Dietary patterns by reduced rank regression predicting changes in obesity indices in a cohort study: Tehran Lipid and Glucose Study

Roya Sherafat-Kazemzadeh MD1,2, Shahryar Egtesadi PhD1, Parvin Mirmiran PhD2, Mahmoud Gohari PhD3, Sara Jalali Farahani BS2, Firoozeh Hosseini Esfahani MS2, Mohammad Reza Vafa PhD1, Mehdi Hedayati PhD4, Fereidoun Azizi MD4

1Nutrition Department, Faculty of Health, Iran University of Medical Sciences, Tehran, Iran
2Nutrition Unit, Obesity Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
3Biostatistics Department, Faculty of Management, Iran University of Medical Sciences, Tehran, Iran
4Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Objective: To examine the association between dietary patterns and obesity indices (BMI, WC, WHR) among Tehranian adults in a 6-year follow-up study. Methods: Within frame of a cohort study in Tehran (mean follow up 6.6±0.9 years), 141 adults were recruited with: two 24 hour dietary recalls at the beginning, as well as obesity indices at the beginning and end of the study period. Dietary intakes were converted into grams of intakes of food items and categorized into 16 groups. Reduced rank regression analysis derived five patterns with total and polyunsaturated-to-saturated fat intake, cholesterol, fiber and calcium intake as response variables. Factors (dietary patterns) were generated retaining a corresponding factor loading ≥ |0.17| on the food groups. Changes in obesity indices were scrutinized within quintiles of factor scores. Results: There were high loadings on refined carbohydrates, whole grain, starchy vegetables, other vegetables, red and refined meat, saturated/trans fat, and egg for the first factor named “traditional”. All obesity indices had increasing trend across quintiles of pattern score. The fifth pattern (namely egg pattern) had high loading for eggs, salty snacks, as well as fruits and dry fruits, and negative loadings for red and processed meat, saturated and trans fat, plant oils, and dairy products. This pattern showed increasing trends for WC and WHR after adjustment for potential confounders. Other patterns showed non-significant trends for obesity indices. Conclusions: The results were indicative of a traditional pattern which is dominated in the Tehran region and associated with increase in obesity indices.

Key Words: dietary patterns, anthropometry, obesity, reduced rank regression, cohort

INTRODUCTION

As many as 250 million people worldwide are reported to be obese (7% of the adult population).1 The national figures in some of the developed countries express a two-to-three fold increase in the prevalence of obesity in two last decades,2,3 and the prevalence exceed 20%. This increase is particularly noticeable in countries where no overweight and obesity problem existed previously, but it is also evident in countries in which obesity has been present for decades.4,5 Obesity and health-related conditions are gaining importance in Iran as this country undergoes nutrition transition due to urbanization and continuous industrialization.6,7 Nevertheless there are different reports dealing with this health problem in this country. Still, rural areas demonstrate a lower rate of obesity (23% central obesity), compared to urban areas with rates as high as 28.5%.8 Nevertheless, diverse figures are reported in different areas of Iran from 49.9% among Tehranian women,9 to 23.1%.10-12 Obesity has been shown to have important metabolic adverse consequences. Nowadays, the increasing incidence of type 2 diabetes during last decades has been speculated to result from obesity, especially, central obesity.13,14,15 In a study by Zabetian et al., they reported a prevalence of over 30% for the metabolic syndrome among Tehranian adults.16 This means that obesity consequences are gaining ground in this city.

The increasing body of scientific evidence emphasized in recent decades on the importance of dietary factors in the development and prevention of non-communicable diseases such as cardiovascular diseases (CVD), atherosclerosis, type 2 diabetes mellitus (T2DM), and different types of cancer and mortality rates.17-20 Different type of dietary analysis has been employed to elucidate and inter-
t against metabolic derangements. A major shortcoming of this approach was the fact that nutrients showed interactions when consumed in combinations and the in-vitro and animal model in this context did not completely comply with the conditions in humans. Besides, adjustment of food intake data would result in a large number of regression parameters to estimate and would impair the power of the analysis and produce wide confidence intervals. A more promising approach was introduced by using dietary quantitative and qualitative scores based on the evidence provided by preceding findings, and the introduction of the dietary pyramid. Diet quality scoring gained attention but showed limited usefulness because it ignored consumption of some food groups. An alternative approach is rather an a-posteriori approach using factor analysis (FA) to reduce intake data in several (often 4 to 5) factors which were then entered into regression models for the study of associations with outcomes. Factor Analysis aims to construct linear combinations of food intakes, which explain a high proportion of the variation in food intakes. However, explaining much variation in food intakes by FA does not necessarily mean that also variations in macro- or micronutrients will also be properly explained. Hoffmann and colleagues introduced reduced rank regression (RRR)-based analysis first in 2004 to attain more focus on the variation in selected nutrients and the mix of energy source when examining the effects of diet on disease. As such, RRR is a more flexible statistical method than FA because it works with two sets of variables. It aims to construct linear combinations of variables belonging to one set by maximizing the explained variation in variables of the other set. They explained how to use a set of food groups as the “predictor variables” and a second set which may be related to the outcome of interest, namely “response variables”. This latter set is thought to be disease specific based on a-priori information from epidemiologic nutritional studies. In a comparison of the two data reduction techniques (RRR, and Principle Component Analysis which is very similar to FA) in a nested-case-control study for prediction of T2DM, Hoffmann and colleagues extracted four factors by either methods and reported that RRR factors explained 93.1% of response variation, whereas the first four factors obtained by principal component analysis accounted for only 41.9% and they concluded that in contrast to principal component analysis and other methods, the new RRR method extracted a significant risk factor for diabetes.

There are several dietary factors in nutritional epidemiology which were accounted for being associated with non-communicable disease risk factors, for instance major energy sources and differences in the mix of energy sources among human populations are correlated with the risk of non-communicable diseases. Previous studies indicated that a higher intake of polyunsaturated fat could be a protective factor against metabolic derangements. Likewise as higher intake of fiber was associated with reduced diabetes risk in some large cohort studies, we chose this item as another response variable. We also included dietary calcium intake, which is believed to alter the susceptibility to obesity and detrimental metabolic effects. Higher risk of non-communicable diseases is thought to be associated with higher intakes of “fat”, “cholesterol” and lower consumption of “polyunsaturated to saturated fat intake”, “calcium intake”, and “fiber intake” in day-to-day diet. Therefore, we chose fat intake, PUFA to SFA ratio, calcium, cholesterol, and fiber intake as the second set of variables in the RRR analysis in order to identify dietary patterns that are meaningful for different components of the metabolic syndrome. In the present study, we applied the RRR method to food intake data from two 24 hours recalls and we focused on the variation in fat intake, polyunsaturated fat/saturated fat ratio, cholesterol, fiber, and calcium intake. We also examined the performance of derived dietary patterns in association with different anthropometric indices.

MATERIALS AND METHODS

Samples

This is a cohort study of dietary patterns among samples within the framework of Tehran Lipid and Glucose Study (TLGS); a population-based cohort study on 15005 representative samples residing in district no. 13, Tehran. The aim of TLGS is to determine the prevalence of non-communicable diseases” risk factors and to develop healthy lifestyle to tackle the risks through appropriate interventions. The TLGS used a multistage cluster random sampling method and subjects were evaluated in the cross-sectional phase 1 of TLGS from 1999 to 2001. A sub-sample of 1476 individuals were recruited in a random sampling method for the study of dietary intakes (among whom there were 862 individuals over 18 years old: 379 males and 483 females). Dietary intake was collected by 24 hour dietary recalls on two non-consecutive days. Data collection was performed by trained dieticians through a face-to-face interview. The TLGS continued to follow up the samples through 3-year intervals, TLGS phase 3 was conducted in the 2005-2007 period. In order to study the contribution of dietary patterns in longitudinal changes of anthropometric indices, we have included all individuals over 18 years old, with accessible dietary intake data from TLGS phase 1 who had records from physical examination both in TLGS phase 1 and 3. We have excluded those individuals who underwent dietary and clinical intervention, those who had reported weight loss of greater than 6 percent of their body weight during the previous six months, and those who were on medications for either cardiovascular or metabolic disorders (including hypertension, hyperlipidemia of any kind, and hyperglycemia), and we have chosen those with accessible biochemical data from both phases. Finally, we came up with a total of 141 individuals for our study and studied anthropometric changes within these two intervals.

Data collection

In the two phases of the study, standing height was measured once using a portable stadiometer (Holtain Ltd, Crymmych, Pembrokeshire, UK) while the subjects were standing while not wearing shoes, and had the shoulders...
in a normal position. Weight was measured while the subjects were minimally clothed and not wearing shoes with digital scales and recorded to the nearest 100 g. Body mass index (BMI) was calculated as weight (kg) divided by height (m²). Waist circumference (WC) was measured at the narrowest level, and hip circumference was measured at the maximum level over light clothing, by using an unstretchable tape measure; all recorded to the nearest 0.1 cm. The procedures were carried out according to the study protocol. The measurements were taken by the same expert, and trained technician. The 24-hour dietary recalls were taken by trained dietitians who had ≥ 5 y experience in the Nationwide Food Consumption Survey project.

In both TLGS phases which were included in the current study, information on age, family history of diabetes or cardiovascular diseases, smoking habits, physical activity level, and medication with drugs altering metabolism and body weights were collected by means of questionnaires. With regard to smoking habits, subjects were categorized as daily smokers, ex-smokers, occasional smokers and non-smokers. Data on physical activity were obtained using the Lipid Research Clinic (LRC) questionnaire. This questionnaire is a simple and comprehensible measure including four questions; no special education is needed to complete this questionnaire. Subjects were classified as having light, moderate or severe physical activity based on their oral responses to the questionnaire according to LRC guidelines. These additional questionnaires were classified intakes of food items into these 16 food groups in grams per day, and the responses are intakes of nutrients in grams per day or ratios of nutrient intakes (as previously derived from Nutritionist IV software). We have used “fat intake”, “polyunsaturated to saturated fat intake”, “calcium intake”, “cholesterol intake” and “fiber intake” as response variables. With the application of the RRR technique we identified five factors (a.k.a. dietary patterns). The factor scores resulting from this analysis was used as a continuous variable and the following association study between dietary factors and anthropometric indices was performed on the quintiles of factor scores.

**Statistical analysis**

In the subsequent application, the predictors are intakes of food groups in grams per day, and the responses are intakes of nutrients in grams per day or ratios of nutrient intakes. The RRR identifies linear functions of predictors, which explain as much response variation as possible. The calculations of pattern scores by using the RRR method is based on the determination of eigenvalue and corresponding eigenvectors of the covariance matrix of predictors and responses, respectively. These dietary patterns were derived in SAS package (SAS ver 9.0, SAS Institute Inc., Cary, NC, USA). We came upon five dietary patterns (factors). The five dietary pattern were studied with respect to the extent each of them explained the variation in response variables. The first pattern explained 38.7% of variation in all food group variables, 74.7% of variation in total fat intake, 70.4% in cholesterol, 29.7% in calcium and 17.9% in fiber intake. Other derived factors explained an amount between 18.7% (for second factor) and 4.8% for the last (Table 1).

Factors loadings were also reported (Table 2). For the sake of sound interpretations we have reported only loadings >|0.17|. A high positive loading indicates a strong direct association between the food group and the pattern, whereas a high negative loading reflects a strong inverse association. First factor consists of sources of hydrogenated and saturated fat, egg, red and processed meat, refined carbohydrates, vegetables, whole grain and starchy vegetables. We called this pattern a “traditional pattern” consisting of a wide range of food items and sources of saturated fat. The second factor pertains to plant oils, starchy vegetables, legumes, other vegetables, salty snacks, fruits and nuts. It shows negative loadings on dairy products. With respect to the loadings on plant oil and high correlation with PUFA to SFA we named this as “fiber and PUFA pattern”. Factor 3 covers fruits and vegetables, dairy and whole grain, as well as negative, high loadings on plant oil and egg. We called this one “fiber

**Table 1.** Percentage of explained variation in dietary intakes (response) variables by each derived factors from reduced rank regression analysis

<table>
<thead>
<tr>
<th>Reduced rank regression</th>
<th>Explained variation in</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fat</td>
</tr>
<tr>
<td>Factor 1</td>
<td>74.7</td>
</tr>
<tr>
<td>Factor 2</td>
<td>77.3</td>
</tr>
<tr>
<td>Factor 3</td>
<td>82.6</td>
</tr>
<tr>
<td>Factor 4</td>
<td>85.6</td>
</tr>
<tr>
<td>Factor 5</td>
<td>96.1</td>
</tr>
</tbody>
</table>

*The explained variations are cumulative figures.
and dairy pattern”. Factor 4 in our analysis showed high, positive loading on dairy, and egg, then with plant oil and had negative loading with saturated and trans fat, refined carbohydrates, vegetables, and fruits. This was called “dairy pattern”. Factor 5 pertained to egg with highest positive loading, then with fruits and salty snacks, and negative loading to dairy, plant and saturated oil, and red meat. We called this as “egg pattern”.

Derived dietary pattern scores were treated as continuous variables and partial correlation was assessed with the response variables. In the next step, quintiles of pattern scores were determined and the trends for different anthropometric variables, as well as dietary variables were assessed across quintiles. These trend analyses were all controlled for several confounding variables including sex, age, total calorie intake, BMI in phase 1 of the study, and separately for physical activity level, smoking habits, medication in phase 1, and family history of diabetes or CVD. Two of the patterns (the first and the fifth pattern) showed significant change for anthropometry and hence, were tabulated furthermore to depict nutrient intakes across quintiles of factor scores. These secondary analyses were performed by SPSS ver 15 (SPSS, Chicago, USA).

RESULTS
A total of 141 individuals were included in this study (females comprised 67.4% of the sample). The mean±SD of age among male participants were 35.5±11.6 and for females, it was 34.7±14.9 years (Table 3). The changes of age among male participants were 35.5±11.6 and for females comprised 67.4% of the sample. The mean±SD across quintiles of factor scores. These secondary analyses were performed by SPSS ver 15 (SPSS, Chicago, USA).

Table 2. Loading from any of the five derived factors by reduced rank regression analysis for the 16 defined foods groups

<table>
<thead>
<tr>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
<th>Factor 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refined carbohydrates</td>
<td>0.32</td>
<td>NR†</td>
<td>NR</td>
<td>-0.28</td>
</tr>
<tr>
<td>Whole grains</td>
<td>0.25</td>
<td>NR</td>
<td>0.31</td>
<td>NR</td>
</tr>
<tr>
<td>Vegetables</td>
<td>0.26</td>
<td>0.31</td>
<td>0.35</td>
<td>-0.21</td>
</tr>
<tr>
<td>Starchy vegetables</td>
<td>0.20</td>
<td>0.33</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Legumes</td>
<td>NR</td>
<td>0.30</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Plant oil</td>
<td>NR</td>
<td>0.63</td>
<td>-0.40</td>
<td>0.39</td>
</tr>
<tr>
<td>Saturated and trans fat sources</td>
<td>0.59</td>
<td>NR</td>
<td>NR</td>
<td>-0.38</td>
</tr>
<tr>
<td>Egg</td>
<td>0.34</td>
<td>NR</td>
<td>-0.41</td>
<td>0.27</td>
</tr>
<tr>
<td>Red meat, Processed meat</td>
<td>0.30</td>
<td>NR</td>
<td>-0.18</td>
<td>NR</td>
</tr>
<tr>
<td>White meat</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Milk, milk product</td>
<td>NR</td>
<td>-0.30</td>
<td>0.34</td>
<td>0.65</td>
</tr>
<tr>
<td>Fruits, dried fruits</td>
<td>NR</td>
<td>0.26</td>
<td>0.42</td>
<td>-0.17</td>
</tr>
<tr>
<td>Jam, canned fruit and sweet snacks</td>
<td>NR</td>
<td>NR</td>
<td>-0.18</td>
<td>NR</td>
</tr>
<tr>
<td>Salty snacks</td>
<td>NR</td>
<td>0.25</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Nuts and seeds</td>
<td>NR</td>
<td>0.21</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Coffee and tea</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>

† NR; Not Reported: loading factors < |0.17| are not depicted in the table for the ease of interpretation and reading.

except for fiber intake. Female individuals generally consumed a greater ratio of polyunsaturated fat to saturated fat and the difference was significant (for p-values refer to Table 3).

The five dietary patterns were studied with respect to the extent that each of them explained the variation in response variables (Table 1). In comparison with table 4 which depicts the correlation coefficients adjusted for age, total energy intake and BMI, the first pattern shows significant, positive correlation with total fat, cholesterol and calcium intake. The second pattern explains less than 10% of variation in total fat, cholesterol and calcium intake (Table 1), but nevertheless, shows significant negative correlation with cholesterol and calcium (Table 4). This pattern explains a variation of 56.4% in PUFA/SFA ratio, and 26.3% in fiber intake with which correlation is positive and significant. The third pattern explains 10.2% of cholesterol, and 23.3% of fiber and 19% of calcium intake (Table 1). This pattern only shows significant (and negative) correlation with PUFA/SFA ratio to which it explains less than 10% of variation (Table 4). The correlation with fat, fiber and calcium intake were positive and significant and with cholesterol and PUFA/SFA ratio were negative. The fourth factor explained 12% of PUFA/SFA ratio, and 20.7% of calcium intake. The explained variations for other response variables were less than 10%. This pattern showed correlations for these two variables in positive direction, as well as cholesterol intake and showed negative correlation with fiber, with the two latter the correlations were poor. Finally, the fifth pattern explained 10.6% of fat and 10.1% of cholesterol intake, showed positive correlation with cholesterol and fiber, and negative with fat.

The first three patterns explained 70.4% of variations to all response variables and the number grows to 83.8% for all patterns. The last pattern only explains 4.8% of variation to all responses (Table 1). Table 2 shows the factor loadings for each food group in detail. For further reference on the relation between the factor loadings and
the assigned names please refer to Table 2 in comparison with Table 4.

In the next step, we analyzed the change in three anthropometric indices with respect to these derived patterns. In the complete sample, the change in BMI (±SD), WHR and WC were respectively as follows: 1.7±2.2, 0.043±0.070, and 6.1±8.4 (data not shown in Tables). Table 5 depicts the change in anthropometry across quintiles of scores for the five patterns. Here we employed sex, age, physical activity, smoking status, family history of diabetes mellitus or cardiovascular diseases, total energy intake, BMI and medication as covariates. We reported three sets of analysis; first one only adjusted these covariates, the second one additionally adjusted for WHR in the first phase of TLGS. And the third model contained the previous covariates and waist circumference in TLGS-1 (Table 5).

Pattern 1 (traditional pattern) showed an increase in BMI, WHR, and WC across quintiles of intake which was statistically significant. This means as the score in the traditional pattern increased, the change in anthropometric indices were positive and greater in magnitude. For changes in WHR and WC, we performed further adjustments by adding initial WHR and WC into the analysis, respectively. The effect for the traditional pattern was attenuated for WC. By more contemplation in other patterns, it seems that the dairy pattern also showed an increase in trend across quintiles, but the magnitude of effect was not statistically significant. For the fifth pattern (egg pattern), WHR and WC showed an increasing trend after further adjustments. It should be noted that, the fifth pattern showed a decreasing significant trend in WC among female individuals which was not shown in presented data. Decreasing trend for WHR and WC was seen for the fiber and PUFA pattern, and the fiber and dairy pattern, respectively which were not statistically significant.

Table 4. Correlation coefficients between reduced rank regression factors and the five dietary (response) variables

<table>
<thead>
<tr>
<th>Factor</th>
<th>fat</th>
<th>Cholesterol</th>
<th>Polyunsaturated fat intake/saturated fat intake</th>
<th>Fiber</th>
<th>Calcium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1</td>
<td>0.479*</td>
<td>0.626*</td>
<td>0.590*</td>
<td>-0.067</td>
<td>0.486*</td>
</tr>
<tr>
<td>Factor 2</td>
<td>-0.124</td>
<td>-0.452*</td>
<td>-0.228**</td>
<td>0.682*</td>
<td>-0.255**</td>
</tr>
<tr>
<td>Factor 3</td>
<td>-0.543*</td>
<td>-0.433*</td>
<td>0.359*</td>
<td>-0.204**</td>
<td>0.721*</td>
</tr>
<tr>
<td>Factor 4</td>
<td>-0.051</td>
<td>0.177*</td>
<td>0.009</td>
<td>0.447*</td>
<td>-0.115</td>
</tr>
<tr>
<td>Factor 5</td>
<td>-0.458*</td>
<td>0.356*</td>
<td>0.009</td>
<td>0.447*</td>
<td>-0.115</td>
</tr>
</tbody>
</table>

* p<0.001, ** p<0.05
In order to better elucidate the relationship between the derived patterns (which showed significant effect on anthropometric indices) and the dietary variables, we presented trends in total energy and the response variables across quintiles of pattern scores in table 6. Traditional pattern shows an increase in total energy, fat, cholesterol, fiber and calcium intake with significant \( p \)-value for the trend. The fifth pattern, named egg pattern, shows an increasing trend in cholesterol and decreasing in total fat intake. The indicated trend was more pronounced among female individuals (data not shown). The consumption of fiber and calcium decreased across these quintiles significantly.

**DISCUSSION**

All anthropometric indices increased through the 6-year follow-up of our study, which is in line with the report from the same samples on a 3.6 year follow-up by Hadaegh et al.\(^a\) With regard to analysis of anthropometric changes across quintiles of factor scores, only the first and the last patterns (namely “traditional” and “egg pattern”) gained significant trends. The first dietary pattern derived by the means of RRR analysis (traditional pattern) was significantly associated with increase in anthropometric indices at 6-year follow-up of our study samples. The first RRR pattern consisted of a mixed diet containing refined carbohydrates, starchy vegetables, other types of vegetables, whole grains, red and processed meat, saturated and trans fat sources, as well as eggs. This is the dietary pattern mostly commonly practiced by Tehranian adults. Traditional pattern shows an increase in total energy, fat, cholesterol, fiber and calcium intake with significant \( p \)-value for trend across quintiles in terms of pattern score. Since this pattern consisted of foods containing carbohydrates, fiber, fat and cholesterol, this finding could be justified. Note that the anthropometric indices increased across quintiles of the traditional pattern score, and 74.7% of fat intake, 70.4% of cholesterol and 38.7%
Dietary patterns and obesity in TLGS cohort

of all response variables were explained by this pattern. One should bear in mind that RRR analysis generally shows a lower percentage of variance in food groups (as predictor variables), but its strength is in explaining a high percentage of variance in response variables.

In a cross-sectional report from Tehranian middle aged females, Esmaillzadeh et al. found that a pattern high in refined grains, potato, tea, whole grain, hydrogenated fat, legumes and broth showed a decrease in BMI (and not WHR). In our analysis this trend is positive (longitudinal study), maybe because our pattern correlates highly with fat and cholesterol intake; besides in the mentioned study, the energy intake decreased across quintiles of this pattern score. The main factor loading in our study was on saturated and trans fat (0.6 loading) and moderate loadings on other components, while in the other study the highest loadings were on potato, refined, and whole grains. In a separate study by Pryer et al., the traditional pattern was found to include high consumption of white bread and refined cereals, butter, margarines, confectionaries, tea, high fat dairy, red and processed meats, chips, as well as vegetables, fruits and nuts only among males, and eggs only among females; and higher factor scores were associated with increased BMI. Although named as traditional pattern by these authors, this does not completely share the specifications of our first pattern. This is due to rather arbitrary labeling of the patterns derived according to the culinary use of nutrients; for instance, in the study by Heitmann and colleagues, the traditional pattern was identified by high consumption of wax beans, green beans, potatoes, with a speculated high intake of fiber, and as the pattern was associated with lower BMI.

For fiber and PUFA, as well as fiber and dairy patterns there was a decreasing trend (although not significant) in WHR and waist circumference changes at 6-year follow-up. Liu in a sample of 74091 followed for 12 years found that high and increasing fiber intake prevents weight gain among women and this is independent from initial weight. Schulze et al. studied dietary pattern on large samples in Germany in a cohort study of 4-year follow-up by means of RRR. They used dietary intakes of total fat, carbohydrate and fiber as response variables. They have found a dietary pattern which was distinguished by high intakes of fiber and low-fat foods. Mean annual weight gain was negatively associated with increasing pattern score. This one factor contributed to the changes in weight in their follow-up. Wirfält et al. also derived six factors by means of cluster analysis and found a “fiber bread” pattern inversely associated with central obesity. This is in line with our results. Note that fiber and PUFA pattern was highly correlated with PUFA/SFA and fiber intake, and fiber and dairy pattern with fiber, and calcium. There is another non-significant trend between dairy pattern with BMI (increasing across quintiles

Table 6. Intake of different dietary variables across quintiles of pattern score

<table>
<thead>
<tr>
<th>Dietary variable</th>
<th>Quintiles of pattern score</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total energy intake (Kcal/day)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pattern 1 (traditional pattern)</td>
<td>1805±246</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Pattern 5 (egg pattern)</td>
<td>1946±292</td>
<td>0.521</td>
</tr>
<tr>
<td>Fat intake (mg/day)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pattern 1 (traditional pattern)</td>
<td>47.7±15.9</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Pattern 5 (egg pattern)</td>
<td>85.7±27.6</td>
<td>0.019**</td>
</tr>
<tr>
<td>Cholesterol intake (mg/day)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pattern 1 (traditional pattern)</td>
<td>184±41</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Pattern 5 (egg pattern)</td>
<td>300±96</td>
<td>0.050</td>
</tr>
<tr>
<td>Polyunsaturated fat intake/saturated fat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pattern 1 (traditional pattern)</td>
<td>0.266±0.047</td>
<td>0.308</td>
</tr>
<tr>
<td>Pattern 5 (egg pattern)</td>
<td>0.258±0.069</td>
<td>0.958</td>
</tr>
<tr>
<td>Fiber intake (mg/day)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pattern 1 (traditional pattern)</td>
<td>10.4±2.0</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Pattern 5 (egg pattern)</td>
<td>14.8±3.2</td>
<td>0.014**</td>
</tr>
<tr>
<td>Calcium intake (mg/day)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pattern 1 (traditional pattern)</td>
<td>514±67</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Pattern 5 (egg pattern)</td>
<td>639±109</td>
<td>0.037**</td>
</tr>
</tbody>
</table>

*significant as p-value is <0.001
**significant as p-value is <0.05
of pattern score) and WHR (opposite trend). This pattern was found to explain only 8.6% of variance of food intakes, and despite its 20.7% explanation of calcium intake, previous studies associated the use of high calcium-content products with decrease in obesity indices\textsuperscript{41-44} so the results should be further scrutinized. Note that apart from dairy products, there were moderate loadings for plant oil and egg groups, and the pattern explained 12% of variance in PUFA/SFA intake and 3% in fat. The fat component of this pattern might in part explain the contradictory results.

Since the first pattern alone had the highest percentage of variation in food groups, it deserved further scrutiny. Nevertheless, the last pattern was also included in analysis of food intake across quintiles of pattern scores (Table 6), but bear in mind that this explained only 4.8% of variation in food intake.

The fifth pattern, named egg pattern, shows an increasing trend in cholesterol and decreasing in total fat intake (which are in accordance by the respective correlation coefficients). The consumption of fiber and calcium decreased across these quintiles significantly. This pattern is recognized by highest loadings to egg, and moderate to salty snacks and fruits, versus negative loadings with plant oil, saturated and trans oil, red and processed meat, and dairy which leads to the negative correlation with fat intake, and positive with cholesterol and fiber, for the first two showing more than 10% of explained variation, which shows high effect on these response variables. Yet, the effect of this pattern on anthropometric parameters is paradoxical; decrease in BMI especially among women, increase in WHR in both genders. This might in part be related to the dual association of this pattern with cholesterol, on one side, and total fat, on the other. Other investigators did not identify a pattern similar to. Usually, high consumption of eggs is coupled with eating more fat from animal sources, for example, there is a reported high fat pattern that consist of eggs, bacon and sausage which was associated with increased BMI\textsuperscript{56} and another example is the western pattern characterized by red and processed meat and eggs, again associated positively with BMI in the literature\textsuperscript{57}.

As stated previously, the investigators assigned names on either derived factors based on current dietary guidelines and culinary specifications in certain populations. Typically, a healthy dietary pattern is characterized by high intake of high-fiber grains and cereals, reduced fat dairy, fruits, and vegetables, and by low intake of refined grains, meats, soda, and high-fat dairy\textsuperscript{57} While a diet based on refined grains, red and processed meat, butter, whole milk, sweets, fast food, and hydrogenated fat is considered as the western diet\textsuperscript{58}. The analysis on our samples revealed a more mixed-type diet patterns\textsuperscript{59} as fiber-rich whole grain and vegetables made up the major dietary pattern named traditional in accordance to studies from other samples in Tehran\textsuperscript{58}. Another point to pay attention to was the presence of fiber rich food groups especially grains and vegetables, in three of the patterns, yet the average fiber intake was below the recommendations. The same concern is raised about dairy consumption that in spite of presence in two of the patterns, the calcium intake is well below recommendations in these samples (note that two other patterns were associated with negative loadings on dairies). Finally, it should be noted that the consumption of salty snacks was also high in fiber and PUFA, and in the egg patterns. Hence, there are several challenging characteristics to the dietary patterns of Tehranian adults which need special intervention/education in the years to come to tackle the diet-related health problems of this community.

Determination of the number of factors is rather arbitrary and according to the justification of investigator from the percentage variation and the following association results. The RRR analysis has been proved to be a strong method to analyse dietary patterns in comparison to other data reduction techniques\textsuperscript{24,28} For instance, Netton compared Principle Component Analysis and RRR in relation to markers of atherosclerosis\textsuperscript{60}. Only the first RRR pattern showed association with coronary artery calcification. In a comparison by Hoffmann and his team on the effects of diet on all-cause mortality rate\textsuperscript{28} RRR outran PCA method on determining the associations of diet and the outcome. It is worth highlighting that the extent of variance explained by the derived factors in this study was 38.7% for the first factor and 83.8% for all factors, which is similar to the studies using the same method thus far\textsuperscript{24,28}. Hoffmann and co-workers used percentage of energy from different macronutrients as response variables in RRR and reported that the first RRR pattern explained 30.8% of variation in energy sources\textsuperscript{28}. In another study by Hoffmann et al\textsuperscript{59} published in 2004, the researchers used RRR to derive the patterns in their study on type 2 diabetes mellitus and four factors were derived. The RRR factors explained as much as 93.1% of variation in responses. In our case, RRR-derived factors accounts for 83.8% of variation in total.

The special strength of the current study was its prospective nature which enabled us to assess the change of anthropometric indices through the 6-year follow up. Yet, there are several limitations: first of all the duration of follow-up might not be as long as to reveal excessive changes in anthropometry. The reproducibility of the results in other populations is questioned, which will be tackled by utilization of a unified food item grouping, but may not be attainable because of the difference between culinary items in different cultures, yet may be attenuated if more unified patterns are employed. The selection of response variables, although based on previous evidences, is arbitrary; one could decide to use another set of variables as response set. However, the five response variables are those with higher citation, but the authors cannot exclude other potentially important factors as response variables to be included in such analysis. Finally, we should also bear in mind the error that the employment of the 24-hour dietary recall would impose on our findings. This is by no means a complete tool to assess dietary intakes and the conclusions should be made accordingly. Since this is a historical cohort study, we had to trim down the samples to those whose with anthropometric data of both phase 1 and 3 of TLGS as well as dietary intake data at hand. This particular consideration reduced our sample size to 141 individuals and is a limitation of our study. In conclusion, the RRR method is a powerful tool to derive dietary patterns in nutritional epi-
Demiology. It is a flexible method that combines the strength of factor analysis to consider inter-correlations of dietary intakes and the advantage of diet-quality scores to allow for use of biologically plausible prior information. This method has enabled us to detect a major dietary pattern (namely traditional) which shows positive association with anthropometric indices in a 6-year follow-up.

**AUTHOR DISCLOSURES**

The TLGS was supported by grant no 121 from the National Research Council of the Islamic Republic of Iran and by the combined support of the National Research Council of Islamic Republic of Iran and Endocrine Research Center of Shahid Beheshti University of Medical Sciences. This study was also supported by Iran University of Medical Sciences. There is no conflict of interest to declare.

**REFERENCES**


Original Article

Dietary patterns by reduced rank regression predicting changes in obesity indices in a cohort study: Tehran Lipid and Glucose Study

Roya Sherafat-Kazemzadeh MD1,2, Shahryar Eghtesadi PhD1, Parvin Mirmiran PhD2, Mahmoud Gohari PhD3, Sara Jalali Farahani BS2, Firoozeh Hosseini Esfahani MS2, Mohammad Reza Vafa PhD1, Mehdi Hedayati PhD4, Fereidoun Azizi MD4

1Nutrition Department, Faculty of Health, Iran University of Medical Sciences, Tehran, Iran
2Nutrition Unit, Obesity Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
3Biostatistics Department, Faculty of Management, Iran University of Medical Sciences, Tehran, Iran
4Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

減維度迴歸導出的膳食模式預測肥胖指標的改變：德黑蘭脂質與血糖的世代研究

目的：檢測德黑蘭成人於 6 年的追蹤研究中，其膳食模式和肥胖指標（身體質量指數、腰圍、腰臀比）之間的相關性。方法：在德黑蘭世代研究的架構內，選出 141 個成人（平均追蹤時間為 6.6±0.9 年），各有 2 次起始點的 24 小時飲食回憶資料，及起始點和追蹤結束時的肥胖指標數據。膳食攝取被轉換為各食物項目的克數重量，並歸類成 16 個食物群組。以總脂肪、多元不飽和/飽和脂肪比率、膽固醇、膳食纖維及鈣的攝取量為依變項，使用減維度迴歸分析衍生出五類膳食模式。以在食物群組中持有對應的因素負荷 ≥|0.17| 來導出各因素（膳食模式）。將因素分數劃為五分位，詳細地檢測肥胖指標的改變。結果：在第一類「傳統膳食模式」的因素中，以精制醣類、全穀類、富含澱粉的蔬菜、其他蔬菜、紅肉和加工肉類、飽和/反式脂肪、蛋類的因素負荷較高。所有的肥胖指標都隨著此膳食模式分數五分位的增加而有增加的趨勢。第五類膳食模式（蛋類模式）在蛋類、鹹的零食類、水果類和乾燥水果有較高的因素負荷，而在紅肉和加工肉類、飽和及反式脂肪、植物油及乳製品則有負的因素負荷。在校正過可能的干擾因子後，該膳食模式顯示出了腰圍和腰臀比有增加的趨勢。其他的膳食模式和肥胖指標則無明顯的趨勢現象。結論：結果顯示在德黑蘭地區，以傳統的膳食模式為主，且與肥胖指標的增加相關。

關鍵字：膳食模式、人體測量、肥胖、減維度迴歸分析、世代研究