British Journal of Nutrition (2005), **93**, 573–574
DOI: 10.1079/BJN20051453

Invited commentary

New approaches to the study of dietary patterns

The relationship between diet and health can be examined at the level of food components, foods and dietary patterns. Until recently, the study of food components, particularly nutrients, has been the dominant approach in nutritional epidemiology. This approach has clear advantages. If the development of a disease is causally related to the intake of a food component, the examination of that food component will be the approach with the greatest power to identify its effect. In addition, results for food components can be compared with associations observed in other populations, data from mechanistic studies and health effects found in intervention studies (Willett & Buzzard, 1998). Knowledge on the level of food components can then be used to produce foods with higher or lower levels of the component. For example, reduction of the amount of *trans*-fatty acids in margarines has probably resulted in substantial health benefits for populations (Oomen *et al.* 2001). However, the effect of a food component can differ depending on the food that it is derived from due to interactions between food components or physical characteristics of foods. For example, folate from beer may provide less health benefits than the same amount of folate from bread, because alcohol reduces intestinal folate uptake, interferes with folate metabolism and increases urinary loss of folate (Jiang *et al.* 2003). Effects of food consumption on disease risk can be different from predictions based on known effects of food components, because our knowledge on the myriad possible beneficial or detrimental aspects of foods (Jacobs & Murtaugh, 2000) is still limited. The study of foods and food groups accounts for interactions between different components of a food, and for effects of physical characteristics and unknown components.

In addition to different components of a food, synergy or antagonism may also exist for components of different foods and drinks that are included in the dietary pattern of an individual. As a result, health effects of dietary patterns may be greater than for individual foods or nutrients. Results from the Dietary Approaches to Stop Hypertension (DASH) intervention study illustrate the importance of considering dietary patterns (Appel *et al.* 1997). The DASH diet, a diet characterized by high consumption of vegetables, fruits, low-fat dairy products and whole grains, resulted in a greater reduction in blood pressure than had been found for individual minerals in these foods. Another advantage of the study of dietary patterns in epidemiological studies is that potential dietary confounders are largely incorporated in the dietary pattern variable.

Different methods have been used to study dietary patterns in epidemiological studies. Two main approaches can be distinguished. First, an exploratory approach can be used that identifies combinations of foods and drinks as they are consumed in reality in a particular population. Principal components analysis is a frequently used exploratory approach to identify dietary patterns (Hu *et al.* 2000). This data reduction technique constructs new variables that are linear combinations of the original variables and explain as much of the variation in the original variables as possible. Applied to dietary data, new dietary pattern variables are derived on the basis of the correlation matrix of the original food variables. The use of principal components analysis to derive dietary pattern variables requires several arbitrary choices such as the original variables to include and the number of dietary patterns to identify (Martinez *et al.* 1998). This underlines the importance of conducting sensitivity analyses to examine the robustness of the findings (Hu *et al.* 2000). Dietary patterns identified by the explanatory approach reflect dietary behaviour and are not based on known health effects of diet. As a result, the identified dietary patterns are not necessarily relevant for disease risk. However, results can increase insight into possibilities for dietary changes and can provide information for setting priorities for changing dietary patterns in a population by public health initiatives. Second, a hypothesis-oriented approach that uses predefined criteria to construct dietary pattern scores can be used. These scores reflect the degree to which a person's diet conforms to a dietary pattern that was defined *a priori* based on presumed health effects. Scores based on dietary recommendations (McCullough *et al.* 2002) and characteristics of the traditional Mediterranean diet (Trichopoulou *et al.* 2003) have been used. Deciding what individual components to include, the cut-off points and the weights of different components of an *a priori* score still requires subjective choices. This approach does not have the advantages related to studying existing dietary behaviour or to the identification of new dietary patterns that may affect disease risk. However, the approach can capture the greater effects of the overall diet as compared with individual components, and can be used to test the validity of dietary recommendations.

Hoffmann and colleagues have introduced a new method in nutritional epidemiology that combines characteristics of the explanatory and hypothesis-oriented approaches to dietary patterns (Hoffmann *et al.* 2004*a*). This reduced rank regression (RRR) method has similarities with principal components analysis, but it uses a set of response variables in addition to a set of food intake variables. The identified dietary patterns are linear combinations of the original food intake variables that maximally explain variation in the response variables (Hoffmann *et al.* 2004*a*). *A priori* knowledge is introduced by using a set of response variables that is known to predict the disease of interest. In a paper in this issue of BJN, Hoffmann *et al.* (2005) have used macronutrient intakes as response variables, resulting in dietary pattern variables that are linear combinations of the original food variables and maximally explain variation in macronutrient intakes. The identified dietary pattern was characterized by higher intakes of meats, poultry, butter, sauces and eggs, and lower intakes of bread and fruit, and was associated with a higher risk of premature mortality. Although the approach used invites interpretation of this association in terms of macronutrients only, other characteristics of the identified dietary pattern may still have contributed to the observed association. For the RRR method, the number of dietary patterns that can be identified is limited by the number of response variables. However, arbitrary decisions such as the choice of original food intake variables remain necessary, and sensitivity analyses should be conducted. An alternative approach that is of potential interest is the use of biomarkers of dietary intake that are relevant for the disease of interest as response variables. This application would identify combinations of foods that maximally explain variation in biomarkers of dietary intake, avoiding measurement error related to food composition data or to lack of information on bioavailability for different combinations of foods.

In a case–control study of coronary heart disease, Hoffmann and colleagues used biomarkers of disease as response variables in the RRR analysis (Hoffmann *et al.* 2004*b*). A dietary pattern characterized by higher intakes of meat, margarine, poultry and sauce, and lower intakes of vegetarian dishes, wine, vegetables and whole grains, maximally explained variation in the five selected biomarkers of CVD (HDL-cholesterol, LDL-cholesterol, lipoprotein(a), C-peptide and C-reactive protein concentrations). The identified dietary pattern was strongly associated with CHD. The use of biomarkers of disease as response variables is of interest because it may aid the pathophysiological interpretation of observed associations between dietary patterns and disease. However, the explorative nature of the approach should be kept in mind: in the theoretical situation where the biomarkers used perfectly predict the disease, the method simply identifies the combination of food variables that is most highly correlated with the disease in the study population. Thus, an important consideration is whether findings can be confirmed in diverse populations. After randomly splitting the sample in half, dietary pattern variables identified in the two sub-populations remained associated with CHD in the other sub-population, but the associations were much weaker (Hoffmann *et al.* 2004*b*).

The RRR method may prove to be a useful tool in nutritional epidemiology, because it can provide insight into the pathophysiological pathway that links dietary patterns to disease. Use of the same set of response variables in different study populations may facilitate the comparison of results from diverse populations. In addition, confirmation of results for RRR or principal components analyses in other populations can be performed by calculating a dietary pattern score using the same weights and food variables or by using a simplified dietary pattern score (Schulze *et al.* 2003*a*). An application of dietary pattern analyses is to obtain insight into intercorrelations between dietary variables and the potential for confounding. In addition, adjustment for dietary pattern scores may be an efficient method to address confounding by other dietary variables in studies of individual foods and food components (Schulze *et al.* 2003*b*). The combination of analyses of food components, foods and dietary patterns is likely to provide most insight into the relationship between diet and disease risk. The methodology to study dietary patterns is still developing and future studies will further clarify what methods will be most useful for addressing confounding, formulating new hypotheses on the link between diet and disease, and providing information for public health interventions.

Rob M. van Dam

Department of Nutrition and Health
Faculty of Earth and Life Sciences
Vrije Universiteit Amsterdam
De Boelelaan 1085
1081 HV Amsterdam, The Netherlands

References

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

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

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

Hoffmann K, Boeing H, Boffetta P, Nagel G, Orfanos P, Ferrari P & Bamia C (2005) Comparison of two statistical approaches to predict all-cause mortality by dietary patterns in German elederly. *Br J Nutr* **93**, 709–716.

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

Jacobs DR Jr & Murtaugh MA (2000) It's more than an apple a day: an appropriately processed, plant-centered dietary pattern may be good for your health. *Am J Clin Nutr* **72**, 899–900.

Jiang R, Hu FB, Giovannucci EL, Rimm EB, Stampfer MJ, Spiegelman D, Rosner BA & Willett WC (2003) Joint association of alcohol and folate intake with risk of major chronic disease in women. *Am J Epidemiol* **158**, 760–771.

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

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

Oomen CM, Ocke MC, Feskens EJ, van Erp-Baart MA, Kok FJ & Kromhout D (2001) Association between trans fatty acid intake and 10-year risk of coronary heart disease in the Zutphen Elderly Study: a prospective population-based study. *Lancet* **357**, 746–751.

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

Schulze MB, Manson JE, Willett WC & Hu FB (2003*b*) Processed meat intake and incidence of Type 2 diabetes in younger and middle-aged women. *Diabetologia* **46**, 1465–1473.

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

Willett WC & Buzzard IM (1998) Foods and nutrients. In *Nutritional Epidemiology*, 2nd ed. pp. 18–32 [WC Willett, editor]. New York: Oxford University Press.