Current Research

Comparative Strategies for Using Cluster Analysis to Assess Dietary Patterns

REGAN L. BAILEY, MS, RD; MELISSA D. GUTSCHALL, MS, RD; DIANE C. MITCHELL, MS, RD; CARLA K. MILLER, PhD, RD; FRANK R. LAWRENCE, PhD; HELEN SMICIKLAS-WRIGHT, PhD

ABSTRACT

Objectives To characterize dietary patterns using two different cluster analysis strategies.

Design In this cross-sectional study, diet information was assessed by five 24-hour recalls collected over 10 months. All foods were classified into 24 food subgroups. Demographic, health, and anthropometric data were collected via home visit.

Subjects One hundred seventy-nine community-dwelling adults, aged 66 to 87 years, in rural Pennsylvania.

Statistical analysis Cluster analysis was performed.

Results The methods differed in the food subgroups that clustered together. Both methods produced clusters that had significant differences in overall diet quality as assessed by Healthy Eating Index (HEI) scores. The clusters with higher HEI scores contained significantly higher amounts of most micronutrients. Both methods consistently clustered subgroups with high energy contribution (eg, fats and oils and dairy desserts) with a lower HEI score. Clusters resulting from the percent energy method were less likely to differentiate fruit and vegetable subgroups. The higher diet quality dietary pattern derived from the number of servings method resulted in more favorable weight status.

Conclusions Cluster analysis of food subgroups using two different methods on the same data yielded similarities and dissimilarities in dietary patterns. Dietary patterns characterized by the number of servings method of analysis provided stronger association with weight status and was more sensitive to fruit and vegetable intake with regard to a more healthful dietary pattern within this sample. Public health recommendations should evaluate the methodology used to derive dietary patterns.

J Am Diet Assoc. 2006;106:1194-1200.

R. L. Bailey is a National Institutes of Aging pre-doctoral research fellow, M. D. Gutschall is a graduate research assistant, D. C. Mitchell is the coordinator of the Diet Assessment Center, F. R. Lawrence is an assistant professor in Human Development and Family Studies, and H. Smiciklas-Wright is a professor in the Nutritional Sciences Department, all at The Pennsylvania State University, University Park. C. K. Miller is the associate director for the Penn State Diabetes Research Center, University Park.

Address correspondence to: Regan L. Bailey, MS, RD, Department of Nutritional Sciences, The Pennsylvania State University, 5 Henderson Bldg, University Park, PA 16801. E-mail: rkb119@psu.edu

0002-8223/06/10608-0003$32.00/0
doi: 10.1016/j.jada.2006.05.012

Epidemiologic studies of diet and health increasingly focus on dietary patterns rather than nutrients (1-5). Nutrients and foods can exhibit synergistic and inhibitory properties that may attenuate single nutrient and health relationships (6). Dietary patterns reflect whole foods and/or combinations of consumption, temporal distribution of intake, and habitual patterns (eg, snacking and food preparation methods). Examination of the totality of dietary patterns provides a more accurate description of actual dietary exposure. Dietary pattern analysis is an ideal tool to identify those who may be at nutritional risk for appropriate intervention (7).

Common methodology for assessing dietary patterns is through data reduction techniques, such as factor or cluster analysis (8). Factor analysis creates food groupings based on correlations of dietary intake; this is beneficial in determining frequency of food types consumed (9). Cluster analysis creates latent variables of people with similar mean dietary intakes (6). Cluster analysis is useful in nutrition studies because it can be used to create groups of people with homogenous dietary patterns. These clusters can then be treated as independent variables for further analysis of associations between dietary patterns and markers of disease risk or health outcomes.

Various methods can be employed to conduct cluster analysis (3,5,7,9-22). We previously reported on cluster analysis using food servings, a novel approach in dietary pattern analysis that is consistent with the *Dietary Guidelines for Americans* (23). However, many published reports employ percent energy contribution from food subgroups as the clustering variable (7,10-12). Number of servings is an absolute number strategy whereas the percent energy method is a proportion; that is, a high number of fat servings does not necessarily mean lower intake of other foods. On the other hand, a high percentage of energy from fat from one or more groups would be accompanied by a lower percent energy from other subgroups. This is particularly salient for fruits and vegetables. Given their low overall contribution to daily energy intake, using a percent energy approach may not be sensitive enough to discern real differences for some fruit and vegetable subgroup intakes. Dietary recommendations of fruits and vegetables are given in terms of servings (23,24). Number of servings of fruits and vegetables has been associated with lower incidence of certain cancers (1,25,26), cardiovascular disease (27-29), and overall higher quality diets (16).

Given that there is no gold-standard technique, Kant (30) suggests a need to conduct comparative evaluations of dietary patterns derived from different methods in the same cohort with similar confounders. Conducting analytic methodology of cluster analysis would result in dietary patterns of greater confidence. Thus, our study was an explorative methodological statistical analysis of dietary patterns determined by cluster analysis based on two different strategies: number of servings from food subgroups and percent energy contribution from food subgroups using the same sample of older adults. We hypothesized that quantifying the same foods in these two different techniques using the same cohort could result in inconsistent dietary patterns.

METHODS

Subject Recruitment, Study Design, and Data Collection

Data were obtained from a cohort from the Geisinger Rural Aging Study, details of which have been previously published (31). The Geisinger Rural Aging Study is a longitudinal study of more than 20,000 rural older adults in Pennsylvania enrolled within a Medicare-managed health maintenance organization. All participants were older than age 65 years at enrollment. The study protocol was approved by the human investigation review boards at both the Pennsylvania State University and the Geisinger Health System.

Subject Selection. From the entire Geisinger Rural Aging Study population, 944 people were randomly contacted to participate in an intensive cross-sectional research study. Of the 797 reached via telephone, 210 agreed to a home visit at which time anthropometric, health history, and demographic data and a blood sample were collected. Participants (n=10) were excluded from participation if they were clinically depressed or had functional limitation that precluded their participation. Depression was assessed with the Geriatric Depression Scale (32) and cognitive function was determined with the Mini-Mental State examination (33). Complete data for this report were available from 179 participants; the study participants were not significantly different from the larger Geisinger Rural Aging Study population with regard to demographic or anthropometric characteristics but were less likely to have a poor appetite or need assistance with activities of daily living (31).

Anthropometry. At the home visit, subjects were weighed with normal clothing but without shoes using a portable digital scale (UC300, A&D Engineering, Miltpas, CA) and height was obtained by a stadiometer (Infant/Child/Adult Height Measuring Board, Shorr Productions, Olney, MD). Waist circumference was determined with a flexible nonelastic measuring tape just above the iliac crest. Body mass index (BMI) was calculated as kg/m^2.

Dietary Intake

Five 24-hour dietary recalls were collected by telephone over a 10-month time period using a multiple-pass technique by trained interviewers at the Pennsylvania State University Diet Assessment Center (34). Diet recalls were conducted on unannounced, random, nonconsecutive days throughout the 10 months. The average of the 5 days of dietary intake data was used for analysis. Dietary data were analyzed for nutrient composition using the Nutrition Data System software (food database version 12A, nutrient database version 28, 1996, Nutrition Coordinating Center, Minneapolis, MN). The daily food intakes were categorized into main groups according to the Food Guide Pyramid and main groups were further disaggregated into 29 subgroups with similar nutrient compositions (35). Four food groups that were not highly consumed by the cohort were excluded from analysis (eg, pancakes/waffles/French toast, yogurt, dried fruit, and meat substitutes). The consumption of these foods was reported by 6% to 20% of the sample but with a very low frequency (ie, mentioned only once by users during the 5 days of recall data) and thus created skewed distributions. Diet quality was calculated using the Healthy Eating Index (HEI), a tool created by the US Department of Agriculture. The HEI is a measured score of 10 dietary components corresponding to the Food Guide Pyramid (24) and *Dietary Guidelines for Americans* (36). Components one through five measure the degree to which a person's diet conforms to recommendations for the grain, vegetable, fruit, milk, and meat groups, including nonanimal protein sources. The five remaining components measure total fat and saturated fat consumption as a percentage of total food energy intake, cholesterol intake, sodium intake, and dietary variety. Each component has a potential score of zero to 10 with a range of possible HEI scores from zero to 100. HEI score was used as an indicator of diet quality.

Statistical Analysis

Cluster analysis was employed using two different strategies: number of servings from food subgroups and percent daily energy contribution from food subgroups. Proc Fastclus (version 8, 1999, SAS Institute, Cary, NC), was used to create disjoint clusters based on least-squares estimation. The number of clusters is specified a priori. In this study we specified and examined solutions from two to six clusters. The number of clusters is first specified, followed by the creation of cluster means or centroids. Each subject is classified into a cluster based on the Euclidean distance of data points from the cluster centroid in an iterative process such that the Euclidean distance between a point and the cluster centroid is minimized. Cluster membership is exclusive and dependent on minimizing the Euclidean distance within a cluster and maximizing differences between clusters. Thus, a cluster represents a group of individuals who consume a similar dietary pattern. We chose a two-cluster solution based on examination of scree plots of eigenvalues of the data and examination of plots of the within-cluster sum of squares against the number of clusters. We selected a two-cluster solution because additional clusters were fragments of the two larger clusters. Discriminant function analysis also was performed on the clusters to evaluate reproducibility and stability using Proc Discrim (version 8, 1999, SAS Institute, Cary, NC). This procedure examines the number of correct predictions of group placement with the clusters as the group.

Table 1. Food group intake of community-dwelling adults in rural Pennsylvania aged 66 to 87 years (n=179) by dietary pattern cluster using two methods of analysis[a]

	No. of Servings (n/day)		Energy Contribution (%)	
Nutrient	**Cluster 1 (n=107)**	**Cluster 2 (n=72)**	**Cluster 1 (n=68)**	**Cluster 2 (n=111)**
	← *Least square mean (95% confidence interval)* →			
Bread	**3.15 (2.95-3.34)**	2.22 (1.98-2.46)***	14.95 (13.68-16.25)	13.58 (12.54-14.55)
Cereal	0.52 (0.41-0.64)	**1.06 (0.91-1.20)***	7.84 (6.48-9.17)	8.53 (7.48-9.36)
Pasta, noodles, rice	0.61 (0.47-0.76)	0.64 (0.50-0.82)	**10.08 (8.26-11.90)**	7.21 (5.79-8.63)**
Sweet bread desserts	**0.94 (0.82-1.05)**	0.67 (0.52-0.81)**	7.66 (5.92-9.41)	**15.73 (14.37-17.09)*****
Snacks	0.45 (0.35–0.55)	0.40 (0.28-0.53)	4.67 (3.44-5.90)	**5.27 (4.30-6.22)***
Dark-green, deep-yellow vegetables	0.20 (0.14-0.26)	**0.52 (0.45-0.60)*****	2.01 (1.59-2.42)	1.37 (1.05-1.69)
Other vegetables	0.83 (0.71-0.96)	**1.19 (1.04-1.35)*****	1.70 (1.49-1.94)	1.74 (1.56-1.94)
Starchy vegetables	1.02 (0.91-1.14)	0.87 (0.72-1.01)	**9.44 (8.51-10.39)**	7.15 (6.41-7.88)***
Vegetable soup/sauce/juice	0.23 (0.17-0.28)	0.23 (0.16-0.30)	**3.94 (3.13-4.74)**	1.93 (1.31-2.56)**
Citrus, melon, berries	0.27 (0.19-0.36)	**0.47 (0.36-0.58)*****	2.75 (2.34-3.38)	2.54 (2.06-3.03)
Fruit juice	0.32 (0.23-0.41)	**0.58 (0.47-0.69)*****	3.38 (2.60-4.17)	3.46 (2.84-4.07)
Other fruit	0.76 (0.63-0.90)	**1.33 (1.16-1.50)*****	7.05 (5.90-8.18)	**9.34 (8.44-10.23)*****
Milk	0.83 (0.67-1.00)	**1.39 (1.19-1.60)*****	6.82 (5.62-8.02)	7.83 (6.89-8.77)
Cheese	0.30 (0.24-0.37)	0.29 (0.21-0.37)	**7.15 (6.03-6.28)**	5.53 (4.65-6.41)*
Dairy desserts	**0.46 (0.38-0.54)**	0.30 (0.20-0.41)*	**12.45 (10.77-14.12)**	6.24 (4.93-7.55)***
Red meat, beef, lamb	0.58 (0.51-0.66)	0.51 (0.42-0.61)	**13.11 (11.63-14.60)**	10.34 (9.18-11.50)***
Poultry	0.18 (0.14-0.23)	**0.31 (0.25-0.37)*****	**9.87 (8.46-11.28)**	5.05 (3.92-6.15)***
Fish	0.10 (0.05-0.15)	**0.20 (0.13-0.26)****	1.91 (1.00-3.48)	**3.92 (2.95-4.89)*****
Processed meats	**0.50 (0.41-0.59)**	0.22 (0.10-0.33)***	0.99 (.85-1.12)	0.69 (0.58-0.79)***
Nuts	0.09 (0.06-0.12)	0.06 (0.02-0.09)	2.31 (1.26-3.37)	3.42 (2.59-4.24)
Beans	0.09 (0.05-0.14)	**0.21 (0.15-0.27)***	**5.70 (4.03-7.37)**	3.23 (1.92-4.53)*
Eggs	**0.33 (0.26-0.40)**	0.20 (0.11-0.29)*	3.86 (2.91-4.81)	3.36 (2.61-4.10)
Fats and oils	**4.68 (4.26-5.09)**	3.22 (2.71-3.73)**	**10.28 (9.27-11.29)**	8.44 (7.61-9.19)**
Sweets	2.66 (2.28-3.05)	2.09 (1.61-2.57)	7.06 (5.92-8.20)	**9.46 (8.53-10.34)*****
Healthy Eating Index[b]	67 (65-69)	**74 (72-76)*****	67 (65.7-69.7)	**71.0 (69.4-72.5)****

[a]Bold denotes higher value within the cluster. A cluster represents a group of individuals who consume a similar dietary pattern.
[b]The Healthy Eating Index is a measured score of 10 dietary components corresponding to the Food Guide Pyramid and Dietary Guidelines for Americans. Each component has a range of zero to 10, total Healthy Eating Index score range zero to 100. Higher scores indicate more compliance to guidelines.
*$P<0.05$.
**$P<0.01$.
***$P<0.001$.

Differences between the clusters with regard to food subgroups were examined using t tests, χ^2, and analysis of covariance. The covariates included in this analysis were age, sex, tobacco use, and regular alcohol consumption. We used sex as a practical proxy to control for effects related to total energy intake; this removes the effect of total energy consumed when examining nutrient profiles. No daily energy expenditure data were collected. Any non-normal data were log transformed before analysis. The significance level was set at $P<0.05$.

RESULTS

The sample consisted of 81 men and 98 women, aged 66 to 87 years, with a mean age of 73±5 years. Participants were primarily white (99%) and married (73%), and with at least a high school education (80%). More descriptive data on the sample is published elsewhere (31,37). Two clusters were derived by each method of analysis (Table 1).

Regardless of the cluster strategy, there was a cluster with a significantly lower HEI score (Cluster 1) and a cluster with a higher HEI score (Cluster 2). Using the servings method, Cluster 1 was represented by higher amounts of bread, sweet breads, dairy desserts, processed meats, eggs, and fats/oils (Table 1). Using the servings method, Cluster 2 had significantly higher intakes of most fruit/vegetable subgroups, fish, milk, and poultry. With the percent energy method, Cluster 1 had higher amounts of pasta/noodles/rice, starchy vegetables, vegetable soups/sauces/juices, dairy desserts, cheese, most meat subgroups, and fats/oils. Cluster 2, using the percent energy method also had a significantly higher HEI but was less representative of a healthful dietary pattern consistent with dietary guidelines. This cluster resulted in significantly higher amounts of sweet breads (eg, cookies, muffins, and doughnuts), snacks, other fruit, fish, and sweets.

Both of the higher HEI clusters revealed dietary patterns with higher intakes of multiple vitamins and minerals as well as a higher dietary fiber intake (Table 2). Regardless of method employed, Cluster 1 was characterized by higher percentage of daily energy intake from

Table 2. Energy, nutrient intake, waist circumference, and body mass index of community-dwelling adults in rural Pennsylvania aged 66 to 87 years (n=179) by dietary pattern cluster using two methods of analysis[ab]

Nutrients	No. of Servings (n/day)		Energy Contribution (%)	
	Cluster 1 (n=107)	Cluster 2 (n=72)	Cluster 1 (n=68)	Cluster 2 (n=111)
	←Least square mean (95% confidence interval)→			
Total energy	**1,736 (1,657-1,816)**	1,383 (1,286-1,481)***	1,466 (1,362-1,570)	**1,672 (1,591-1,754)****
Carbohydrate (g)	205.68 (199.8-211.6)	227 (220.0-234.7)**	189.7 (173.62-203.33)	**228.9 (217.38-240.54)*****
% Energy from carbohydrate	**57.3 (55.5-59.1)**	52.1 (50.6-53.5)***	51.8 (50.0-53.5)	**55.7 (54.3-57.0)*****
Protein (g)	58.73 (56.89-60.56)	**65.47 (63.19-67.76)*****	59.2 (55.3-63.1)	62.4 (59.4-65.5)
% Energy from protein	**14.87 (14.4-16.8)**	16.84 (16.2-17.5)***	**16.5 (15.84-17.18)**	15.1 (14.62-15.66)***
Fat (g)	**60.22 (58.02-62.41)**	49.81 (47.08-52.53)***	53.5 (48.34-58.80)	57.2 (53.44-61.61)
% Energy from fat	**33.83 (32.6-35.0)**	27.43 (25.9-28.9)***	**32.5 (30.98-34.15)**	30.4 (29.21-31.69)*
% Energy saturated fat	**11.80 (11.2-12.4)**	9.60 (8.9-10.4)***	**11.5 (10.76-12.27)**	10.5 (9.95-11.14)*
Fiber (g)	14.29 (13.4-15.2)	**19.58 (18.4-20.7)*****	14.9 (13.4-16.3)	**17.2 (16.2-18.4)****
Vitamin C (mg)	78.67 (70.1-87.2)	**130.90 (120.3-141.4)*****	93.6 (82.93-106.56)	104.5 (93.47-111.94)
Thiamin (mg)	1.47 (1.4-1.5)	**1.62 (1.5-1.7)***	1.4 (1.28-1.51)	**1.6 (1.53-1.71)****
Magnesium (mg)	221.59 (211.6-231.4)	**293.5 (281.1-305.8)*****	229.83 (210.99-248.67)	**263.11 (248.39-277.83)***
Vitamin K (mg)	2,218.70 (2,136.1-2,301.2)	**2,890 (2,788.2-2,993.7)*****	2,357.94 (2,193.26-2,522.62)	**2,569.47 (2,440.81-2,198.13)***
Zinc (mg)	8.52 (7.8-9.2)	**10.0 (9.1-10.9)****	8.49 (7.57-9.41)	9.44 (8.71-10.1)
Vitamin B-12 (μg)	3.56 (2.85-4.27)	**4.89 (4.01-5.77)***	3.30 (2.45-4.15)	**4.58 (3.91-5.24)***
Vitamin D (μg)	4.00 (3.54-4.45)	**5.38 (4.81-5.96)****	3.92 (3.30-4.53)	**4.94 (4.46-5.42)****
Calcium (mg)	609.81 (565.3-654.4)	**780 (724-836)*****	626 (558-690)	**706 (658-760)***
Iron (mg)	12.27 (11.40-13.14)	**15.22 (14.1-16.2)*****	11.9 (10.69-13.20)	**14.4 (13.39-15.35)***
Folate (dietary folate equivalents)	425.77 (391.9-459.6)	434.81 (393.3-476.2)	383.11 (342.0-424.2)	**455.81 (423.9-487.6)*****
Vitamin B-6 (mg)	1.45 (1.4-1.5)	**2.07 (1.9-2.2)*****	1.55 (1.41-1.71)	**1.77 (1.65-1.89)****
Riboflavin (mg)	1.56 (1.5-1.7)	**1.88 (1.8-2.0)*****	1.51 (1.36-1.64)	**1.79 (1.69-1.91)****
Niacin (mg)	17.95 (17.1-18.8)	**20.66 (19.6-21.7)****	17.6 (16.32-19.11)	**19.7 (18.76-20.94)***
Weight status				
Waist circumference (cm)	**38.3 (37.4-39.1)**	36.8 (35.7-37.8)*	37.8 (36.8-38.9)	37.8 (36.9-38.6)
Body mass index	28.9 (28.1-29.7)	27.6 (26.6-28.7)	28.8 (27.5-29.9)	28.6 (27.6-29.5)

[a]Bold denotes higher value within the cluster. A cluster represents a group of individuals who consume a similar dietary pattern.
[b]Energy-adjusted.
*$P<0.05$.
**$P<0.01$.
***$P<0.001$.

total fat and saturated fat, as well as lower intakes of most vitamins and minerals (Table 2).

There were similar food groups among the higher and lower HEI clusters between the two methods. With both methods of analysis, the cluster with a lower HEI score had significantly more dairy desserts, processed meat, and fats/oil. Fruit and fish intakes were consistently higher in the higher HEI clusters across both methods. However, several food groups were different between clusters only when analyzed by number of servings. Higher servings of cereal, dark-green/deep-yellow vegetables, other vegetables, citrus/melons/berries, fruit juice, and milk but fewer servings of bread and eggs are shown in Cluster 2. Similarly, several food groups (ie, pasta/noodles/rice, starchy vegetables, vegetable soups/sauces/juices, cheese, red meats, and sweets) were different between clusters only when analyzed for energy contribution. Finally, some food groups differed by cluster with both analysis methods but had conflicting cluster placement. For example, poultry intake was higher in Cluster 2 when number of servings was the unit of analysis but with the percent energy method, poultry was lower in Cluster 2.

There were no significant differences between the clusters using the percent energy contribution from food subgroups with regard to sex, age, educational attainment, or use of tobacco products. There were significantly more women in the higher HEI score cluster using the servings method (data not shown), with no other differences noted with demographic variables. Notable differences in weight status, specifically, BMI and waist circumference data, are presented in Table 2. Only with the analysis based on number of servings method did those in Cluster 2 have a lower waist circumference with differences in BMI approaching significance ($P=0.08$).

DISCUSSION

We used different cluster analysis schema for analyzing the same dietary data from a sample of older adults, specifically, food group servings method and food group contribution to energy intake method. Both approaches yielded two dietary patterns; in both cases, one pattern had a higher HEI and more favorable nutrient profiles. Detection of only two patterns is intuitive to this sample, given the sample size and homogeneity; all participants were older rural adults residing in Pennsylvania. The lack of diversity and the size of this sample is a limitation of this study. However, other investigators with larger, more diverse samples have also found two dietary patterns (4,38-40).

Other studies have found two distinct dietary patterns within samples, generally one more healthful that has been termed "prudent" and one less healthful called "Western" (4,29,40-42). The more healthful or prudent dietary patterns have been related to lower risk for cardiovascular diseases (29,39) and markers of cardiovascular disease (39) and selected cancers (42), as well as lower risk for all-cause mortality (43). The Western or higher-fat cluster has been related to higher risk for type 2 diabetes mellitus (4). However, Togo and colleagues (44) found no consistent relationship between dietary patterns and obesity status in a recent review; this may be due to the various statistical methodologies employed across studies to derive dietary patterns. Newby and colleagues (5) found diets characterized by higher intakes of fruits, vegetables, and low-fat dairy and lower intakes of processed meats, sweetened carbonated beverages, and fast-food items to be associated with smaller increases in waist circumference and BMI among participants in the Baltimore Longitudinal Aging Study. The more healthful cluster described by Newby and colleagues (5) is very similar to the higher HEI score cluster that we found with the servings methods. Furthermore, we showed an association between dietary patterns and weight status only when analyzing data by food group servings.

When the servings method was employed, all of the fruit subgroups (ie, citrus, melons and berries, fruit juice, and other fruits) were associated with the more healthful eating pattern. However, only the "other" fruit subgroup (eg, bananas and apples) was associated with the more healthful pattern with the percent energy method. Furthermore, using the servings method, dark-green and deep-yellow vegetables and other vegetables were related to the more healthful pattern whereas no vegetable subgroups were associated with the more healthful pattern with the percent energy method. In fact, starchy vegetables and vegetable soups, sauces, and juices were associated with the less healthful cluster with the percent energy method. Fruits and vegetables generally do not contribute substantially to energy intake; this combined with our sample size may have limited our ability to discern differences with the percent energy method. Total fruit and vegetable intake has been identified with healthful diets in many studies and intake of these foods has been associated with more desirable weight status (2,5,45). We also noted cluster differences for all fruits and all vegetables with the servings method but for all fruit only with the percent energy methods (data not shown). Many other studies have indicated higher fruit and vegetable intakes to be associated with better health outcomes and more favorable weight status, only the servings methods of analysis was consistent with these findings.

Increased interest in dietary patterns has led to greater scrutiny of statistical methodology (30,46,47). Methods of analysis have been based on various presentations of food groups, including mean energy contribution of food groups (7), number of food group servings (13), gram weight of foods (9,14-16), and frequency of food group servings (18). One of the most common strategies for analyzing dietary patterns is via daily contribution of energy from food groups (3,10-12). When this method was applied to our data, the clusters were not as interpretable or comparable to a healthful dietary pattern as outlined by dietary guidelines. The goal of cluster analysis is to classify data according to a set of rules derived from the aims of the analysis. The groups formed from the classification are expected to be useful for describing behavior or phenomena. The groups reflect the rules chosen for the classification. To the extent that the rules are useful and valid, the groupings are useful and defensible.

Analytic epidemiology is needed within the same cohort to determine dietary patterns and elucidate which procedures yield the best description of the sample or are most predictive of health outcomes (30). The number of servings analysis strategy yielded results consistent with a more healthful diet and better weight status for our cohort. This is a novel approach and one that needs to be further explored within other groups because results from cluster analysis are sample specific. Public health messages are given in terms of servings of food or amounts of food. Thus, clusters derived from number of servings can be used to explore relationships of these recommendations with health outcomes.

Comparative strategies are recommended when determining dietary patterns of groups. Using number of servings as the clustering variable is a strategy that best described more the healthful eating patterns among our sample and it is consistent with dietary strategies outlined by national guidelines. Determining dietary patterns within specific populations can help guide successful interventions that can be translated into practice.

The research contained within this article was supported in part by National Institutes of Health R21AG023179-01A1, US Department of Agriculture no. 58-1950-6019, and National Institute of Aging training grant, Interdisciplinary Training in Gerontology, no. T32 AG00048.

References

1. Chen H, Ward MH, Graubard BI, Heineman EF, Markin RM, Potischman NA, Russell RM, Weisenburger DD, Tucker KL. Dietary patterns and adenocarcinoma of the esophagus and distal stomach. *Am J Clin Nutr*. 2002;75:137-144.
2. Quatromoni PA, Copenhafer DL, D'Agostino RB, Millen BE. Dietary patterns predict the development of overweight in women: The Framingham nutrition studies. *J Am Diet Assoc*. 2002;102:1239-1246.
3. Tucker KL, Chen H, Hannan MT, Cupples LA, Wilson PW, Felson D, Kiel DP. Bone mineral density and dietary patterns in older adults: The Framingham Osteoporosis Study. *Am J Clin Nutr*. 2002;76:245-252.
4. van Dam RM, Rimm EB, Willett WC, Stampfer MJ, Hu FB. Dietary patterns and risk for type 2 diabetes mellitus in US men. *Ann Intern Med*. 2002;136:201-209.
5. 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.
6. Newby PK, Tucker KL. Empirically derived eating patterns using factor or cluster analysis: A review. *Nutr Rev*. 2004;62:177-203.
7. 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.
8. Willett WC, Buzzard MI. Foods and nutrients. In: Willett WC, ed. *Nutritional Epidemiology*. 2nd ed. New York, NY: Oxford University Press; 1998
9. 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.
10. 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.
11. 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.
12. Rasanen M, Lehtinen JC, Niinikoski H, Keskinen S, Ruottinen S, Salminen M, Ronnemaa T, Viikari J, Simell O. 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.
13. Ledikwe JH, Smiciklas-Wright H, Mitchell DC, Miller CK, Jensen GL. Dietary patterns of rural older adults are associated with weight and nutritional status. *J Am Geriatr Soc*. 2004;52:589-595.
14. Pryer JA, Nichols R, Elliott P, Thakrar B, Brunner E, Marmot M. Dietary patterns among a national random sample of British adults. *J Epidemiol Community Health*. 2001;55:29-37.
15. 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.
16. 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.
17. 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.
18. Fraser GE, Welch A, Luben R, Bingham SA, Day NE. The effect of age, sex, and education on food consumption of a middle-aged English cohort-EPIC in East Anglia. *Prev Med*. 2000;30:26-34.
19. 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.
20. Hulshof KF, Wedel M, Lowik MR, Kok FJ, Kistemaker C, Hermus RJ, ten Hoor F, Ockhuizen T. Clustering of dietary variables and other lifestyle factors (Dutch Nutritional Surveillance System). *J Epidemiol Community Health*. 1992;46:417-424.
21. Farchi G, Mariotti S, Menotti A, Seccareccia F, Torsello S, Fidanza F. Diet and 20-year mortality in two rural population groups of middle-aged men in Italy. *Am J Clin Nutr*. 1989;50:1095-1103.
22. Schroll K, Carbajal A, Decarli B, Martins I, Grunenberger F, Blauw YH, de Groot CP. Food patterns of elderly Europeans. SENECA Investigators. *Eur J Clin Nutr*. 1996;50(suppl 2):S86-S100.
23. US Department of Health and Human Services, US Department of Agriculture. *Nutrition and Your Health: Dietary Guidelines for Americans*. 6th ed. Washington, DC: US Government Printing Office; 2005. Home and Garden Bulletin No. 232.
24. US Department of Agriculture. *The Food Guide Pyramid: A Guide to Daily Food Choices*. Washington, DC: US Department of Agriculture, Human Nutri-

tion Services; 1992. Home and Garden Bulletin No. 252.
25. Block G, Patterson B, Subar A. Fruit, vegetables, and cancer prevention: A review of the epidemiological evidence. *Nutr Cancer*. 1992;18:1-29.
26. Horn-Ross PL, Morrow M, Ljung BM. Diet and the risk of salivary gland cancer. *Am J Epidemiol*. 1997; 146:171-176.
27. Appel LJ, Miller ER 3rd, Jee SH, Stolzenberg-Solomon R, Lin PH, Erlinger T, Nadeau MR, Selhub J. Effect of dietary patterns on serum homocysteine: Results of a randomized, controlled feeding study. *Circulation*. 2000;102:852-857.
28. Appel LJ, Moore TJ, Obarzanek E, Vollmer WM, Svetkey LP, Sacks FM, Bray GA, Vogt TM, Cutler JA, Windhauser MM, Lin PH, Karanja N. A clinical trial of the effects of dietary patterns on blood pressure. DASH Collaborative Research Group [comments]. *N Engl J Med*. 1997;336:1117-1124.
29. 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.
30. Kant AK. Dietary patterns and health outcomes. *J Am Diet Assoc*. 2004;104:615-635.
31. Ledikwe JH, Smiciklas-Wright H, Mitchell DC, Jensen GL, Friedmann JM, Still CD. Nutritional risk assessment and obesity in rural older adults: A sex difference. *Am J Clin Nutr*. 2003;77:551-558.
32. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO. Development and validation of a geriatric depression screening scale: A preliminary report. *J Psychiatr Res*. 1982;17:37-49.
33. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state.” A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12:189-198.
34. Jonnalagadda SS, Mitchell DC, Smiciklas-Wright H, Meaker KB, Van Heel N, Karmally W, Ershow AG, Kris-Etherton PM. Accuracy of energy intake data estimated by a multiple-pass, 24-hour dietary recall technique. *J Am Diet Assoc*. 2000;100:303-308.
35. Mitchell DC. Database requirements to meet food grouping needs. *J Food Comp Analysis*. 2001;14:279-285.
36. Kennedy ET, Ohls J, Carlson S, Fleming K. The Healthy Eating Index: Design and applications. *J Am Diet Assoc*. 1995;95:1103-1108.
37. Jensen GL, Friedmann JM, Coleman CD, Smiciklas-Wright H. Screening for hospitalization and nutritional risks among community-dwelling older persons. *Am J Clin Nutr*. 2001;74:201-205.
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. Fung TT, Rimm EB, Spiegelman D, Rifai N, Tofler GH, Willett WC, Hu FB. Association between dietary patterns and plasma biomarkers of obesity and cardiovascular disease risk. *Am J Clin Nutr*. 2001;73:61-67.
40. Tsai YY, McGlynn KA, Hu Y, Cassidy AB, Arnold J, Engstrom PF, Buetow KH. Genetic susceptibility and dietary patterns in lung cancer. *Lung Cancer*. 2003; 41:269-281.
41. Osler M, Helms Andreasen A, Heitmann B, Hoidrup S, Gerdes U, Morch Jorgensen L, Schroll M. 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.
42. 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.
43. 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.
44. 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.
45. Lin H, Bermudez OI, Tucker KL. Dietary patterns of Hispanic elders are associated with acculturation and obesity. *J Nutr*. 2003;133:3651-3657.
46. Newby PK, Muller D, Tucker KL. Validation of empirically derived eating patterns with plasma lipid biomarkers: Comparision of cluster and factor analysis methods. *Faseb J*. 2004;18:108.
47. Hoffmann K, Schulze MB, Schienkiewitz A, Nothlings U, Boeing H. Application of a new statistical method to derive dietary patterns in nutritional epidemiology. *Am J Epidemiol*. 2004;159:935-944.