Original Article

The association between dietary patterns before pregnancy and gestational diabetes mellitus: a matched case-control study in China

Xinxin Li MM[1†], Ting Kang MM[2†], Zhenwei Cui MD[3], Yacong Bo MD[1], Yanhua Liu MD[4], Amin Ullah MD[1], Xiangying Suo MD[5], HuaNan Chen MD[6], Quanjun Lyu MD[1,7]

[1]*Department of Nutrition and Food Hygiene, School of Public Health of Zhengzhou University, Zhengzhou, China*
[2]*Department of Infection and Prevention and Control, The Eighth People's Hospital of Zhengzhou, Zhengzhou, China*
[3]*Centre for Sport Nutrition and Health, School of Physical Education (main campus), Zhengzhou University, Zhengzhou, China*
[4]*Department of Nutrition, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China*
[5]*Department of Epidemiology and Health Statistics, School of Public Health of Zhengzhou University, Zhengzhou, China*
[6]*Department of Medical Record Management and Statistics, Shandong Provincial Qianfoshan Hospital, The First Affiliated Hospital of Shandong First Medical University, Jinan, China*
[7]*Department of Public Health, Zhengzhou Shuqing Medical College, Zhengzhou, China*
[†]*Both authors contributed equally to this manuscript*

Background and Objectives: We aimed to explore the relationship between dietary patterns and gestational diabetes mellitus (GDM) during pre-pregnancy six months using principal component analysis (PCA) and the geometric framework for nutrition (GFN). **Methods and Study Design:** We conducted a case-control study that included 210 GDM pregnant women and 210 controls. The dietary intake of all participants was assessed by a validated semi-quantitative food frequency questionnaire (FFQ). Major dietary patterns were extracted by PCA. A conditional logistic regression model was used to determine whether specific dietary patterns are associated with the risk of GDM. Meanwhile, the relationship between dietary patterns and GDM was visualized using GFN. **Results:** Four major dietary patterns were identified: "protein-rich pattern," "plant-based pattern," "oil-pickles-desserts pattern," and "cereals-nuts pattern." After adjustment for confounders, the "plant-based pattern" was associated with decreased risk of GDM (Q4 vs. Q1: OR = 0.01, 95% CI: 0.00~0.08), whereas no significant association was found in other dietary patterns. Moreover, there was no dietary intake of ice cream cones and deep-fried dough sticks for the population, which would produce fewer patients with GDM. Deep-fried dough sticks had statistically significant differences in the case and control groups ($p < 0.001$), while ice cream cones had the opposite result. **Conclusions:** The "plant-based pattern" may reduce the risk of GDM. Besides, although the "cereals-nuts pattern" had no association with GDM risk, avoiding the intake of deep-fried dough sticks could decrease GDM risk.

Key Words: dietary patterns, gestational diabetes mellitus, principal component analysis, geometric framework for nutrition, matched case-control study

INTRODUCTION

Gestational diabetes mellitus (GDM) is a prevalent condition characterized by the onset of hyperglycemia during pregnancy.[1] The prevalence of GDM has been increasing in China, making it a major public health concern.[2,3] GDM has been linked to various adverse health outcomes for both women and their offspring, including an increased risk of developing cardiometabolic disorders later in life.[4,5] A meta-analysis conducted in mainland China found that the total incidence of GDM in mainland China was 14.8% between 2010 and 2017.[6] It has already been shown that the diet before pregnancy is a modifiable factor that may influence the risk of developing GDM.[7-9]

Corresponding Author: Prof. Quanjun Lyu, Department of Nutrition and Food Hygiene, School of Public Health of Zhengzhou University, No.100, Science Avenue, Zhengzhou 450001, Henan, China
Tel: +86-37167781923
Email: lvquanjun666@163.com

Manuscript received 16 December 2023. Initial review completed 06 May 2024. Revision accepted 14 June 2024.
doi: 10.6133/apjcn.202409_33(3).0013

Although dietary factors are relevant modifiable risk factors for GDM, the exact association between dietary patterns and GDM remains controversial.[10] For example, a study has shown that the Western dietary pattern positively correlates with the risk of developing GDM.[11] In contrast, another study showed no significant association between adherence to the Western dietary pattern and the risk of GDM.[12] Principal component analysis (PCA) is a common method for analyzing dietary patterns.[13,14] It is a technique used to reduce many correlated variables into a smaller components,[14] revealing the underlying structure within population diets. However, the PCA has a limitation: there is no analysis of the relationship between food, macronutrients, and disease. Therefore, we adopted a novel approach of the geometric framework for nutrition (GFN) to analyze diet and disease, filling the PCA limitation.

GFN is used to examine how mixtures of nutrients (and other dietary components) influence biological outcomes like health and disease.[15] It utilizes a right-angled mixed triangle model to integrate nutrients, food, and dietary patterns, which simulates interactions in nutrients, meals, diets, dietary patterns, and disease.[16] GFN has been used by many studies, including the new theories of obesity based on the “protein leverage hypothesis”,[17] diet and aging,[18] cardiovascular metabolism,[19] liver,[20,21] and kidney diseases,[22] and new approaches to precision medicine.[23] It is evident from these studies that the nutritional geometry model demonstrates strengths in analyzing the relationship between diet and disease risk.

Moreover, insufficient evidence exists on the associations between dietary patterns before pregnancy and the risk of GDM. Therefore, the present study aimed to examine the association between maternal pre-pregnancy dietary patterns and the risk of developing GDM using the PCA method and GFN approach. Meanwhile, this study will contribute a perspective on the relationship between dietary habits and GDM in Chinese pregnant women, thereby establishing a new groundwork for protective strategies in GDM prevention based on dietary patterns.

METHODS

Study participants

This study was conducted by the Department of Obstetrics of the First Affiliated Hospital of Zhengzhou University from December 2020 to December 2021. Cases (n = 210) with GDM pregnant women aged 18~45 years and singleton pregnancy were included. The diagnostic criteria for GDM: a 75-g oral glucose tolerance test (OGTT) at 24~28 weeks of gestational age (wkGA), fasting plasma glucose ≥ 5.1 mmol/L and < 7 mmol/L, 1-h plasma glucose ≥ 10.0 mmol/L, or 2-h plasma glucose ≥ 8.5 mmol/L and < 11.1 mmol/L, blood glucose reaches any one of the above points was used to diagnose GDM.[24]

Controls were pregnant women (n = 210) whose OGTT at 24~28 weeks of pregnancy was in the normal range. The exclusion criteria were as follows: (1) women with a history of GDM and a history of pre-gestational diabetes; (2) hypertensive disorders during pregnancy, including gestational hypertension and pre-eclampsia; (3) artificial impregnation; (4) endocrine diseases (hyperthyroidism or hypothyroidism) or other diseases, such as severe heart, kidney, neurological complications (5) genetic family history, (6) women who had cognitive disorders such as mental illnesses, (7) refused to participate, or an incomplete questionnaire. Moreover, participants were excluded if they had improbable energy intakes of < 500 kcal/d and > 3500 kcal/d (n=1).[25] Based on a 1:1 case to control ratio study formula for individual matching according to age (±1) and gestational week (±2). The present study reported the data of 420 pregnant women. The adequacy of the sample size of this study is that hypothesized compliance with better dietary patterns by approximately 30% of control participants could reduce the incidence of GDM by 50%,[26] and we also took into account a 10% non-response rate and a 90% qualified rate, with 80% statistical power and 5% significance. The study flowchart is shown in Figure 1.

This study was approved by the Ethics Committee of Zhengzhou University Life Science (Ethics Number: ZZUIRB2020-32), and all the participants signed an informed consent form.

Dietary intake

The dietary intake of participants was recorded by a validated semiquantitative food frequency questionnaire (FFQ) six months before pregnancy during hospitalization through a face-to-face interview.[27] The FFQ contained seventy-nine food items, including seventy-one food items, five beverage items, and three other items (soup, water, and edible oil). For each food item, the participants were asked to report their consumption frequency and serve of a given serving of each food item. The FFQ has been previously shown to be valid and reproducible for the usual intake of nutrients and major foods by women in urban Shanghai and the usual consumption of major nutrients and food groups among Chinese women in Guangdong.[27]

The daily energy, nutrients, and food consumption for participants was calculated according to the Chinese Food Composition Tables 2009.[28] The data obtained this way was transformed into daily dietary doses and expressed as g/day or mL/day. To calculate micronutrient density, the formula: nutrient density = (amount of a nutrient in a given amount of a food/energy contained in the same amount of that food) * 1000.

Assessment of nondietary exposures

Basic characteristics of pregnant women who participated in the study six months before pregnancy were collected. This information included height, weight, body mass index (BMI), physical activity, family history of diabetes, pre-pregnancy BMI, education level, income, and daily energy intake. Weight and height were measured using a weight machine by a trained investigator, with minimum clothes and without shoes, and were reported to the nearest 0.1 kg. Besides, a physical activity questionnaire was used to report common activities during the six months before pregnancy, including type and cumulative duration

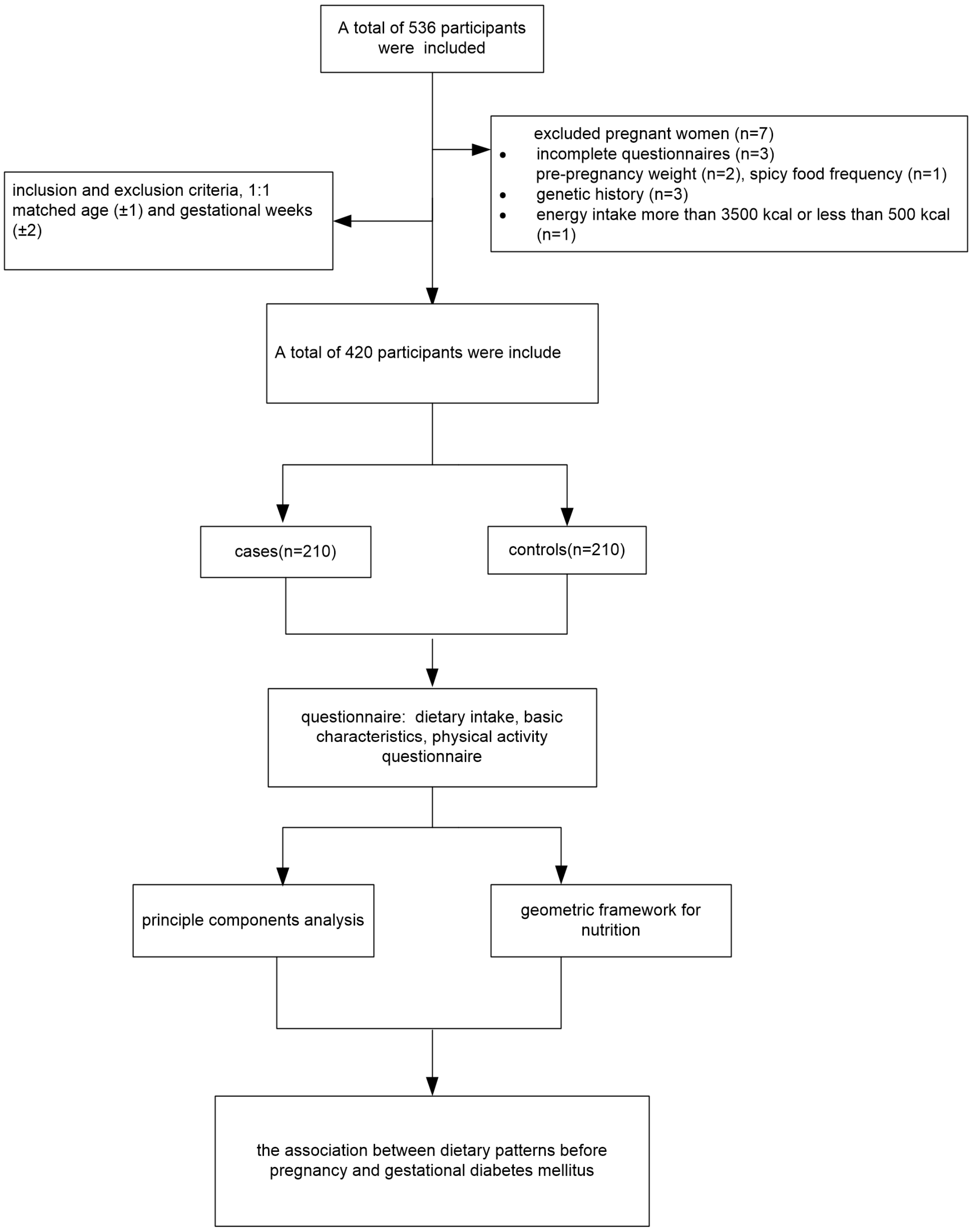


Figure 1. Study flowchart

of activity per week or day. The metabolic equivalent (MET) of tasks of each activity was multiplied by the frequency and duration of physical activity to calculate the daily physical activity level in MET hours per day (MET-h/d).

Statistical analysis

All analyses were performed using SPSS software version 25 (IBM Corporation, USA), and a two-sided p value < 0.05 was considered significant. The normal distribution of the data was evaluated using a histogram and the Shapiro-Wilk test, and analysis of variance was performed using the F-test. A paired samples t-test and a Wilcoxon matched-pairs signed rank test were applied. Normally distributed continuous variables were presented as means and standard deviations. If the data did not conform to a normal distribution, the data were expressed as median and quartiles. Categorical variables were compared using the McNemar or McNemar-Bowker test for paired observations and presented as counts and percentages.

Methods analysis

Dietary patterns were identified through the applied varimax rotation to the component matrix to maximize the variance explained on each component. Dietary patterns were identified from the reported intakes of individual foods using PCA, and the factor score for each dietary pattern was calculated by summing the food intakes of that group in terms of their factor loading, and each participant received a score for each pattern in terms of factor scores. The participants were divided into quartiles based on the dietary pattern scores of the controls. Conditional logistic regression models were used to compute the odds ratio (OR) and 95% confidence interval (CI) between dietary pattern scores and the risk of GDM. Linear trends were tested by entering the within-group median transformation as continuous parameters in the models. Raw scoring data were transformed into a z score to assess participants for the risk of GDM according to each standard deviation (SD) increase. p values for trend across the quartile of dietary patterns adherence were determined. Two models were created: crude model and adjusted model. Covariates were determined according to the previous studies, and categorical variables included in the model as follows: education level (junior high school and below, senior high school and undergraduate, master and above), family history of diabetes (yes/no) and income (≤3000 RMB, 3001~3999 RMB, 4000~5999 RMB, ≥6000 RMB, unknown), while age, gestational week, pre-pregnancy BMI, physical activity and energy intake were included in the model as continuous variables. Additionally, covariates were adjusted into the conditional logistic regression model when there was a significant difference between the case and control group.

The right-angled mixture triangle depicts the relationship between dietary patterns and GDM using R version 4.2.3.[15] By choosing to model macronutrients, we were able simplify the complex makeup of foods and diets, which consist of numerous dimensions.[16] Dietary patterns extracted from PCA were visualized by GFN. The detailed operation is as follows. The food groups were presented in each dietary pattern by calculating the percentage of total energy provided by protein, fats, and carbohydrates from the food items. Second, a polygon function was used to plot dietary patterns according to the food groups.

Quality control

First, a pre-survey was conducted before the formal survey to ensure that the questionnaire was easily understood and accepted within the questionnaire. Second, the questioning and survey methods were standardized for case and control groups to help respondents recall information in as much detail as possible and reduce survey bias. Third, the investigators were all master's degree students from the School of Public Health and underwent rigorous training to ensure that participants understood the questionnaire content in detail and the potential benefits, making it easy to obtain informed consent and reducing reporting bias. Fourth, the questioning was used one-to-one, face-to-face, with the investigator asking and completing the questionnaire to ensure the accuracy of information collection. Fifth, the pre-prepared food atlas was used for better estimation of the food intake consumed each time by the participants, such as bowls, cups, spoons, slices, and palms of hand. Finally, the content of the questionnaire was verified at the end of the survey to ensure the completeness of the data.

RESULTS

The demographics of study participants (Table 1)

There were statistically significant differences in pre-pregnancy BMI, physical activity, family history of diabetes, education, and income in the case group and control group. At the same time, other variables did not differ significantly. The case group, in terms of pre-pregnancy BMI and energy intake were higher compared to the control group, which had lower gestational week and physical activity.

The nutrient concentrations and nutrient density of study participants (Table 2)

All variables, except for carbohydrate, total energy, non-protein, non-carbohydrate, and vitamin E, other variables were found to have a statistically significant relationship between the case group and control group ($p < 0.05$). In the case groups, fat (kcal/day), total energy, non-protein, fat (%E), sodium, copper, and vitamin B-12 consistently were higher than the control group. However, other vari-

Table 1. Demographic data of the studied population

Variables[†]	Cases (n = 210)	Controls (n = 210)	p-value
Age (year)	31.4 ± 3.99	31.4 ± 4.05	0.359
Gestational week (week)	36.0 ± 3.77	36.1 ± 3.62	0.320
Pre-pregnancy BMI (kg/m^2)	23.4 ± 3.43	22.2 ± 3.20	<0.001
Physical activity (MET-hour/day)	30.4 ± 3.96	32.3 ± 5.41	<0.001
Energy intake (kcal/day)	1835 ± 430	1817 ± 350	0.651
Family history of diabetes			<0.001
No	173 (82.4)	200 (95.2)	
Yes	37 (17.6)	10 (4.8)	
Education (%)			0.023
Junior high school and below	35 (16.7)	36 (17.1)	
Senior high school and undergraduate	134 (63.8)	155 (73.8)	
Master and above	41 (19.5)	19 (9.0)	
Income (%)			0.026
≤3000 RMB	84 (40.0)	119 (56.7)	
3001~3999 RMB	47 (22.4)	36 (17.1)	
4000~5999 RMB	34 (16.2)	20 (9.5)	
≥6000 RMB	35 (16.7)	24 (11.4)	
unknown	10 (4.8)	11 (5.2)	

pre-pregnancy BMI: pre-pregnancy body mass index; MET: metabolic equivalent; RMB: renminbi (Chinese yuan)

[†] Paired chi-squared test, categorical variables were compared using the McNemar or McNemar-Bowker test, and continuous variables were compared using paired samples t-test. p-value < 0.05 was considered significant.

Table 2. The nutrient concentrations and nutrient density of study participants

Nutrients[†]	Cases (n = 210)	Controls (n = 210)	*p-value*
Macronutrients			
Protein (kcal/day)	193 ± 71.6	223 ± 60.5	<0.001
Fat (kcal/day)	901 ± 205	832 ± 160	<0.001
Carbohydrate (kcal/day)	756 ± 211	783 ± 207	0.192
Total energy (kcal/day)	1851 ± 435	1837 ± 355	0.731
Non-protein (kcal/day)	289 ± 68.5	288 ± 60.9	0.861
Non-fat (kcal/day)	237 ± 67.7	251 ± 63.0	0.028
Non-carbohydrate (kcal/day)	138 (119, 169)	140 (127, 166)	0.426
Protein (%E)	9.86 (8.85, 11.5)	11.9 (10.7, 13.0)	<0.001
Fat (%E)	49.1 ± 5.01	45.6 ± 5.32	<0.001
Carbohydrate (%E)	40.7 ± 4.77	42.4 ± 5.17	0.001
Fat-soluble vitamins			
Vitamin A (μgRE/1000 kcal)	227 (174, 295)	338 (257, 429)	<0.001
Retinol (μg/1000 kcal)	59.8 (46.4, 83.3)	83.9 (59.7, 108)	<0.001
Vitamin E (mg/1000 kcal)	21.7 ± 5.75	21.8 ± 5.16	0.805
Vitamin D (IU/1000 kcal)	29.1 (21.7, 45.4)	36.5 (26.4, 57.3)	0.001
Water-soluble vitamins			
Thiamin (mg/1000 kcal)	0.27 (0.23, 0.32)	0.34 (0.30, 0.39)	<0.001
Riboflavin (mg/1000 kcal)	0.42 (0.36, 0.49)	0.54 (0.48, 0.60)	<0.001
Niacin (mg/1000 kcal)	6.53 (5.83, 7.43)	7.52 (6.82, 8.37)	<0.001
Vitamin C (mg/1000 kcal)	45.9 (34.9, 56.5)	55.7 (46.0, 71.7)	<0.001
Vitamin B-6 (mg/1000 kcal)	0.35 (0.30, 0.39)	0.40 (0.35, 0.45)	<0.001
Vitamin B-12 (μg/1000 kcal)	0.45 (0.25, 0.71)	0.31 (0.18, 0.53)	<0.001
Folate (μg/1000 kcal)	91.5 (79.5, 108)	112 (97.2, 126)	<0.001
Major minerals			
Calcium (mg/1000 kcal)	183 (151, 230)	253 (211, 299)	<0.001
Phosphorus (mg/1000 kcal)	384 (343, 445)	481 (437, 541)	<0.001
Potassium (mg/1000 kcal)	710 (622, 824)	913 (833, 1032)	<0.001
Sodium (mg/1000 kcal)	366 (328, 436)	350 (308, 396)	<0.001
Magnesium (mg/1000 kcal)	130 (116, 146)	161 (146, 174)	<0.001
Trace minerals			
Iron (mg/1000 kcal)	8.12 (7.65, 8.77)	9.53 (8.86, 10.3)	<0.001
Zinc (mg/1000 kcal)	3.98 (3.62, 4.47)	4.71 (4.34, 5.14)	<0.001
Selenium (mg/1000 kcal)	16.3 (14.3, 18.6)	19.2 (16.7, 21.5)	<0.001
Copper (mg/1000 kcal)	2.00 (0.99, 3.15)	1.59 (0.88, 2.70)	0.038
Manganese (mg/1000 kcal)	2.36 (2.18, 2.64)	2.72 (2.46, 2.96)	<0.001
Other nutrients			
Soluble fiber (g/1000 kcal)	4.15 ± 0.96	5.43 ± 1.16	<0.001
Methionine (mg/1000 kcal)	522 (466, 577)	591 (538, 647)	<0.001
Total choline (mg/1000 kcal)	111 ± 29.4	139 ± 33.1	<0.001

[†] Continuous variables were compared using paired samples t-test and Wilcoxon matched-pairs signed rank test. *p*-value < 0.05 was considered significant

ables were the opposite.

Dietary patterns construction based on PCA (Table 3)

We grouped food items into 16 food groups according to the similarity of foods (Supplementary Table 1). All variables, except cereals and red meat, reached statistical significance ($p < 0.05$). In the case group, the consumption of cereals, potatoes, beverages, desserts, pickles, edible oil, and red meat was consistently higher than in the control groups. In contrast, the intake of soybean and soybean products, fruits, aquatic products, eggs, nuts, edible fungi, whole grains and legumes, vegetables, dairy and dairy products were found to be lower than in the control group ($p < 0.05$) (Supplementary Table 2).

The results of the Kaiser-Meyer-Olkin (KMO) test is 0.717, and Bartlett's test of sphericity was significant ($p < 0.001$), indicating that factor analysis was appropriate for these data. The dietary patterns, whether included or not, used a factor loading threshold of 0.50. Four major dietary patterns were extracted based on the eigenvalue criteria (≥ 1), the scree plot, and the maximum variance method (Supplementary Figure 1), accounting for 45.0% of the total variance in the data. These dietary patterns were: "protein-rich pattern," "plant-based pattern," "oil-pickles-desserts pattern," and "cereals-nuts pattern." Each dietary pattern was named according to the composition of the predominant food groups. The "protein-rich pattern" consisted of 19.2 % of the variance and contained soybean and soybean products, aquatic products, dairy and dairy products, as well as red meat. The "plant-based pattern" consisted of 10.5 % of the variance and mainly contained high factor loadings for whole grains and legumes, vegetables, and fruits. The "oil-pickles-desserts pattern" consisted of 8.40 % of the variance and included desserts, pickles, and edible oil. The "cereals-nuts pattern" consisted of 6.91 % of variance and mainly included cereals and nuts.

Table 3. Dietary patterns derived from the maximum variance method and their rotated component matrix

Food groups[†]	Protein-rich pattern	Plant-based pattern	Oil-pickles-desserts pattern	Cereals-nuts pattern
Soybean and soybean products	0.640	—	—	—
Aquatic products	0.724	—	—	—
Dairy and dairy products	0.551	—	—	—
Red meat	0.545	—	—	—
Whole grains and legumes	—	0.672	—	—
Vegetables	—	0.787	—	—
Fruits	—	0.519	—	—
Desserts	—	—	0.600	—
Pickles	—	—	0.629	—
Edible oil	—	—	0.662	—
Cereals	—	—	—	0.690
Nuts	—	—	—	0.522
Total variance	19.2%	10.5%	8.40%	6.91%

[†] Only food groups with absolute factor loadings ≥ 0.50 were retained in each pattern for simplicity.

Analysis of dietary pattern and GDM risk using conditional logistic regression (Table 4)

Q1 represents the lowest quartile, while Q4 represents the highest quartile. The Q4 group in the "protein-rich pattern" reduced the risk of GDM by 50% compared to the Q1 group, but it was not statistically significant. The trend of the dietary pattern becomes closer as the dietary pattern score increases. The linear trend was found to be statistically significant ($p = 0.017$). Furthermore, each SD increase in the "protein-rich pattern" score was associated with a 28% reduction in the risk of GDM in the adjusted model, but it was not statistically significant. For the "plant-based pattern," a higher dietary pattern score was associated with a decreased risk of GDM, and it was statistically significant. Compared to the Q1 group, the Q4 group showed an almost 100% reduction in the risk of GDM in the adjusted model. Moreover, the linear trend was statistically significant ($p < 0.001$) in the adjusted model. Additionally, each SD increase in the "plant-based pattern" score was associated with an 83% reduced risk of GDM, and it was statistically significant. In the "oil-pickles-desserts pattern," compared to the Q1 group, the Q4 group showed a 7.39-fold relative increase in the risk of GDM, and it was statistically significant in the crude model. The linear trend was statistically significant in the crude model ($p < 0.001$), whereas this linear trend relationship disappeared in the adjusted model ($p = 0.082$). In addition, each SD increase in the "oil-pickles-desserts pattern" score was associated with a 1.96-fold risk of GDM in the adjusted model, and it was statistically significant. We did not find any association between the "cereals-nuts pattern" and GDM risk.

Visualized the association between dietary patterns and GDM using GFN (Figure 2)

According to the proportion of total energy provided by protein, we used 10% as a cutoff to identify the case and control group population. The proportion less than 10% indicated the majority of the case group population, while the majority of the control group population was identified as the proportion greater than or equal to 10%. Most participants in the control group can be seen clustered within the polygon (Figure 2a), representing a "protein-rich pattern." The food groups for the "protein-rich pattern" were represented by hollow triangles. In this diet, the closer the triangle was to the origin of the coordinate axis, the higher the proportion of total energy provided by carbohydrates. Interestingly, no data points from either group fell within the polygon (Figure 2b), which represented a "plant-based pattern." Hollow diamonds represented vegetables, fruits, whole grains and legumes. They were distributed in the lower-left area of the picture. In addition, according to the proportion of total energy provided by protein, we used 10% as a cutoff to identify food species from a "plant-based pattern." The proportion less than 10% indicated fruits, while vegetables, whole grains and legumes were identified as the proportion greater than or equal to 10%. Additionally, we observed that some individuals in the case group are distributed within the polygon for the "oil-pickles-desserts pattern" (Figure 2c). Unfortunately, we did not observe any consistent distribution pattern in the "oil-pickles-desserts pattern." Moreover, we observed that when there was no dietary intake of ice cream cones for the population, it would produce fewer patients with GDM. However, no significant differences were noticed for ice cream cones in the case and control groups ($p = 0.302$) (Supplementary Table 2). The distribution of individuals in both the case and control groups within the region was similar, representing a "cereals-nuts pattern" (Figure 2d). The "cereals-nuts pattern" was represented by hollow squares and was predominantly distributed in the 10% proportion of total energy provided by protein, with a wide range of the total energy provided by fat. We observed that when there was no dietary intake of deep-fried dough sticks for the population, it would produce fewer patients with GDM. Deep-fried dough sticks had statistically significant differences in the case group and control group ($p < 0.001$) (Supplementary Table 2).

DISCUSSION

The present study investigated the relationship between dietary patterns and GDM using two analytical methods. The PCA method was employed, which identified four common factors named "protein-rich pattern," "plant-based pattern," "oil-pickles-desserts pattern," and "cereals-nuts pattern" based on the major food groups. After adjusting for confounding factors, the results from conditional logistic regression showed that each SD increase in the "plant-based pattern" was associated with an 83%

Table 4. The relationship between dietary pattern score quartiles and the risk of GDM based on conditional logistic regression analysis

Dietary patterns	Dietary pattern score quartiles (OR, 95%CI)			
	Q1	Q2	Q3	Q4
Protein-rich pattern				
N (cases/controls)	121/52	26/53	24/53	39/52
Crude model	1	0.17 (0.06, 0.47)**	0.28 (0.10, 0.78)*	0.78 (0.28, 2.19)
Adjusted model†	1	0.16 (0.05, 0.52)**	0.31 (0.09, 1.14)	0.50 (0.12, 2.06)
Plant-based pattern				
N (cases/controls)	138/52	39/53	19/53	14/52
Crude model	1	0.21 (0.09, 0.51)**	0.10 (0.03, 0.30)***	0.06 (0.02, 0.19)***
Adjusted model†	1	0.10 (0.03, 0.34)***	0.04 (0.01, 0.19)***	0.01 (0.00, 0.08)***
Oil-pickles-desserts pattern				
N (cases/controls)	32/52	20/53	69/53	89/52
Crude model	1	0.78 (0.30, 2.06)	2.16 (0.94, 4.97)	8.39 (3.04, 23.16)***
Adjusted model†	1	0.49 (0.14, 1.74)	1.05 (0.34, 3.22)	3.02 (0.64, 14.36)
Cereals-nuts pattern				
N (cases/controls)	72/52	49/53	44/53	45/52
Crude model	1	1.00 (0.41, 2.45)	0.66 (0.25, 1.75)	0.89 (0.33, 2.45)
Adjusted model†	1	1.48 (0.50, 4.39)	0.91 (0.28, 2.92)	0.86 (0.20, 3.79)

Dietary patterns	*p*-trend	Per 1 SD
Protein-rich pattern		
N (cases/controls)		
Crude model	0.001	0.74 (0.50, 1.11)
Adjusted model†	0.017	0.72 (0.41, 1.25)
Plant-based pattern		
N (cases/controls)		
Crude model	<0.001	0.24 (0.15, 0.39)***
Adjusted model†	<0.001	0.17 (0.09, 0.32)***
Oil-pickles-desserts pattern		
N (cases/controls)		
Crude model	<0.001	3.81 (2.37, 6.11)***
Adjusted model†	0.082	2.96 (1.41, 6.23)**
Cereals-nuts pattern		
N (cases/controls)		
Crude model	0.836	1.04 (0.71, 1.54)
Adjusted model†	0.801	1.06 (0.62, 1.81)

†Adjusted model: adjusted for age, gestational week, pre-pregnancy BMI, family history of diabetes, education, income, physical activity, energy intake.

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

reduction. Conversely, each SD increase in the "oil-pickles-desserts pattern" showed a 1.96-fold increase in the risk of GDM. Moreover, the study utilized the GFN to visualize the association between dietary patterns and GDM. The results showed that the proportion less than 10% indicated the majority of the case group population, whereas the majority of the control group population was identified for the proportion greater than or equal to 10%. Besides, there was no dietary intake of ice cream cones for the population, which would produce fewer patients with GDM for the "oil-pickles-desserts pattern." However, the ice cream cones had no statistically significant differences between the case and control groups (p = 0.302). There was no dietary intake of deep-fried dough sticks for the population, it would produce fewer patients with GDM for the "cereals-nuts pattern." Deep-fried dough sticks had statistically significant differences between the case and control groups ($p < 0.001$).

We did not observe a significant association between the risk of GDM and "protein-rich pattern." The "protein-rich pattern" in this study was characterized by a high consumption of soybean and soybean products, aquatic products, dairy and dairy products, and red meat. Previous research has shown that dietary protein intake contributes to the development of GDM, particularly red and processed meat.[29-31] A study found that a higher intake of protein was associated with a higher risk of GDM in Asian women.[32] However, our findings showed that most of the population included in the "protein-rich pattern" was the control group, and the majority of the population in the control group was distributed in the range of 10% or more protein percentage in a state of balanced protein intake. Conversely, in the case group, the majority of the population was distributed in the range of less than 10%, which was a state of protein deficiency, compared to the nationally recommended lowest protein level (10%). When the protein was in a deficiency state, food intake and food preferences showed adaptive changes that suggested that compensatory mechanisms were induced to restore adequate protein status,[33] which meant the intake of more foods and, consequently, more non-protein energy (fat and carbohydrate) might lead to poor blood sugar control. Moreover, a study showed that a minor decrease in the percentage of protein in the diet would lead to a significant increase in the percentage of fat and carbohydrate intake.[17] Another study further demonstrated that

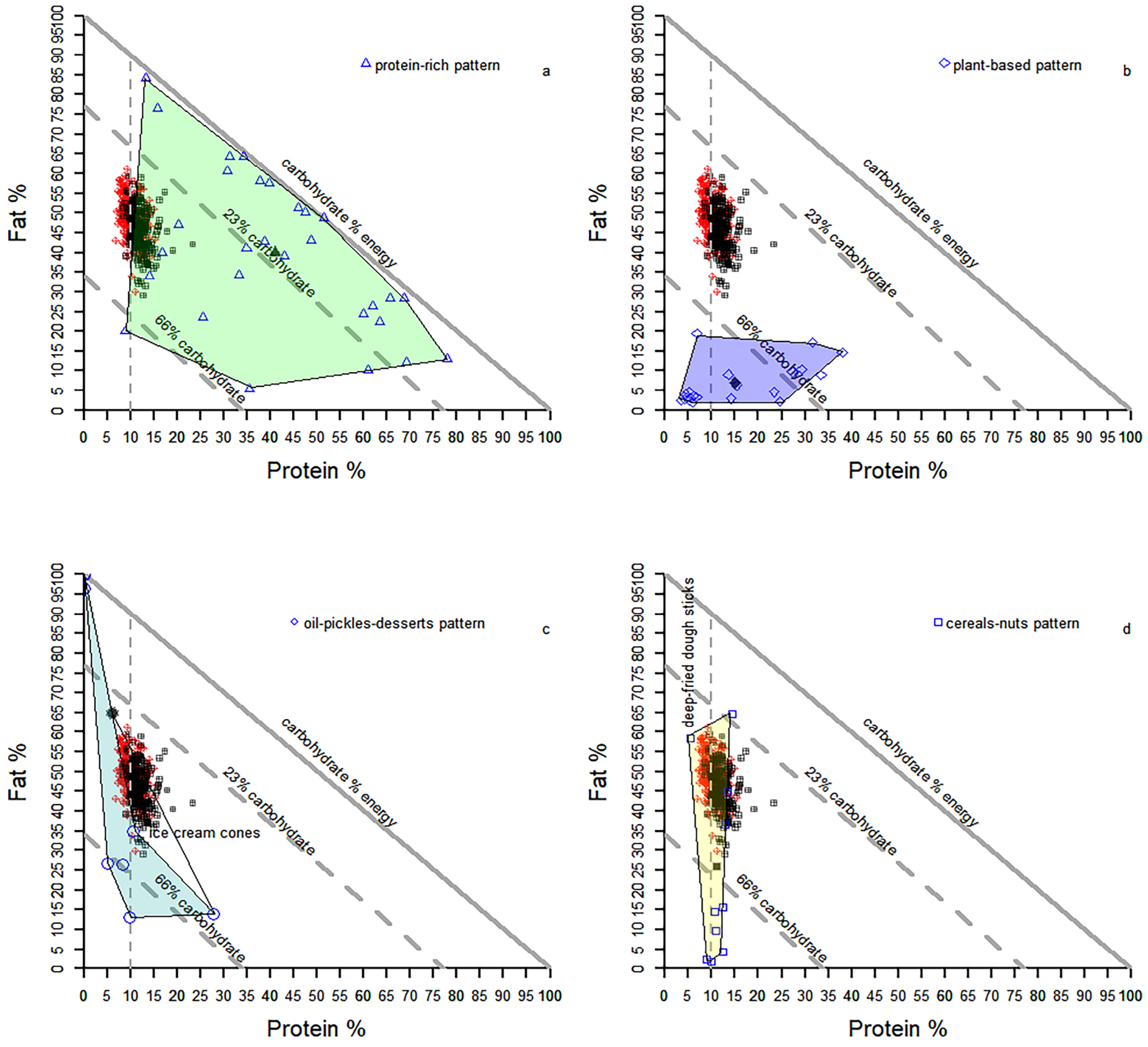


Figure 2. The four dietary patterns of the PCA were related to GDM. The combination of red circles and plus represents the case group, and the combination of black boxes and plus represents the control group. Other graphical elements with different shapes represent different food types from four dietary patterns. The black solid elements of the picture are mean values in different diets. The four shapes correspond to the four different patterns. The polygons of the figure (a, b, c, and d) represent the “protein-rich pattern,” “plant-based pattern,” “oil-pickles-desserts pattern,” and “cereals-nuts pattern,” respectively. Each polygon is composed of different shapes.

when the percent protein of total energy in the diet increased, dietary protein intake remained relatively more stable, whereas the non-protein had the opposite trend.[34] In other words, if we can ensure that the protein intake in the diet is adequate and balanced, avoiding excessive intake of non-protein in the diet could improve the abnormality of blood glucose caused by diet. The “protein-rich pattern” appears to have a preventive effect on the development of GDM risk from this point of view. However, conditional logistic regression analysis findings demonstrated that the “protein-rich pattern” was not associated with a risk of GDM. We supposed that the branched-chain amino acids from proteins of animal origin and certain amino acids rich in proteins of plant origin, such as arginine, had diverse effects mechanisms, leading to potentially affecting diabetes-related metabolic pathways or potentially promoting the body’s insulin metabolism, which might have produced the reason for a part of the results differences.[35-39]

In a “plant-based pattern,” the consumption of whole grains and legumes, vegetables, and fruits which are rich in carbohydrates. The proportion of total energy provided by protein is less than 10%, which indicates fruits, while vegetables, whole grains, and legumes were identified as the proportion greater than or equal to 10%. It did not significantly increase serum glucose levels due to the inclusion of low glycemic index foods. Foods with a low glycemic index slowly release glucose and do not cause a significant increase in blood glucose levels.[40] In addition, fruits and vegetables provide fiber, micronutrients, and antioxidants, further highlighting the benefits of including them in the diet.[7] Therefore, the dietary pattern is regarded as a protective factor for GDM. The closer the dietary pattern that includes vegetables and fruits is considered a protective factor against GDM, which supports previous findings. Another study also showed that the “fruits and dairy products” dietary pattern may decrease GDM risk.[26] This may be because the micronutrients in leafy green vegetables, fruit, and milk may have a significant protective effect against the development of GDM.[41] These results were consistent with our findings.

We also did not observe a significant association between the risk of GDM and the "oil-pickles-desserts pattern," and the "cereals-nuts pattern." Although women who consumed diets high in sugar and fats are at a higher risk of developing GDM,[10,42] whereas the "oil-pickles-desserts pattern" showed no association for GDM. This dietary pattern was similar to the Western dietary pattern, which comprised a high consumption of sugar-sweetened beverages, salty snacks, biscuits, and saturated oils.[12] However, a prospective study of pregnant Malaysian women found that a significant association between a diet which is high in sugar, spreads and creamers, spices, and condiments and GDM risk was found only among obese women and significantly associated with a reduced risk of GDM among women with high adherence.[43] For the "cereals-nuts pattern," our findings showed that the proportion of the case group in this diet was similar to the control group population in the "cereals-nuts pattern." The results were consistent with the findings of conditional logistic regression analysis, which showed no association with GDM risk. However, a study revealed that a higher intake of vegetable protein, specifically nuts, was associated with a significantly lower GDM risk.[31] This may be attributed to the dietary fibers in nuts, which can slow gastric depletion and glucose absorption, thus impacting the risk of developing gestational diabetes. Moreover, there was no dietary intake of deep-fried dough sticks for the population, which would produce fewer patients with GDM for the "cereals-nuts pattern," and deep-fried dough sticks had statistically significant differences in the case group and control group. A large prospective cohort study showed a significant and positive association between pre-pregnancy fried food consumption and the risk of GDM.[44] Another study also indicated that higher pre-pregnancy consumption of fried food was associated with an increased risk of GDM.[45] In addition, we found that the "cereals-nuts pattern" maintained a stable protein content. Most foods had a protein content of approximately 10%, while the intake of carbohydrates and fats varied significantly. This pattern aligns with the protein leverage hypothesis, which suggests that in nutritionally imbalanced diets, individuals prioritize protein consumption over carbohydrates and fats.[17] The lack of association between the "cereals-nuts pattern" and GDM may be explained by the complementary relationship between high-carbohydrate and low-fat diets, as well as high-fat foods, which help regulate excessive intake of both fats and carbohydrates.

The current study has several significant strengths. First, we introduced a new model-based GFN to visualize the association between dietary patterns and disease. This framework allows us to examine the relationship between individual food intake and overall dietary patterns and disease outcomes. Second, compared with the PCA method of analyzing dietary patterns, GFN better explained the relationship between dietary patterns and disease by presenting a more objective and concise picture, avoiding the variability of results due to subjectivity during the methodological analysis. Third, the PCA is combined with the GFN to analyze dietary patterns. Dietary patterns in PCA were visualized by GFN to analyze the association among nutrients, food, and dietary patterns, which better explains the relationship between diet and disease.

However, there are also limitations to mention. First, due to the observational nature of our study, we cannot establish a causal relationship between pre-pregnancy dietary patterns and the risk for GDM.[42] Second, the current study used self-reported FFQ data to measure dietary intake. In addition, the time range is too broad (first six months of pregnancy), which may bias the results. Whereas dietary data is strengthened by the similarities between the daily mean energy intake reported in our study (case group: 1850 kcal/day; control group: 1837 kcal/day) and that reported in a representative sample of pregnant women among Chinese women (light physical activity level in women: 1800 kcal/day) in 2017.[46] The distributions of macronutrients were also similar to the nationwide population data before pregnancy.[46] Third, there may be several other confounding factors that have an impact on the results, such as changes in weight during pregnancy. Therefore, we will emphasize the effect of this indicator on the results in future studies.

Conclusion

Our research findings suggest that the "plant-based pattern" before pregnancy can reduce the probability of developing GDM. Moreover, since most cases are in a protein-deficient state, the "protein-rich pattern" may restore a protein-sufficient state. The finding suggests the importance of ensuring adequate protein intake in our diets. Although we found no significant association between "cereals-nuts pattern" and GDM risk, when there is no dietary intake of deep-fried dough sticks for the population, it could decrease the risk of GDM. We aim to further explore the relationship between preconception diet and GDM based on the GFN using a larger sample size in future studies.

ACKNOWLEDGEMENTS

This study would not have been possible without the support of the participants and the assistance of the Department of Obstetrics of the First Affiliated Hospital of Zhengzhou University. Moreover, we sincerely thank the reviewers and editors for their valuable comments.

CONFLICT OF INTEREST AND FUNDING DISCLOSURES

The authors declare no conflict of interest.

This study was funded and supported by the Chinese Nutrition Society-Feihe Physical Nutrition and Health Research Fund project (Project No. CNS-Feihe2019A34).

REFERENCES

1. Mahajan A, Donovan LE, Vallee R, Yamamoto JM. Evidenced-based nutrition for gestational diabetes mellitus. Curr Diab Rep. 2019; 19: 94. doi: 10.1007/s11892-019-1208-4.
2. Zhu Y, Zhang C. Prevalence of gestational diabetes and risk of progression to type 2 diabetes: a global perspective. Curr Diab Rep. 2016; 16: 7. doi: 10.1007/s11892-015-0699-x.
3. Juan J, Yang H. Prevalence, prevention, and lifestyle intervention of gestational diabetes mellitus in China. Int J Environ Res Public Health. 2020; 17: 9517. doi: 10.3390/ijerph17249517.

4. Ye W, Luo C, Huang J, Li C, Liu Z, Liu F. Gestational diabetes mellitus and adverse pregnancy outcomes: systematic review and meta-analysis. BMJ. 2022; 377: e067946. doi: 10.1136/bmj-2021-067946.
5. Leybovitz-Haleluya N, Wainstock T, Landau D, Sheiner E. Maternal gestational diabetes mellitus and the risk of subsequent pediatric cardiovascular diseases of the offspring: a population-based cohort study with up to 18 years of follow up. Acta Diabetol. 2018; 55: 1037-42. doi: 10.1007/s00592-018-1176-1.
6. Gao C, Sun X, Lu L, Liu F, Yuan J. Prevalence of gestational diabetes mellitus in mainland China: a systematic review and meta-analysis. J Diabetes Investig. 2019; 10: 154-62. doi: 10.1111/jdi.12854.
7. Donazar-Ezcurra M, López-Del Burgo C, Bes-Rastrollo M. Primary prevention of gestational diabetes mellitus through nutritional factors: a systematic review. BMC Pregnancy Childbirth. 2017; 17: 30. doi: 10.1186/s12884-016-1205-4.
8. Raghavan R, Dreibelbis C, Kingshipp BL, Wong YP, Abrams B, Gernand AD et al. Dietary patterns before and during pregnancy and maternal outcomes: a systematic review. Am J Clin Nutr. 2019; 109: 705s-28s. doi: 10.1093/ajcn/nqy216.
9. Jarman M, Mathe N, Ramazani F, Pakseresht M, Robson PJ, Johnson ST, Bell RC, APrON and ENRICH study teams. Dietary patterns prior to pregnancy and associations with pregnancy complications. Nutrients. 2018; 10: 914. doi: 10.3390/nu10070914.
10. Schoenaker DAJM, Mishra GD, Callaway LK, Soedamah-Muthu SS. The role of energy, nutrients, foods, and dietary patterns in the development of gestational diabetes mellitus: a systematic review of observational studies. Diabetes Care. 2016; 39: 16-23. doi: 10.2337/dc15-0540.
11. Quan W, Zeng M, Jiao Y, Li Y, Xue C, Liu G, Wang Z, Qin F, He Z, Chen J. Western dietary patterns, foods, and risk of gestational diabetes mellitus: a systematic review and meta-analysis of prospective cohort studies. Adv Nutr. 2021; 12: 1353-64. doi: 10.1093/advances/nmaa184.
12. Asadi M, Shahzeidi M, Nadjarzadeh A, Hashemi Yusefabad H, Mansoori A. The relationship between pre-pregnancy dietary patterns adherence and risk of gestational diabetes mellitus in Iran: a case-control study. Nutr Diet. 2019; 76: 597-603. doi: 10.1111/1747-0080.12514.
13. Ocké MC. Evaluation of methodologies for assessing the overall diet: dietary quality scores and dietary pattern analysis. Proc Nutr Soc. 2013; 72: 191-9. doi: 10.1017/s0029665113000013.
14. Santos RdO, Gorgulho BM, Castro MAd, Fisberg RM, Marchioni DM, Baltar VT. Principal component analysis and factor analysis: differences and similarities in nutritional epidemiology application. Rev Bras Epidemiol. 2019; 22: e190041. doi: 10.1590/1980-549720190041.
15. Raubenheimer D. Toward a quantitative nutritional ecology: the right-angled mixture triangle. Ecol Monogr. 2011; 81: 407–27. doi: 10.1890/10-1707.1.
16. Raubenheimer D, Simpson SJ. Nutritional ecology and human health. Annu Rev Nutr. 2016; 36: 603-26. doi: 10.1146/annurev-nutr-071715-051118.
17. Simpson SJ, Raubenheimer D. Obesity: the protein leverage hypothesis. Obes Rev. 2005; 6: 133-42. doi: 10.1111/j.1467-789X.2005.00178.x.
18. Lee KP, Simpson SJ, Clissold FJ, Brooks R, Ballard JW, Taylor PW, Soran N, Raubenheimer D. Lifespan and reproduction in Drosophila: new insights from nutritional geometry. Proc Natl Acad Sci U S A. 2008; 105: 2498-503. doi: 10.1073/pnas.0710787105.
19. Wali JA, Raubenheimer D, Senior AM, Le Couteur DG, Simpson SJ. Cardio-metabolic consequences of dietary carbohydrates: reconciling contradictions using nutritional geometry. Cardiovasc Res. 2021; 117: 386-401. doi: 10.1093/cvr/cvaa136/5835276.
20. Simpson SJ, Raubenheimer D, Cogger VC, Macia L, Solon-Biet SM, Le Couteur DG, George J. The nutritional geometry of liver disease including non-alcoholic fatty liver disease. J Hepatol. 2018; 68: 316-25. doi: 10.1016/j.jhep.2017.10.005.
21. Berná G, Romero-Gomez M. The role of nutrition in non-alcoholic fatty liver disease: pathophysiology and management. Liver Int. 2020; 40: 102-08. doi: 10.1111/liv.14360.
22. Fotheringham AK, Solon-Biet SM, Bielefeldt-Ohmann H, McCarthy DA, McMahon AC, Ruohonen K et al. Kidney disease risk factors do not explain impacts of low dietary protein on kidney function and structure. iScience. 2021; 24: doi: 10.1016/j.isci.2021.103308.
23. Simpson SJ, Le Couteur DG, James DE, George J, Gunton JE, Solon-Biet SM, Raubenheimer D. The geometric framework for nutrition as a tool in precision medicine. Nutr Healthy Aging. 2017; 4: 217-26. doi: 10.3233/nha-170027.
24. Chinese Nutrition Society. Guideline for the prevention and treatment of type 2 diabetes mellitus in China (2020 edition). Chin J Diabetes Mellitus. 2020; 13: 315-409. doi: 10.3760/cma.j.cn115791-20210221-00095. (in Chinese)
25. Willett WC.Nutritional epidemiology.2nd. New York: Oxford University Press; 1998.
26. Roustazadeh A, Mir H, Jafarirad S, Mogharab F, Hosseini SA, Abdoli A, Erfanian S. A dietary pattern rich in fruits and dairy products is inversely associated to gestational diabetes: a case-control study in Iran. BMC Endocr Disord. 2021; 21: 41. doi: 10.1186/s12902-021-00707-8.
27. Zhang CX, Ho SC. Validity and reproducibility of a food frequency questionnaire among Chinese women in Guangdong province. Asia Pac J Clin Nutr. 2009; 18: 240-50. doi: 10.6133/apjcn.2009.18.2.13.
28. Yang YX, Wang GY, Pan XC. Chinese food composition tables. Beijing: Peking University Medical Press; 2009. (in Chinese)
29. Mijatovic-Vukas J, Capling L, Cheng S, Stamatakis E, Louie J, Cheung NW et al. Associations of diet and physical activity with risk for gestational diabetes mellitus: a systematic review and meta-analysis. Nutrients. 2018; 10: 698. doi: 10.3390/nu10060698.
30. Zhang C, Schulze MB, Solomon CG, Hu FB. A prospective study of dietary patterns, meat intake and the risk of gestational diabetes mellitus. Diabetologia. 2006; 49: 2604-13. doi: 10.1007/s00125-006-0422-1.
31. Bao W, Bowers K, Tobias DK, Hu FB, Zhang C. Prepregnancy dietary protein intake, major dietary protein sources, and the risk of gestational diabetes mellitus: a prospective cohort study. Diabetes Care. 2013; 36: 2001-8. doi: 10.2337/dc12-2018.
32. Pang WW, Colega M, Cai S, Chan YH, Padmapriya N, Chen L-W et al. Higher maternal dietary protein intake is associated with a higher risk of gestational diabetes mellitus in a multiethnic asian cohort. J Nutr. 2017; 147: 653-60. doi: 10.3945/jn.116.243881.
33. Griffioen-Roose S, Mars M, Siebelink E, Finlayson G, Tomé D, de Graaf C. Protein status elicits compensatory changes in food intake and food preferences. Am J Clin Nutr. 2012; 95: 32-38. doi: 10.3945/ajcn.111.020503.
34. Raubenheimer D, Simpson SJ. Protein appetite as an integrator in the obesity system: the protein leverage

hypothesis. Philos Trans R Soc Lond B Biol Sci. 2023; 378: 20220212. doi: 10.1098/rstb.2022.0212.

35. Xu L, Lin X, Li X, Hu Z, Hou Q, Wang Y, Wang Z. Integration of transcriptomics and metabolomics provides metabolic and functional insights into reduced insulin secretion in MIN6 β-cells exposed to deficient and excessive arginine. FASEB J. 2022; 36: e22206. doi: 10.1096/fj.202101723R.
36. Halperin F, Mezza T, Li P, Shirakawa J, Kulkarni RN, Goldfine AB. Insulin regulates arginine-stimulated insulin secretion in humans. Metabolism. 2022; 128: 155117. doi: 10.1016/j.metabol.2021.155117.
37. White PJ, McGarrah RW, Herman MA, Bain JR, Shah SH, Newgard CB. Insulin action, type 2 diabetes, and branched-chain amino acids: a two-way street. Mol Metab. 2021; 52: 101261. doi: 10.1016/j.molmet.2021.101261.
38. Zhou M, Shao J, Wu C-Y, Shu L, Dong W, Liu Y et al. Targeting bcaa catabolism to treat obesity-associated insulin resistance. Diabetes. 2019; 68: 1730-46. doi: 10.2337/db18-0927.
39. Nie C, He T, Zhang W, Zhang G, Ma X. Branched chain amino acids: beyond nutrition metabolism. Int J Mol Sci. 2018; 19: 954. doi: 10.3390/ijms19040954.
40. Ali AM, Kunugi H. Intermittent fasting, dietary modifications, and exercise for the control of gestational diabetes and maternal mood dysregulation: a review and a case report. Int J Environ Res Public Health. 2020; 17: doi: 10.3390/ijerph17249379.
41. Sahariah SA, Potdar RD, Gandhi M, Kehoe SH, Brown N, Sane H et al. A daily snack containing leafy green vegetables, fruit, and milk before and during pregnancy prevents gestational diabetes in a randomized, controlled trial in Mumbai, India. J Nutr. 2016; 146: 1453S-60S. doi: 10.3945/jn.115.223461.
42. Sedaghat F, Akhoondan M, Ehteshami M, Aghamohammadi V, Ghanei N, Mirmiran P, Rashidkhani B. Maternal dietary patterns and gestational diabetes risk: a case-control study. J Diabetes Res. 2017; 2017: 1-8. doi: 10.1155/2017/5173926.
43. Yong HY, Mohd Shariff Z, Mohd Yusof BN, Rejali Z, Appannah G, Bindels J, Tee YYS, van der Beek EM. The association between dietary patterns before and in early pregnancy and the risk of gestational diabetes mellitus (GDM): data from the Malaysian SECOST cohort. PLoS One. 2020; 15: e0227246. doi: 10.1371/journal.pone.0227246.
44. Bao W, Tobias DK, Olsen SF, Zhang C. Pre-pregnancy fried food consumption and the risk of gestational diabetes mellitus: a prospective cohort study. Diabetologia. 2014; 57: 2485-91. doi: 10.1007/s00125-014-3382-x.
45. Cui Y, Liao M, Xu A, Chen G, Liu J, Yu X et al. Association of maternal pre-pregnancy dietary intake with adverse maternal and neonatal outcomes: a systematic review and meta-analysis of prospective studies. Crit Rev Food Sci Nutr. 2023; 63: 3430-51. doi: 10.1080/10408398.2021.1989658.
46. Pu JH, Su YX, Yang XG, Yang YX, Zhang J, Cheng YY et al. Chinese dietary reference intakes— part 1: macronutrient. National Health and Family Planning Commission of the People's Republic of China. 2017/09/14 [cited 2023/08/16]; Available from: http://www.nhc.gov.cn/wjw/yingyang/201710/fdade20feb8144ba921b412944ffb779/files/0fa10dfb812a48b483d931972df1ccb8.pdf.

Supplementary Tables and Figures

Supplementary Table 1. Categorization of food groups in dietary pattern analysis

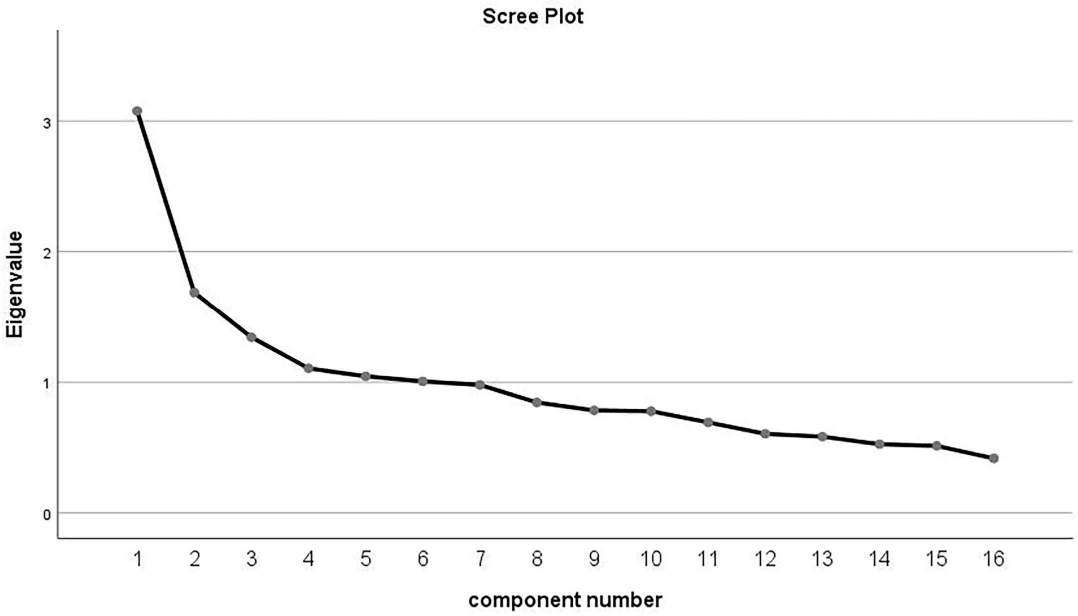
Food category	Food items
Cereals	Rice, porridge, noodles, whole wheat bread[‡], steamed cake, white bread, meat bun, deep-fried dough sticks and dumplings, etc.
Whole grains and legumes	Mung bean and corn, etc.
Potatoes	Potato and sweet potato, etc.
Soybean and soybean products	Soybean, tofu, soy milk, dried bean curd, bean curd puff, dried bean curd, bean cream skim, jellied bean curd, etc.
Vegetables	Cabbage, broccoli, cauliflower, pakchio, lettuce, spinach, Chinese cabbage, onion, green onion, garlic, dark green leafy vegetables, bean seedling, asparagus, eggplant, white gourd, cucumber, towel gourd, pumpkin, white turnip, carrot, tomato, yard-long bean, green beans, lentil, bean sprouts, etc.
Fruits	Oranges, pomelo, citrus, apples, pears, watermelon, cantaloupe, melon, grapes, banana, longan, lychee, pitaya, durian, etc.
Red meat	Pork, beef, mutton, organ meat (e.g., liver, kidney, brain, intestines), poultry meat (e.g., chicken, duck meat, goose meat, chicken feet), processed meat products (e.g., cured meat, sausages, ham, lunch meat), etc.
Aquatic products	Freshwater fish, seawater fish (e.g., hairtail, yellow croaker), mollusks (e.g., squid) crustaceans, etc.
Eggs	Eggs, duck eggs, goose eggs, quail eggs, etc.
Dairy and dairy products	Whole milk, skim milk, whole milk powder, skimmed milk powder yoghurt, milk tea, cheese.
Beverages	Carbonated beverages, fruit juice and fruit drink, tea beverages, etc.
Nuts	Peanuts, cashews, walnuts, pistachios, melon seeds, etc.
Edible fungi	Mushrooms, agaric, kelp, etc.
Desserts	Carrot cake, cake, biscuit, ice cream cones
Edible oil	Peanut oil, rapeseed oil, corn oil, blending oil, olive oil, etc.
Pickles	Salty vegetables, tuber mustard, sauerkraut, etc.

[†] Since the primary ingredient of whole wheat bread in the market is wheat flour, we classify it as a grain product based on practical considerations. However, it is fundamentally part of the whole grain category

Supplementary Table 2. Daily different food intake daily

Food groups[†]	Cases (n = 210)	Controls (n = 210)	*p*-value
Cereals	614 ± 252	608 ± 220	0.781
Deep-fried dough sticks	8.00 (0.00, 17.1)	4.00 (0.00, 12.0)	<0.001
Potatoes	50.3 ± 29.0	43.8 ± 30.8	0.032
Soybean and soybean products	58.2 ± 61.8	84.0 ± 68.2	<0.001
Fruits	118 ± 79.4	156 ± 73.7	<0.001
Aquatic products	6.55 ± 8.68	11.6 ± 12.1	<0.001
Eggs	31.9 ± 18.2	38.9 ± 16.4	<0.001
Nuts	2.31 ± 6.22	4.46 ± 8.86	0.006
Beverages	55.3 ± 84.4	37.1 ± 81.1	0.031
Desserts	12.9 ± 14.7	6.76 ± 14.8	<0.001
Ice cream cones	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.302
Edible fungi	10.6 ± 13.2	16.1 ± 15.4	<0.001
Pickles	0.85 ± 2.27	0.26 ± 0.95	0.001
Edible oil	66.9 ± 11.3	60.9 ± 12.2	<0.001
Red meat	67.8 ± 43.5	66.7 ± 40.7	0.794
Whole grains and legumes	0.00 (0.00, 13.8)	28.7 (6.67, 65.9)	<0.001
Vegetables	186 (141, 227)	241 (189, 312)	<0.001
Dairy and dairy products	56.3 (5.33, 117)	100 (40.0, 167)	<0.001

[†] Continuous variables were compared using paired samples t-test and Wilcoxon matched-pairs signed rank test.
p-value < 0.05 was considered significant



Supplementary Figure 1. The scree plot.