Identifying the pattern of unhealthy dietary habits among an Iranian population: A latent class analysis

Ali Gholami1,2, Masoudreza Sohrabi3, Abbas Abbasi-Ghahramanloo4,5,2, Farhad Moradpour5,2, Saeid Safiri5, Mansooreh Maadi5, Maryam Khazaee-Pool7, Ali Ghanbari5, Farhad Zamani*3

Received: 1 Mar 2018    Published: 10 Aug 2018

Abstract

Background: An unhealthy diet is one of the most important risk factors for chronic diseases. The goal of this study was to use the latent class analysis (LCA) modeling to define unhealthy diet habits among an Iranian population.

Methods: This cross-sectional study was conducted within the framework of Amol (North of Iran) cohort health study (Phase 1). The participants aged 10 to 90 years. All participants provided written informed consent. Latent class analysis was used to classify the participants of the study. All analyses were conducted by PROC LCA in SAS 9.2 software. Significance level was set at 0.05.

Results: The mean age of the participants was 42.58±17.23 years. Four classes of individuals with different diet habits were identified using LCA modeling: class 1: individuals with healthy diet patterns (92.6%); class 2: individuals with slightly unhealthy diet habits (6.3%); class 3: individuals with relatively unhealthy diet habits (0.8%); and class 4: individuals with unhealthy diet habits (0.2%). Being female and alcohol consumption increased the odds of membership in latent classes 2, 3, and 4 compared to class 1. Physical activity decreased the odds of membership in classes 3 and 4 compared to class 1.

Conclusion: Overall, almost more than 7.4% of all participants had some degree of unhealthy dietary habits, and some variables acted as risk factors for membership in risky classes. Therefore, focusing on these variables may help design and execute effective preventive interventions in groups with unhealthy dietary habits.

Keywords: Unhealthy diets, Latent class analysis, Diet pattern

Copyright© Iran University of Medical Sciences


Introduction

An unhealthy diet is one of the most important risk factors for most chronic conditions, such as coronary heart disease, diabetes, and cancers worldwide (1).

Nowadays, many attempts have been done to produce dietary guidelines for chronic disease prevention. However, there is less concern about measuring the overall diet quality, and measuring nutrients is usually used to assess dietary quality (2).

Traditionally, the majority of studies on nutrition, including interventional, clinical, and epidemiological studies, focused on the diagnosis of specific mechanisms and special health effects of the single nutrients. However, the relationship between a special nutrient and chronic diseases may be problematic and confusing (3-6).

In nutritional research, specific traditional methods are selected, which might lead to underestimation of the effect size of natural foods, leading to inaccurate interpretation of the study outcomes (4, 6, 7). People do not eat iso-
lated nutrients; they eat meals with plenty of nutrients and there might be synergism among them (8). As it is difficult to show collinearity between individual nutrient intakes, multiple tests may be needed to find the acceptable effect size, and this may lead to increased probability of type 2 error in the studies (8, 9).

Most recent studies have emphasized the realistic dietary pattern as a determinant of health and diseases in different populations (10-14). These dietary pattern studies better indicate people’s real-life scenario, as the nutrients are consumed as food and meals. Thus, this scenario allows us to investigate the health impact of the overall diet. In addition, studying dietary patterns could have more impact on public health; hence, the results of this study could be easily used in public health (15, 16).

Dietary pattern cannot be extracted directly; however, by using statistical methods, dietary information may be categorized in mutually exclusive groups. Available techniques for determining dietary patterns are relatively new and still developing. Three methods are proposed in the literature: factor analysis, cluster analysis, and diet indexes (8).

The aim of cluster analysis is to create the K-mean exclusive classes of people according to their shared dietary information. Among the techniques of cluster analysis, model-based clustering methods, such as latent class analysis (LCA), were used rarely despite their advantages (17-21). The present study was conducted to use LCA modeling to define unhealthy diet patterns among the Iranian population.

Methods
This cross-sectional study was conducted within the framework of Amol (North of Iran) cohort health study (Phase 1). Amol cohort study was designed and performed by Gastro Intestinal and Liver Disease Research Center (GILDRC) in Firoozgar hospital (Tehran). A total of 7104 participants aged 10-90 years were selected from Amol health care centers using random sampling. Of all participants at baseline, 961 were excluded, and data of 6143 participants were used in the present study (Figure 1).

Six binary variables were used to assess dietary pattern. These variables were (a) skipping breakfast (yes, no), (b) skipping dairy products (yes, no), (c) skipping fruit and vegetables (yes, no), (d) skipping fish (yes, no), (e) skipping healthy snacks (if any) and fast food intake (yes, no). Such variables as gender of the participants, having regular physical activity, current cigarette smoking, alcohol use, and having a chronic disease were used as covariates. In this paper, alcohol use was assessed by 3 questions: (1) How much do you drink alcohol per month? (about _____ times per month) (3) and for How long? The answers were converted into 2 categories: “positive history (1)”, those in this category drank alcohol at least once per month and for at least 6 months; and “negative history” (0), those in this category drank alcohol for less than 6 months or had never tried drinking.

Statistical analysis
In this study, LCA was used for data analysis. This analysis is a latent categorical variable model and, in the final step, it classifies the participants based on their homogeneity in the response pattern. In addition to find the measurement error, it determines whether correlations between observed variables can be explained by latent variables. To determine the best model and calculate the indices, LCA could be run by various iterations for the number of identified classes and compare the frequencies of the observed response pattern with expected one. LCA determined a statistic called G2 and 2 indices could be calculated based on this statistic: Akaike information criterion (AIC) and Bayesian information criterion (BIC). These indices are used for model selection; and, for all of the them, a smaller value represents a more optimal balance of model fit and parsimony. In other words, a model with the minimum AIC or BIC might be selected. To run LCA, 6 binary observable variables (indicators) were used for subgrouping the participants based on an unhealthy dietary pattern as a latent variable. These indicators were skipping breakfast, dairy products, fruit and vegetables, fish consumption, and healthy snack, and fast food intake.

To perform simple statistical analysis, descriptive analysis was used in SPSS 16. To assess multiple associations of all the covariates with latent classes of unhealthy dietary habits, multinomial logistic regression analysis was used.

To report the frequency of demographic factors and observable variables, SPSS 16 and to subgroup the participants, the PROC LCA were used in SAS 9.2 software (SAS Institute Inc. Cary, NC, USA). Significant level was set at 0.05.

Results
The mean age of the participants was 42.58±17.23 years (range: 10-90). More information about participants’ characteristics is presented in Table 1. The prevalence of skipping breakfast, dairy products, fruit and vegetables, fish consumption, skipping snack, and fast food intake is illustrated in Table 2. This table shows that some of the unhealthy dietary components are more prevalent than others. For example, the prevalence of eating fast food was 20.9%.

A total of 64 possible response patterns can exist with 6 dichotomous variables. The LCA was fitted with classes ranging from 1 to 8. For model selection, the G2, AIC, and BIC indices were computed separately for each model (Table 3). The LCA model is based on contingency table and is similar to the cross tabulation tests of independence and the expected and observed frequencies of the response pattern. The expected cell counts are estimated according to the specified model and estimated parameters. Based on this test, a p-value could be calculated for the goodness of the model. According to this value, we should select a nonsignificant model. Among 5 nonsignificant models (3, 4, 5, 6 and 7 latent class models), considering G2, AIC, and BIC indices and interpretability of the results of each model, the authors concluded that 4 latent classes were appropriate. As demonstrated in Table 3, the results of the
selected model (4-latent class) indicated that the difference between the expected and observed frequencies of response pattern was not statistically significant ($G^2 = 47.61, df = 36, p = 0.093$).

Table 4 presents a description of latent classes. The prevalence of each latent class appeared in the first section of this table. The second section of this table shows the probabilities of the “yes” response to unhealthy dietary components. To achieve the probabilities of “no” response, the item-response probabilities should be subtracted from 1. These probabilities are used for labeling and interpretation of the findings. Larger than 0.5 probabilities appear in bold font to indicate the overall pattern.

### Table 1. Characteristics of the study population (n = 6122)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Characteristic</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>Male</td>
<td>3497</td>
<td>57.1</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>2625</td>
<td>42.9</td>
</tr>
<tr>
<td>Physical activity</td>
<td>No</td>
<td>4621</td>
<td>75.5</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>1501</td>
<td>24.5</td>
</tr>
<tr>
<td>History of cigarette smoking</td>
<td>No</td>
<td>5269</td>
<td>86.1</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>853</td>
<td>13.9</td>
</tr>
<tr>
<td>Alcohol use</td>
<td>No</td>
<td>5796</td>
<td>94.7</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>326</td>
<td>5.3</td>
</tr>
<tr>
<td>Having a chronic disease</td>
<td>No</td>
<td>2252</td>
<td>36.8</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>3870</td>
<td>63.2</td>
</tr>
</tbody>
</table>

### Table 2. Percentages of participants responding “Yes” to questions about unhealthy diet pattern

<table>
<thead>
<tr>
<th>Items</th>
<th>Male (n=3497)</th>
<th>Female (n=2625)</th>
<th>Total (n=6122)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breakfast skipping</td>
<td>217 (6.2)</td>
<td>201 (7.7)</td>
<td>418 (6.8)</td>
</tr>
<tr>
<td>Skipping dairy products (At least one serving per day)</td>
<td>46 (1.3)</td>
<td>48 (1.8)</td>
<td>94 (1.5)</td>
</tr>
<tr>
<td>Skipping fruit and vegetable (At least one serving per day for fruit and vegetable.)</td>
<td>40 (1.1)</td>
<td>30 (1.1)</td>
<td>70 (1.1)</td>
</tr>
<tr>
<td>Skipping fish consumption (At least one serving per week.)</td>
<td>272 (7.8)</td>
<td>197 (7.5)</td>
<td>469 (7.7)</td>
</tr>
<tr>
<td>Skipping healthy snack</td>
<td>77 (2.2)</td>
<td>47 (1.8)</td>
<td>124 (2.0)</td>
</tr>
<tr>
<td>Fast food consumption</td>
<td>699 (20.0)</td>
<td>580 (22.1)</td>
<td>1279 (20.9)</td>
</tr>
</tbody>
</table>

### Table 3. Comparison of LCA models with different latent classes based on model selection statistics

<table>
<thead>
<tr>
<th>Number of latent class</th>
<th>Number of parameters estimated</th>
<th>$G^2$</th>
<th>df</th>
<th>AIC</th>
<th>BIC</th>
<th>P-value</th>
<th>Maximum log-likelihood</th>
<th>Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>280.21</td>
<td>57</td>
<td>292.21</td>
<td>332.53</td>
<td>&lt;0.001</td>
<td>-7793.15</td>
<td>1.00</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>79.99</td>
<td>50</td>
<td>105.99</td>
<td>193.35</td>
<td>0.004</td>
<td>-7693.04</td>
<td>0.88</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>56.02</td>
<td>43</td>
<td>96.02</td>
<td>230.41</td>
<td>0.087</td>
<td>-7681.05</td>
<td>0.79</td>
</tr>
<tr>
<td>4</td>
<td>27</td>
<td>47.61</td>
<td>36</td>
<td>96.61</td>
<td>278.04</td>
<td>0.093</td>
<td>-7674.35</td>
<td>0.89</td>
</tr>
<tr>
<td>5</td>
<td>34</td>
<td>33.31</td>
<td>29</td>
<td>101.31</td>
<td>329.78</td>
<td>0.265</td>
<td>-7669.70</td>
<td>0.87</td>
</tr>
<tr>
<td>6</td>
<td>41</td>
<td>27.86</td>
<td>22</td>
<td>109.86</td>
<td>385.36</td>
<td>0.180</td>
<td>-7666.97</td>
<td>0.80</td>
</tr>
<tr>
<td>7</td>
<td>48</td>
<td>24.07</td>
<td>15</td>
<td>120.07</td>
<td>442.61</td>
<td>0.063</td>
<td>-7665.08</td>
<td>0.74</td>
</tr>
<tr>
<td>8</td>
<td>56</td>
<td>20.14</td>
<td>8</td>
<td>130.14</td>
<td>499.72</td>
<td>0.009</td>
<td>-7663.11</td>
<td>0.61</td>
</tr>
</tbody>
</table>

### Table 4. The Four latent classes’ model of components of healthy/unhealthy diet

<table>
<thead>
<tr>
<th>Latent classes</th>
<th>Healthy diet</th>
<th>breakfast skippers</th>
<th>Fruit, vegetable and fish skimmen</th>
<th>Unhealthy diet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy diet</td>
<td>0.926</td>
<td>0.063</td>
<td>0.008</td>
<td>0.002</td>
</tr>
<tr>
<td>breakfast skippers</td>
<td>Probability of a “Yes”</td>
<td>0.702</td>
<td>0.057</td>
<td>0.539</td>
</tr>
<tr>
<td>Skipping dairy product</td>
<td>0.008</td>
<td>0.036</td>
<td>0.434</td>
<td>0.770</td>
</tr>
<tr>
<td>Skipping fruit and vegetables</td>
<td>0.004</td>
<td>0.022</td>
<td>0.605</td>
<td>0.358</td>
</tr>
<tr>
<td>Skipping fish consumption</td>
<td>0.063</td>
<td>0.180</td>
<td>0.793</td>
<td>0.248</td>
</tr>
<tr>
<td>Skipping healthy snack</td>
<td>0.018</td>
<td>0.035</td>
<td>0.000</td>
<td>0.608</td>
</tr>
<tr>
<td>Fast food consumption</td>
<td>0.203</td>
<td>0.281</td>
<td>0.307</td>
<td>0.132</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Covariates OR(95%CI)</th>
<th>OR(95%CI)</th>
<th>OR(95%CI)</th>
<th>OR(95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (being female) (P=0.007)</td>
<td>Reference</td>
<td>3.05(1.5-5.9)</td>
<td>1.33(0.6-2.8)</td>
</tr>
<tr>
<td>Physical activity (P=0.001)</td>
<td>Reference</td>
<td>3.42(1.9-6.1)</td>
<td>0.21(0.1-0.8)</td>
</tr>
<tr>
<td>History of cigarette smoking (P=0.871)</td>
<td>Reference</td>
<td>1.37(0.7-2.8)</td>
<td>1.24(0.5-2.9)</td>
</tr>
<tr>
<td>Alcohol consumption (P=0.001)</td>
<td>Reference</td>
<td>7.75(3.8-15.9)</td>
<td>4.93(2.0-11.9)</td>
</tr>
<tr>
<td>Having a chronic disease (P=0.047)</td>
<td>Reference</td>
<td>1.20(0.7-2.1)</td>
<td>1.34(0.7-2.7)</td>
</tr>
</tbody>
</table>

Note. LCA = latent class analysis; AIC = Akaike information criterion; BIC = Bayesian information criterion.

Note. The probability of a “No” response can be calculated by subtracting the item-response probabilities shown above from 1. Item-response probabilities >.5 in bold to facilitate interpretation.
Pattern of unhealthy dietary habits

individuals exhibiting a high probability of breakfast skipping. The third class described 0.8% of the individuals and was characterized by individuals exhibiting a high probability of skipping fruit and vegetables and skipping fish. The fourth class described 0.2% of the individuals and was characterized by individuals exhibiting a high probability of skipping breakfast, dairy products, and healthy snacks. In this class, the probability of skipping fruits, vegetables, and fish is high.

Last section of Table 4 shows the odds ratios for each covariate. In other words, latent class 1 was a reference. As demonstrated in this table, being a female increased the odds of membership in classes 2, 3, and 4 compared to class 1. Having physical activity decreased the odds of membership in classes 3 and 4 compared to class 1. Finally, alcohol consumption increased the odds of membership in classes 2 and 3 compared to class 1.

Discussion

The findings of this study demonstrated the prevalence of each dietary habit. Fast food intake was a common habit with a prevalence of 20.9% and skipping fruit and vegetables was uncommon with a prevalence of 1.1%. It seems that the prevalence of dietary habits is different across the world. For example, 76% of the English adult population did not eat enough fruit and vegetables (22).

Our study showed that prevalence of skipping breakfast was 6.8%. The results of the National Health and Nutrition Examination Survey (NHANES) during 1999-2002 indicated that 25% of young adults were breakfast skippers (23). However, the results showed that 1.5% of all participants skipped dairy products. A follow-up study among Australian adolescents showed that in a 14-year follow-up, 0.1% and in a 17-year follow-up, 0.2% of all participants skipped dairy products. These percentages are relatively higher among females (24). Our findings are consistent with those of this study.

Overall, weight gain in adolescents can be considered as a key factor of lifestyle. Fast food intake and skipping breakfast are 2 dietary factors identified as being problematic in this group. These factors may contribute to enhancing energy intake and weight gain (25-29). Our results indicated that the prevalence of fast food intake is remarkably high. Other studies had emphasized the increasing trend of fast food intake. For example, the Continuous Survey of Food Intake by Individuals (CSII) reported that the prevalence of fast food intake during 1977-1978 was 6.5%, and this rate increased to 19.3% during 1994-1996 (30). Another cross-sectional study revealed that among 5000 students nearly 75% reported eating fast food during the last 7 days (26).

According to above-mentioned comparisons, it seems that in some cases, there is a big difference between our results and those of other studies. This broad range may be related to the variety of the definition of each dietary habit, the age difference of the study samples, and the location of studies. Overall, the results of any country can be attributed to the specific lifestyle of the population and other related factors. Findings of our study revealed a high prevalence of some dietary habits, indicating the necessity of interventions of altering dietary habits.

Previous studies reported that considering the co-occurrence of unhealthy behaviors is an effective approach in preventive programs (31-33). The patterning and co-occurrence of dietary habits were analyzed in some studies (22, 34). In previous works, the clustering of lifestyle risk factors has been done by other methods. These studies indicated that there is an association between lifestyle indicators (22, 35-37). For example, Sanchez et al., in a study of patterns and correlates of physical activity and nutrition behaviors in adolescents, concluded that about 80% of the study participants had multiple unhealthy lifestyle behaviors and nearly 50% of them had at least 3 unhealthy lifestyle indicators (38). Other studies, including one conducted by Lotrean et al., found unhealthy alimentary behaviors and co-occurrence of several elements of this behavior (39).

Some studies assessed this co-occurrence differently using LCA. Laska et al., in a similar study on U.S college students, identified 4 latent classes for male and female students. Female classes were as follow: (1) poor lifestyle yet low-risk behaviors (40.0% of the females), (2) high risk (24.3%), (3) moderate lifestyle, few risky behaviors (20.4%), and (4) health conscious (15.4%). Male classes were as follow: (1) poor lifestyle, low risk (9.2%), (2) high risk (33.6%), (3) moderate lifestyle, low risk (51.0%), and (4) classic jocks (6.2%). (34).

In another study, Mathur et al. used LCA to identify separated classes of weight-related health behaviors of college students (40). The authors conducted the analysis separately for sophomore and senior students. This study identified 5 and 4 latent classes for sophomore and senior students. Four classes were similar among the 2 groups: (1) mostly healthy dietary habits, active; (2) moderately high screen time, inactive; (3) moderately healthy dietary habits, inactive; and (4) moderately high screen time, inactive. There were 2 additional classes for sophomore students: (1) moderately high screen time, inactive; (2) moderately healthy dietary habits, high screen time.

We identified 4 latent classes of dietary habits in a specific combination. As demonstrated in Table 4, most of the participants had healthy dietary habits and only 0.2% of the study population had unhealthy habits. In latent class 1 (healthy diet), the probability of fast food intake is 20%, but the probability of other dietary habits is very low. This rate of fast food intake is not low, however, we should consider over 50% probability in naming the classes.

On the other hand, in latent class 4 (unhealthy diet), the probability of most of dietary habits is high. There are 2 different patterns of dietary habits. Participants in latent class 2 (breakfast skippers) had a high probability of reporting skipping breakfast. Latent Class 3 (fruit, vegetables, and fish skippers) had a high probability of reporting fruit, vegetables, and fish skipping.

Our results indicated that alcohol use increases the odds of membership in classes 2 and 3 compared to class 1. A study among university students reported that after adjusting for sociodemographic characteristics, alcohol use was associated with uncommon breakfast consumption and

http://mjiri.iums.ac.ir
fruit/vegetable use, high fast food intake, unhealthy weight control behavior, and other dietary habits (41). Many studies from Iran and other countries have identified an association between cigarette smoking and alcohol consumption (36, 37, 42-44). A study showed that smoking and heavy alcohol use cluster with fruit/vegetable intake (22). Our finding is consistent with that of previous researches; however, in this study, we entered smoking and alcohol use as covariates into a model with 4 latent classes. Nevertheless, the results of the final analysis showed that smoking has no significant effect on the odds membership in latent classes 2, 3, and 4 compared to class 1.

This study had some strength, including large sample size and high response rate, which helped increase the generalization of the findings. The present study had the following limitations: (1) the design of the study was cross-sectional and causality could not be assessed (2). The study relied on self-report data and underreporting was expected about some dietary habits.

Despite using a well-designed methodology and sampling method, generalization of the results is limited only to the people of north of Iran.

**Conclusion**

This study showed the co-occurring nature of dietary habits by subgrouping a relatively large sample of Iranian people into 4 latent classes. Overall, 7.4% of all participants had some degrees of unhealthy habits. We named latent class 1 as "healthy diet", but the probability of eating fast food in this class was 20%. These findings stress the necessity of implementing a preventive intervention for the Iranian population. In addition, we found that being female and alcohol consumption are risk factors for membership in risky classes. Also, we found that having physical activity decreases the odds of membership in latent class 3 and 4 compared to class 1. Focusing on these variables and on probability of these variables can help design and execute effective preventive interventions in groups with unhealthy diet.

**Conflict of Interests**

The authors declare that they have no competing interests.

**References**


http://mjiri.iums.ac.ir

Med J Islam Repub Iran. 2018 (10 Aug); 32.69.
Pattern of unhealthy dietary habits


