rimm EB, Willett WC, Stampfer MJ, Hu FB. Dietary patterns and risk for type 2 diabetes mellitus in U.S. men. *Ann Intern Med.* 2002;136: 201–209.
68. Kumagai S, Shibata H, Watanabe S, Suzuki T, Haga H. Effect of food intake pattern on all-cause mortality in the community elderly: a 7-year longitudinal study. *J Nutr Health Aging.* 1999;3:29–33.
69. van den Bree MB, Eaves LJ, Dwyer JT. Genetic and environmental influences on eating patterns of twins aged >/=50 y. *Am J Clin Nutr.* 1999;70:456–465.
70. Armstrong RW, Imrey PB, Lye MS, Armstrong MJ, Yu MC, Sani S. Nasopharyngeal carcinoma in Malaysian Chinese: salted fish and other dietary exposures. *Int J Cancer.* 1998;77:228–235.
71. Beaudry M, Galibois I, Chaumette P. Dietary patterns of adults in Quebec and their nutritional adequacy. *Can J Public Health.* 1998;89:347–351.
72. Gittelsohn J, Wolever TM, Harris SB, Harris-Giraldo R, Hanley AJ, Zinman B. Specific patterns of food consumption and preparation are associated with diabetes and obesity in a Native Canadian community. *J Nutr.* 1998;128:541–547.
73. Slattery ML, Boucher KM, Caan BJ, Potter JD, Ma KN. Eating patterns and risk of colon cancer. *Am J Epidemiol.* 1998;148:4–16.
74. Whichelow MJ, Prevost AT. Dietary patterns and their associations with demographic, lifestyle and health variables in a random sample of British adults. *Br J Nutr.* 1996;76:17–30.
75. Nolan CC, Gray-Donald K, Shatenstein B, O'Loughlin J. Dietary patterns leading to high fat intake. *Can J Public Health.* 1995;86:389–391.

76. Wolff CB, Wolff HK. Maternal eating patterns and birth weight of Mexican American infants. *Nutr Health.* 1995;10:121–134.
77. Ishimoto H, Nakamura H, Miyoshi T. Epidemiological study on relationship between breast cancer mortality and dietary factors. *Tokushima J Exp Med.* 1994;41:103–114.
78. Barker ME, McClean SI, Strain JJ, Thompson KA. Dietary behaviour and health in Northern Ireland: an exploration of biochemical and haematological associations. *J Epidemiol Commun Health.* 1992; 46:151–156.
79. Barker ME, McClean SI, Thompson KA, Reid NG. Dietary behaviours and sociocultural demographics in Northern Ireland. *Br J Nutr.* 1990;64:319–329.
80. Randall E, Marshall JR, Graham S, Brasure J. Patterns in food use and their associations with nutrient intakes. *Am J Clin Nutr.* 1990;52:739–745.
81. Gex-Fabry M, Raymond L, Jeanneret O. Multivariate analysis of dietary patterns in 939 Swiss adults: sociodemographic parameters and alcohol consumption profiles. *Int J Epidemiol.* 1988;17: 548–555.
82. Iizumi H, Amemiya T. Eleven-year follow-up of changes in individuals' food consumption patterns. *Int J Vitam Nutr Res.* 1986;56:399–409.
83. Iizumi H, Amemiya T. Effect of food consumption pattern on total serum cholesterol level: a methodological approach. *J Nutr Sci Vitaminol (Tokyo).* 1986;32:205–218.
84. Ganganna P, Johnson AA. A new nutrient index for measuring nutritional well-being of Indian states. *Int J Vitam Nutr Res.* 1985;55:315–322.
85. Campain AC, Morgan MV, Evans RW, et al. Sugar-starch combinations in food and the relationship to dental caries in low-risk adolescents. *Eur J Oral Sci.* 2003;111:316–325.
86. Correa Leite ML, Nicolosi A, Cristina S, Hauser WA, Pugliese P, Nappi G. Dietary and nutritional patterns in an elderly rural population in Northern and Southern Italy: (I). A cluster analysis of food consumption. *Eur J Clin Nutr.* 2003;57:1514–1521.
87. Diehr P, Beresford SA. The relation of dietary patterns to future survival, health, and cardiovascular events in older adults. *J Clin Epidemiol.* 2003;56: 1224–1235.
88. Lin H, Bermudez OI, Tucker KL. Dietary patterns of Hispanic elders are associated with acculturation and obesity. *J Nutr.* 2003;133:3651–3657.
89. Martikainen P, Brunner E, Marmot M. Socioeconomic differences in dietary patterns among middle-aged men and women. *Soc Sci Med.* 2003;56: 1397–1410.
90. Newby PK, Muller D, Hallfrisch J, Qiao N, Andres R, Tucker KL. Dietary patterns and changes in body mass index and waist circumference in adults. *Am J Clin Nutr.* 2003;77:1417–1425.
91. Shimizu K, Takeda S, Noji H, et al. Dietary patterns and further survival in Japanese centenarians. *J Nutr Sci Vitaminol (Tokyo).* 2003;49:133–138.
92. Tsai YY, McGlynn KA, Hu Y, et al. Genetic susceptibility and dietary patterns in lung cancer. *Lung Cancer.* 2003;41:269–281.
93. Chen H, Ward MH, Graubard BI, et al. Dietary patterns and adenocarcinoma of the esophagus and distal stomach. *Am J Clin Nutr.* 2002;75:137–144.
94. Rasanen M, Lehtinen JC, Niinikoski H, et al. Dietary patterns and nutrient intakes of 7-year-old children taking part in an atherosclerosis prevention project in Finland. *J Am Diet Assoc.* 2002;102: 518–524.
95. Tucker KL, Chen H, Hannan MT, et al. Bone mineral density and dietary patterns in older adults: the Framingham Osteoporosis Study. *Am J Clin Nutr.* 2002;76:245–252.
96. Haveman-Nies A, Tucker KL, de Groot LC, Wilson PW, van Staveren WA. Evaluation of dietary quality in relationship to nutritional and lifestyle factors in elderly people of the US Framingham Heart Study and the European SENECA study. *Eur J Clin Nutr.* 2001;55:870–880.
97. Millen BE, Quatromoni PA, Copenhafer DL, Demissie S, O'Horo CE, D'Agostino RB. Validation of a dietary pattern approach for evaluating nutritional risk: the Framingham Nutrition Studies. *J Am Diet Assoc.* 2001;101:187–194.
98. Quatromoni PA, Copenhafer DL, Demissie S, et al. The internal validity of a dietary pattern analysis. The Framingham Nutrition Studies. *J Epidemiol Commun Health.* 2002;56:381–388.
99. Millen BE, Quatromoni PA, Nam BH, O'Horo CE, Polak JF, D'Agostino RB. Dietary patterns and the odds of carotid atherosclerosis in women: the Framingham Nutrition Studies. *Prev Med.* 2002;35: 540–547.
100. Quatromoni PA, Copenhafer DL, D'Agostino R B, Millen BE. Dietary patterns predict the development of overweight in women: The Framingham Nutrition Studies. *J Am Diet Assoc.* 2002;102: 1239–1246.
101. Millen BE, Quatromoni PA, Nam BH, et al. Dietary patterns, smoking, and subclinical heart disease in women: opportunities for primary prevention from the Framingham Nutrition Studies. *J Am Diet Assoc.* 2004;104:208–214.
102. Pryer JA, Nichols R, Elliott P, Thakrar B, Brunner E, Marmot M. Dietary patterns among a national random sample of British adults. *J Epidemiol Commun Health.* 2001;55:29–37.
103. Pryer JA, Cook A, Shetty P. Identification of groups who report similar patterns of diet among a representative national sample of British adults aged 65 years of age or more. *Public Health Nutr.* 2001;4:787–795.
104. Greenwood DC, Cade JE, Draper A, Barrett JH, Calvert C, Greenhalgh A. Seven unique food consumption patterns identified among women in the UK Women's Cohort Study. *Eur J Clin Nutr.* 2000; 54:314–320.
105. Wirfalt E, Mattisson I, Gullberg B, Berglund G. Food patterns defined by cluster analysis and their utility as dietary exposure variables: a report from the Malmo Diet and Cancer Study. *Public Health Nutr.* 2000;3:159–173.
106. Wirfalt E, Hedblad B, Gullberg B, et al. Food patterns and components of the metabolic syndrome in men and women: a cross-sectional study within

the Malmo Diet and Cancer cohort. *Am J Epidemiol.* 2001;154:1150–1159.
107. Haveman-Nies A, de Groot LP, van Staveren WA. Snack patterns of older Europeans. *J Am Diet Assoc.* 1998;98:1297–1302.
108. Margetts BM, Thompson RL, Speller V, McVey D. Factors which influence 'healthy' eating patterns: results from the 1993 Health Education Authority health and lifestyle survey in England. *Public Health Nutr.* 1998;1:193–198.
109. Wirfalt AK, Jeffery RW. Using cluster analysis to examine dietary patterns: nutrient intakes, gender, and weight status differ across food pattern clusters. *J Am Diet Assoc.* 1997;97:272–279.
110. Millen BE, Quatromoni PA, Gagnon DR, Cupples LA, Franz MM, D'Agostino RB. Dietary patterns of men and women suggest targets for health promotion: the Framingham Nutrition Studies. *Am J Health Promot.* 1996;11:42–52, 52–53.
111. Schroll K, Carbajal A, Decarli B, et al. Food patterns of elderly Europeans. SENECA Investigators. *Eur J Clin Nutr.* 1996;50:S86–S100.
112. Huijbregts PP, Feskens EJ, Kromhout D. Dietary patterns and cardiovascular risk factors in elderly men: the Zutphen Elderly Study. *Int J Epidemiol.* 1995;24:313–320.
113. Hulshof KF, Wedel M, Lowik MR, et al. Clustering of dietary variables and other lifestyle factors (Dutch Nutritional Surveillance System). *J Epidemiol Commun Health.* 1992;46:417–424.
114. Tucker KL, Dallal GE, Rush D. Dietary patterns of elderly Boston-area residents defined by cluster analysis. *J Am Diet Assoc.* 1992;92:1487–1491.
115. Boeing H, Kleine U, Hendrichs A, Oltersdorf U, Bodenstedt A. Strategies for analysing nutritional data for epidemiological purposes—food patterns by means of cluster-analytical purposes. *Nutr Rep Int.* 1989;40:189–198.
116. Farchi G, Mariotti S, Menotti A, Seccareccia F, Torsello S, Fidanza F. Diet and 20-y mortality in two rural population groups of middle-aged men in Italy. *Am J Clin Nutr.* 1989;50:1095–1103.
117. Akin JS, Guilkey DK, Popkin BM, Fanelli MT. Cluster analysis of food consumption patterns of older Americans. *J Am Diet Assoc.* 1986;86:616–624.
118. Slattery ML, Boucher KM. The senior authors' response: Factor analysis as a tool for evaluating eating patterns. *Am J Epidemiol.* 1998;148:20–21.
119. Togo P, Osler M, Sorensen TI, Heitmann BL. Food intake patterns and body mass index in observational studies. *Int J Obes Relat Metab Disord.* 2001;25:1741–1751.
120. Vega-Lopez S, Conde-Knape K, Vidal-Quintanar RL, Shachter NS, Fernandez ML. Sex and hormonal status influence the effects of psyllium on lipoprotein remodeling and composition. *Metabolism.* 2002;51:500–507.
121. Luz Fernandez M, West KL, Roy S, Ramjiganesh T. Dietary fat saturation and gender/hormonal status modulate plasma lipids and lipoprotein composition(1). *J Nutr Biochem.* 2001;12:703–710.
122. Cos E, Ramjiganesh T, Roy S, Yoganathan S, Nicolosi RJ, Fernandez ML. Soluble fiber and soybean protein reduce atherosclerotic lesions in guinea pigs. Sex and hormonal status determine lesion extension. *Lipids.* 2001;36:1209–1216.
123. Martinez ME, Marshall JR, Sechrest L. Invited commentary: Factor analysis and the search for objectivity. *Am J Epidemiol.* 1998;148:17–19.
124. Jacques PF, Tucker KL. Are dietary patterns useful for understanding the role of diet in chronic disease? *Am J Clin Nutr.* 2001;73:1–2.
125. Kim J-O, Mueller C. *Factor Analysis: Statistical Methods and Practical Issues.* Newbury Park, CA: Sage Publications, Inc.; 1978.
126. Aldenderfer MS, Blashfield RK. *Cluster Analysis.* Newbury Park, CA: Sage Publications, Inc.; 1984.
127. Hoffmann K, Schulze M, Boeing H, Altenburg H. Dietary patterns: report of an international workshop. *Public Health Nutr.* 2000;5:89–90.
128. Jacobs DR Jr, Steffen LM. Nutrients, foods, and dietary patterns as exposures in research: a framework for food synergy. *Am J Clin Nutr.* 2003;78:508S–513S.
129. Willett WC. *Nutritional Epidemiology.* 2nd ed. New York: Oxford University Press; 1998.