A new approach to the study of diet and risk of type 2 diabetes

Jinlin Fu, Binyou W*, Terry C

ABSTRACT

Type 2 diabetes mellitus (T2DM) is a polygenetic disorder resulting from interaction of both hereditary and environmental factors. Diet is an important environmental factor in the development of T2DM. However, there existed inconsistent conclusions among previous studies. The validity of a study of associations between diet and diabetes depends on the method used in the study. Associations between individual nutrients/foods and T2DM have been explored even till today, but the controversy is still unsolved due to the limitations of the methods used in the studies of associations of dietary patterns with T2DM. The advantages and disadvantages of the methods are also explored. In this review, dietary-pattern-related studies since 1990 were identified by searching through Medline and PubMed in order to analyze methods used in the studies of diabetes. A dietary pattern approach was introduced as an alternative complementary way used to test associations of diet with risk of T2DM and the advantages of the studies of T2DM using this new approach are also explored. It is concluded that the dietary pattern approach shows its advantages over that using individual nutrients or foods in studies on associations between the diet and diabetes, but its reproducibility and validity for such effects, however, need to be further verified in different ethnic population-based on diverse eating habits. The long-term effects of a beneficial dietary pattern on T2DM also require clarifying in future studies.

KEY WORDS: Approach, diet, dietary pattern, noninsulin-dependent diabetes mellitus, type 2 diabetes

Diet has been known for many years to play a key role as a risk factor in the development of chronic diseases.[1] Diet is thought to have important effects on the development of diabetes.[2-5] For instance, it was reported that diabetes was related to certain carbohydrate,[6] fatty acid,[6] cereal fibers,[5] meat[7,8] and drinking coffee[9] and sweet beverages.[10] Even so, there are inconsistent conclusions on studies on the relation of single diet to risk of type 2 diabetes mellitus (T2DM).[6] These analyses are valuable but to some extent they are confined by the theory and methods. Studying associations between diet and T2DM using single foods or nutrients is limited. For example, people take in nutrients not in isolation, but they ingest a variety of nutrients at the same time when consuming sorted foods. These nutrients may counteract or synergize each other.[11]

Recently, the emerging dietary pattern analyses have become an alternative and complementary way used to examine associations of diet with risk of chronic disease.[12] Dietary patterns (DP), also called eating patterns or food patterns, are “foods as they are actually consumed in various characteristic combinations”.[13] Since the early 1980s, interest in dietary patterns has been growing with many studies focusing on the associations between dietary patterns and disease outcomes or biomarkers.[14] Dietary patterns represent not only combinations of different foods consumed, but also accumulative effects of overall diet on diseases and moreover they have associations with the risk of diabetes.[7,15,16]

In this review, a dietary pattern approach is introduced as an alternative and complementary way used to explore associations of diet with risk of T2DM. The advantages and disadvantages of the methods are also summarized based on studies on T2DM using both traditional and new approaches. Medline and PubMed were searched to identify diet-related studies published since 1990 by using key words “Type 2 diabetes”, “noninsulin-dependent diabetes mellitus” and “dietary/food/eating pattern”. Only full-text papers which were closely related to the review objectives were included, but those using dietary index to identify dietary patterns were excluded because of their inconsistency in defining a dietary pattern and their flexibility with study purposes.

Methods Used to Define Dietary Patterns

Dietary patterns can be identified in terms of dietary indexes, but here we only discuss those defined using statistical exploratory ways such as factor analysis, cluster analysis and reduced rank regression.
Factor analysis approach

Factor analysis (FA) is a multivariate statistical technique and its dietary information used comes from food frequency questionnaires (FFQ) or dietary records. The analysis process generates pattern scores that summarize and then dissect the correlated structures of original food items. Then, the common and basic dimensions (i.e., factors or patterns) of food consumption are determined with this information. Of most researches, however, original food items measured were divided and reunited in small numbers into input variables, usually called food groups so as to be input into FA, that is, much information of consumed foods were compressed into several important factors by analyzing covariance structures of incepted foods.\[17\]

After statistical analyses, 2-13 pattern variables (also called factor scores) are usually obtained, but 2-4 in the commonest.\[18,19\] By calculating general scores of each DP and applying them to correlation and regression analyses, we are able to examine the relationships of each eating pattern to outcomes of interest, e.g., nutrient intakes, risks of diseases and other biochemical markers related to health.

Factor analysis is a convenient way to be used when DPS are to be identified based on the existing dietary data, but it is not a perfect way to do this. The performance of FA is quite easy. A researcher can simply run the FA procedure in any statistical software package such as SAS, STIXA or SPSS to the input database.

Cluster analysis approach

Cluster analysis (CA) is another multivariate statistical approach to defining specialized dietary patterns. Compared with FA, CA classifies study individuals with relatively similarity, not food items, into subgroups. This is done by distributing individuals into the subgroups which are determined beforehand according to Euclidean Distances, with which the distances within groups are shorter than those between groups.\[17\] That is to say, CA is to attenuate the number of data to enter into the DPS, based on the differences in average food intakes of individuals and in result, the diet of individuals within the same group is relatively identical, whereas that of those between groups is diverse.

Individuals can also be distributed into different subgroups by frequency of food consumption,\[18,19\] percentage of energy supplied by each food or food group,\[20\] mean weight (in gram) of food intakes,\[21\] standardized nutrient intakes\[22\] or integration of both diet and biochemical results.\[23\]

K-means cluster analysis, which is installed in most statistic software packages such as STIXA and SPSS, is the most widely used to define DPS,\[22,24,25\] especially for large sample size. Usually 2-8, but more often 5-6 clusters, are induced after completion of the clustering procedure, as reported in the published literature.\[14\] In addition, further analyses are required, for example, to interpret the defined DPS by comparing dietary features of each pattern.

Cluster analysis is an effective approach to measure exposure to diet and it can be used to explore the relationships among eating patterns, health and diseases.\[26\] Cluster analysis, like FA, can be employed directly with a readymade program, which has been installed in statistical software packages.

Both FA and CA are partially subjective, which makes their application limited. For instance, a researcher needs to determine in advance how many DPS to be retained, how to group foods or input food variables and how to name patterns generated.\[14,27\] The subjectivity of FA also presents with choosing rotation. In addition, either FA or CA induces DPS, independent of a disease, which may vary with the different populations studied. So, poor repeatability may appear in different studies.

Reduced rank regression approach

Reduced rank regression (RRR),\[28,29\] also called maximum redundancy analysis (MRA), is an approach to inducing DP scores using PLS procedure in SAS software package. Reduced rank regression is neither an a priori approach nor an exploratory statistical one, but an a posteriori approach. It works with two different sets of variables, i.e. predictors and responses. The first set of variables may be food groups or the other nutrients and the second, biomarkers related to diseases of interest. Its strong point is that it can induce pattern scores by maximizing variation of explained concentrations of biomarkers and other mediated variables related to the studied diseases. So, RRR may be a useful tool in studying associations of diet with diseases.

For FA, it is necessary to solve for maximum variation in all predictor variables (e.g., food groups) and to identify scores, namely linear combinations (or functions). In contrast, RRR identifies linear functions of predictors, which explain as much response variation as possible.\[30\] Formally, principal component analysis is a special case of RRR in which predictors will be taken as responses; RRR equals multivariate linear regression only in the special occasion when there are merely responses. RRR can be used to interpret variation of nutrients and nutrient-related responses through linear functions of food intakes. Compared with traditional FA, RRR is more flexible and efficient and can be employed to choose disease-specific responses in nutritional epidemiology and to identify the combination forms of food intakes.\[28\]

Reduced rank regression, like FA and CA, has its limitation when used to generate DPS.\[28\] Firstly, factor scores in RRR cannot perfectly measure dietary features, but they are just linear combinations of food intake information collected through FFQs which have a large bias in the measurement of food intakes.\[31\] Next, RRR, similar to FA and CA, uses the existing data to assess a factor score coefficient, which cannot be acquired from the diet data based on another population. This means that distinct diet information will result in somewhat different DPS. Schulze et al.\[32\] recommended that omitting food groups with low score coefficients and ignoring weights of retained food groups should reduce dependence of data on DP variables. Last, although the objective of RRR is to interpret as much response variation as possible, predicted values of all responses may be greater than those of all the RRR factors. For example, Hoffmann and co-workers\[28\]
indicated that the four nutrients (unsaturated/saturated fat, dietary fiber, magnesium and alcohol), in which they were interested, can better predict diabetes mellitus than the four factor scores generated with RRR can, but a diabetes model based on nutrients as independent variables cannot account for which foods are able to attenuate the risks of diabetes.

### Some Studies of the Associations Between Dietary Patterns and T2DM

Most observations dealing with relations between dietary patterns and T2DM depict the associations by conducting cross-sectional surveys with *a posteriori* approaches [Table 1]. Factor analysis is mostly used to identify correlations of dietary patterns with diabetes. Slattery *et al.* first identified two patterns with FA, namely “prudent pattern” and “western pattern”. After that, similar patterns were achieved in studies by others. van Dam and co-workers [15] pointed out that prudent pattern score was related to reduction of risk of T2DM (RR for extreme quintiles, 0.84; 95%CI, 0.70-1.00), whereas the western dietary pattern score was associated with an increased risk of T2DM (RR=1.59, *P*<0.001). The study also showed that people with high prudent pattern scores were older, physically more active, unlikely to smoke and more likely to be of hypercholesterolemia; in contrast, those with high western pattern scores were just opposite, but the associations did not change much after adjustment of physical activity, smoking, drinking, hypercholesterolemia and family history of diabetes. They also found that of all the foods characterized in the prudent pattern, whole grain had a strong inverse relation to, and all the main foods characterized in the western pattern were positively related to, risk of Type 2 diabetes.[19]

Heidemann *et al.* defined, using RRR, dietary patterns related to biomarkers of diabetes and studied their relations with the risk of T2DM. The dietary patterns were deduced from the data of food intakes of 48 food groups as exposure variables and from the data of biomarkers, as responses, including glycosylated hemoglobin (HbA1c), high density lipoprotein cholesterol (HDLC), C-reactive protein (CRP) and adiponectin. Then, they estimated the relation of scores to the risk of diabetes with logistic regression.

Heidemann and co-workers chose the four biomarkers based on the previous studies: 1) HbA1c reflects raised levels of blood glucose, which is usually taken as a glycemic index for long-term control. The impairment of internal glucose homeostasis will cause increased risk of T2DM. 2) Studies[40,41] showed that patients with insulin resistance and T2DM are often low in HDLC level, which can therefore be an indicator of onset of T2DM. 3) It was reported[42] that diabetes is related to sub-clinically chronic inflammation which is characterized by the changes in protein contents at its acute stage. Among the changes is CRP which is currently the best indicator of potential risk of T2DM. 4) There is evidence[43] showing that adiponectin from fat cells is sensitive to insulin and anti-inflammation. Although the association between adiponectin and diabetes is not clear yet, plasma adiponectin level seems to be an independent indicator of risk of T2DM. Heidemann *et al.* pointed out that the above four diabetes-related biomarkers could be affected by dietary components, particularly by alcohol, fat consumption and certain types of carbohydrates.

The dietary patterns defined by Heidemann and co-workers were characterized by high intakes of fresh fruits and low intakes

### Table 1: Dietary patterns and type 2 diabetes mellitus

<table>
<thead>
<tr>
<th>First author (year)</th>
<th>Study</th>
<th>Pattern No.</th>
<th>Outcomes (s)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>van Dam (2002)[15]</td>
<td>Prospective study</td>
<td>2 factors</td>
<td>T2DM</td>
<td>The western pattern score was associated with an increased risk, but the prudent with a modestly lower risk, of T2DM. RR for extreme quintiles: Western: 1.59 (95%CI:1.32-1.93, <em>P</em>&lt;0.001); Prudent: 0.84 (95%CI: 0.70 to 1.00). Western factor associated with T2DM: The highest quintile/lowest: RR=1.49(95%CI:1.26-1.76); <em>P</em>&lt;0.001. For red meat: RR=1.26(95%CI:1.21-1.46). Total processed meat: RR=1.38 (95%CI:1.23-1.56). Bacon: RR=1.73 (95%CI:1.23-2.16); Hot dog: RR=1.43 (95%CI:1.22-1.69). Factor 1 RR=0.76 (95%CI:0.29-1.96) <em>P</em>=0.62; Factor 2 RR= 1.04 (95%CI:0.51-2.12) <em>P</em>=0.77; Factor 3 RR= 1.32 (95%CI:0.70-2.48) <em>P</em>=0.17; Factor 4 RR= 0.49 (95%CI:0.25- 0.94) <em>P</em>=0.01.</td>
</tr>
<tr>
<td>Fung (2004)[17]</td>
<td>Prospective study</td>
<td>2 factors</td>
<td>T2DM</td>
<td></td>
</tr>
<tr>
<td>Hoffmann (2004)[24]</td>
<td>Case-control study</td>
<td>4 factors</td>
<td>T2DM</td>
<td></td>
</tr>
<tr>
<td>Heidemann (2005)[38]</td>
<td>Nested case-control study</td>
<td>1 factor</td>
<td>T2DM</td>
<td>OR (after multivariate adjustment) for DM across increasing quintiles of the dietary pattern score were 1.0, 0.59, 0.51, 0.26 and 0.27, respectively (<em>P</em> = 0.0006 for trend) RR (adjusted for non-dietary confounders) between the extreme quartiles of the pattern scores: Prudent pattern = 0.72; (95%CI: 0.53-0.97); <em>P</em> for trend = 0.03; Conservative pattern =1.49 (95%CI: 1.11-2.00); <em>P</em> for trend = 0.01) Study 1: OR (comparing extreme quintiles) = 0.56 (95% CI: 2.10, 3.12); <em>P</em> for trend &lt; 0.001; Study 2: OR=2.93 (95%CI: 2.18, 3.92); <em>P</em> for trend &lt; 0.001</td>
</tr>
<tr>
<td>Montonen (2005)[16]</td>
<td>Cohort study</td>
<td>2 factors</td>
<td>T2DM</td>
<td></td>
</tr>
<tr>
<td>Schulze (2005)[39]</td>
<td>Nested case-control study</td>
<td>1 factor</td>
<td>T2DM</td>
<td></td>
</tr>
</tbody>
</table>

T2DM - Type 2 diabetes mellitus

J Postgrad Med April 2007 Vol 53 Issue 2
of high-energy drinks, beer, red meat, poultry, meat products, legume and bread (excluding wholegrain bread). Subjects with high dietary pattern scores were high in levels of HDLc and adiponectin, but low in those of HBA1c and CRP. After adjustment of multivariate, odds ratio (OR) of T2DM decreased with increases of dietary pattern scores (OR: 1.000.590.510.26 and 0.27).[38]

Acknowledgement

I gratefully acknowledge my advisors, Prof. Wang and Dr. Terry Coyne for their kind direction to my study. I also appreciate the help from Dr. P.K. Newby.

References


Source of Support: The University of Queensland. Conflict of Interest: None declared.