Supplemental Appendix

Article title

Association of dietary patterns with obesity and metabolically healthy obesity phenotype in Chinese population: a cross-sectional analysis of China Multi-Ethnic Cohort Study

Journal title

British Journal of Nutrition

Authors

Dan Tang1¶, Xiong Xiao1¶, Liling Chen2, Yixi kangzhu3, Wei Deng4, Basang5, Shujuan Yang1, Lu Long1, Xiaofen Xie1, Jiaojiao Lu1, Qun Meng1,6, Jianzhong Yin7,8\*, Feng Hong9\*, Xing Zhao1\*

¶Joint first authors, these authors contributed equally to this work

**\*** Joint corresponding authors, Emails: [xingzhao@scu.edu.cn](mailto:xingzhao@scu.edu.cn) (XZ);

[fhong@gmc.edu.cn](mailto:fhong@gmc.edu.cn) (FH); [yinjianzhong2005@sina.com](mailto:yinjianzhong2005@sina.com) (JY)

**Contents**

[Text S1. The process of choosing eligible participants 2](#_Toc86088112)

[Text S2. The process of constructing directed acyclic graph (DAG) 4](#_Toc86088113)

[Text S3. Results of the covariate balances using various weighting methods 6](#_Toc86088114)

[Text S4. Results of component analysis 8](#_Toc86088115)

[Text S5. Results of sensitivity analysis 10](#_Toc86088116)

**List of figures**

[Figure S1. The flow chart of participants enrollment 2](#_Toc86083538)

[Figure S2. The final constructed DAG 5](#_Toc86083539)

[Figure S3. The covariate balances using various weighting methods across DASH and aMED score quintiles 7](#_Toc86083540)

[Figure S4. Estimated associations by including the self-reported cardiometabolic diseases a or cancer 11](#_Toc86083541)

[Figure S5. Estimated associations by using the strict definition a of metabolic health 12](#_Toc86083542)

[Figure S6. Estimated associations by using waist-hip ratio (WHR) a criteria instead of waist circumference (WC) criteria 13](#_Toc86083543)

[Figure S7. Estimated associations by additionally adjusting for drinking status a 14](#_Toc86083544)

**List of tables**

[Table S1. Odds ratios of obesity and metabolically healthy obesity (MHO) phenotype associated with aMED component scores 8](#_Toc86083545)

[Table S2. Odds ratios of obesity and MHO associated with DASH component scores 9](#_Toc86083546)

## Text S1. The process of choosing eligible participants

To properly assess the associations of two well-known prior dietary patterns with obesity and its metabolic phenotypes, we defined detailed inclusion and exclusion criteria according to our study purpose and existing similar researches. In addition to some conventional aspects such as basic characteristics of population and missing data, we further considered the possibility of reverse causality and tried to minimize it by excluding subjects with self-reported cardiometabolic diseases or cancer. The flow chart of participants enrollment was shown in Figure S1.



**Figure S1. The flow chart of participants enrollment**

Abbreviations: CMEC = the China Multi-Ethnic Cohort; BMI = body-mass index.

a. Diet-related data included information about tea, alcohol, food frequency questionnaires (FFQs), condiment and dietary supplements;

b. Extreme total energy intake was defined as < 600 or > 3500 kcal/day for female, and < 800 or > 4200 kcal/day for male;

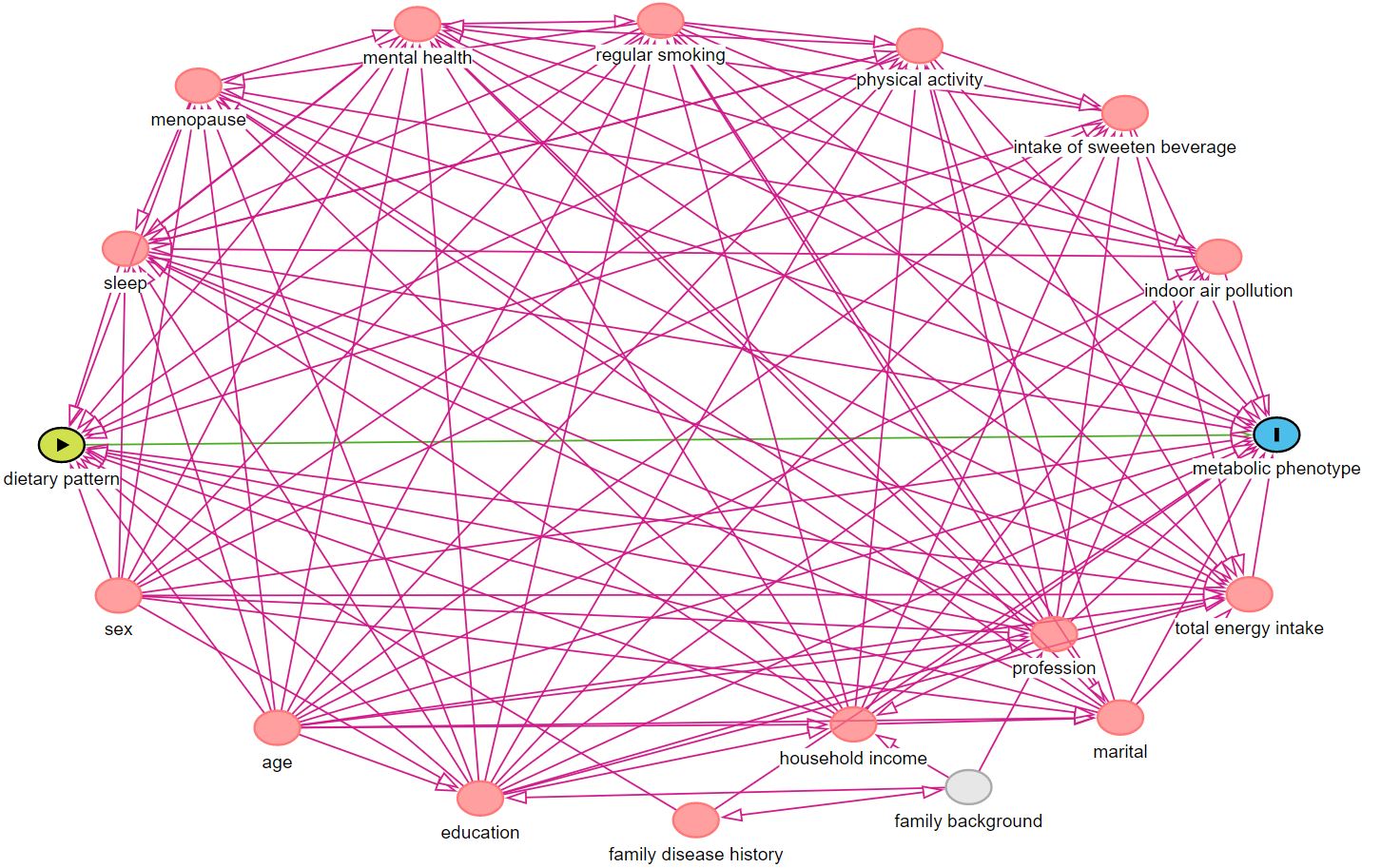
c. Underweight was defined as BMI < 18.5 kg/m2;

d. Cardiometabolic diseases included diabetes, hypertension, hyperlipidemia, coronary heart disease and stroke.

## Text S2. The process of constructing directed acyclic graph (DAG)

The DAG was constructed based on the procedure of “Evidence Synthesis for Constructing Directed Acyclic Graphs” (ESC-DAGs) method.1 First, we determined a pool of potential confounders according to systematic literature review on this study topic. Second, we assumed a saturated DAG by drawing directed or undirected edges between all pairs of variables (including exposure, outcome, and all potential confounders). Third, each edge in the saturated DAG was assessed using several causal criteria (including temporality, validity, and theoretical support) and determined as retained, reversed, bi-directional or deleted. Fourth, based on decisions on each edge, a simplified DAG was constructed, then a series of conditional independences were generated according to the current DAG. Lastly, we continuously performed the independence tests and modified the current DAG until all the implied conditional independences were satisfied in our data and the final DAG was reached. The final constructed DAG was shown in Figure S2. [The DAG was constructed using DAGitty version 3.0]

According to the final constructed DAG and back-door criterion, the minimal sufficient set of confounders includes age, sex, education, household income, profession, marital status, smoking, physical activity, total energy intake, intake of sweeten beverage, sleep status, mental health, menopause status for women, and family history of cardiometabolic diseases. In addition, we further adjusted for several regional-level confounders (including urbanicity and ethnic group) and certain diet-related variables which were not considered in these two dietary patterns (i.e., intake of dietary supplements, spicy food, and peppery food).



**Figure S2.** **The final constructed DAG**

**References**

1. Ferguson KD, McCann M, Katikireddi SV, et al. Evidence synthesis for constructing directed acyclic graphs (ESC-DAGs): a novel and systematic method for building directed acyclic graphs. *Int J Epidemiol* 2020;49(1):322-29. doi: 10.1093/ije/dyz150 [published Online First: 2019/07/22]

## Text S3. Results of the covariate balances using various weighting methods

The propensity score-based weighting is a well-recognized and widely used method in observational studies to balance confounders among groups defined by exposure status and then to estimate the treatment effects of interest. However, as a parametric model, the propensity score model (selection model/exposure model) may suffer from misspecification in many circumstances and result in suboptimal balance level of confounders and violation of the critical (conditional) exchangeability assumption for causal inference. Several new weighting methods have been proposed to improve their balancing properties, such as covariate balancing propensity scores weighting, propensity score weighting using generalized boosted models, propensity score weighting using SuperLearner, entropy balancing weighting and empirical balancing calibration weighting 1. In the present study, we compared propensity score-based weighting with abovementioned five new methods and determined the optimal weighting method which achieve the optimal balance of confounders to estimate the treatment effect of interest. For continuous covariates, we used standardized mean differences as the measurement statistic of covariate balance. For binary variables (multinomial variables were treated as multiple binary variables), we used the raw difference in proportion between two groups as the measurement statistic of covariate balance. A value closer to zero indicates a better covariate balance for both measurement statistics. As the dependent variable of the weighting model was multi-categorical (i.e., quintiles of dietary pattern score), only the maximum difference across all pair comparisons was displayed for simplicity. The results of covariate balances using various weighting methods are shown in Figure S3 and the entropy balancing weighting method produced an optimal covariate balance.



**Figure S3. The covariate balances using various weighting methods across DASH and aMED score quintiles**

ps: propensity score weighting; gbm: propensity score weighting using generalized boosted models; cbps: covariate balancing propensity scores weighting, super: propensity score weighting using SuperLearner; ebcw: empirical balancing calibration weighting; ebal: entropy balancing weighting

**References**

1. Greifer N. WeightIt: Weighting for Covariate Balance in Observational Studies. R package version 0.10.2. ed, 2020.

## Text S4. Results of component analysis

To investigate the association of each individual component of the aMED and DASH diet with obesity and metabolic obesity phenotypes and to assess the relative importance of each component, we conducted component analysis using methods proposed by Trichopoulou 1. Firstly, we evaluated association of each component by including all component scores (as continuous, ranging from 1 to 5 for each component) simultaneously in multivariable logistic regression models (adjusting for the same confounders as in the main analysis) for each dietary pattern, thus the estimated coefficients represented the mutually-adjusted associations of these components, respectively. The results were shown in Table S1 and Table S2.

Subsequently, we assessed the contribution of each of the seven or eight components of the DASH and aMED index on risk for obesity and MHO respectively, by dropping one component at a time from the total score. Specifically, we estimated the associations of original diet scores and scores subtracted each component (continuous, expressed as 25% score range increment) on obesity and MHO (the models of subtracted diet scores further adjust for the corresponding subtracted component score), respectively. To assure comparability, we multiplied the estimated coefficients by 25/29 for DASH and 29/33 for aMED diet before exponentiating them to correct for the different score ranges. Then the relative importance (contribution) of specific component can be calculated as reduction in apparent effect (reduction in magnitude of the overall effect after dropping the component). The results of DASH diet were reported in the manuscript, and the results of aMED were not shown because the overall effect of aMED diet was almost null (OR=1.002 for obesity and OR=1.007 for MHO), thus a little change in OR after dropping one component would result in very large contribution proportions and such results were meaningfulness.

**Table S1****. Odds ratios of obesity and metabolically healthy obesity (MHO) phenotype associated with aMED component scores**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Component scores** | **OR (95% CI) of obesity** | ***P* value** | **OR (95% CI) of MHO** | ***P* value** |
| whole grain | 0.991 (0.971-1.011) | 0.373 | 1.030 (0.992-1.070) | 0.127 |
| fish | 1.023 (1.001-1.045) | 0.039 | 0.983 (0.943-1.025) | 0.430 |
| vegetables | 1.057 (1.038-1.077) | <0.001 | 1.054 (1.017-1.093) | 0.004 |
| legumes | 0.930 (0.912-0.947) | <0.001 | 0.987 (0.951-1.025) | 0.495 |
| fruits | 1.010 (0.989-1.031) | 0.363 | 1.007 (0.967-1.048) | 0.746 |
| MUFA: SFA | 1.119 (1.096-1.142) | <0.001 | 0.952 (0.913-0.992) | 0.018 |
| red & processed meats | 0.901 (0.879-0.924) | <0.001 | 1.008 (0.959-1.060) | 0.747 |
| alcohol | 1.044 (0.995-1.095) | 0.082 | 0.926 (0.842-1.019) | 0.114 |

Abbreviations: aMED = alternative Mediterranean; OR = odds ratio; CI = confidence interval; MUFA = monounsaturated fatty acids; SFA = saturated fatty acids.

**Table S2. Odds ratios of obesity and MHO associated with DASH component scores**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Component scores** | **OR (95% CI) of obesity** | ***P* value** | **OR (95% CI) of MHO** | ***P* value** |
| whole grain | 0.999 (0.979-1.019) | 0.883 | 1.023 (0.985-1.063) | 0.244 |
| fruits | 1.017 (0.996-1.039) | 0.109 | 0.999 (0.959-1.040) | 0.950 |
| vegetables | 1.056 (1.037-1.076) | <0.001 | 1.054 (1.017-1.094) | 0.004 |
| legumes | 0.926 (0.909-0.944) | <0.001 | 0.987 (0.950-1.024) | 0.476 |
| red & processed meats | 0.949 (0.928-0.971) | <0.001 | 0.988 (0.945-1.034) | 0.610 |
| dairy products | 0.916 (0.900-0.932) | <0.001 | 1.044 (1.009-1.081) | 0.014 |
| sodium | 0.997 (0.979-1.016) | 0.780 | 1.052 (1.015-1.091) | 0.006 |

Abbreviations: DASH = Dietary Approaches to Stop Hypertension; OR = odds ratio; CI = confidence interval.

**References**

1. Trichopoulou A, Bamia C, Trichopoulos D. Anatomy of health effects of Mediterranean diet: Greek EPIC prospective cohort study. *BMJ* 2009;338:b2337. doi: 10.1136/bmj.b2337 [published Online First: 2009/06/25]

## Text S5. Results of sensitivity analysis

We re-estimated the associations of these two diet scores with obesity or/and MHO under four different situations to assess the robustness of our results. The first sensitivity analysis aimed to assess the existence of inverse causality, the second and third analyses aimed to evaluate the influence of different definitions of metabolic health, and the last analysis further adjusted for drinking status which was a commonly used confounder but was not considered in DASH diet. See the results in Figure S4 to Figure S7.

In conclusion, our results were generally robust to all kinds of sensitivity analyses. Specifically, the positive association of DASH diet scores with MHO was moderately attenuated after including individuals with self-reported cardiometabolic diseases or cancer, which indicated that inverse causality did exist and should be considered. Using different criteria to define metabolic health or further adjusting for drinking status did not change the results obviously.



**Figure S4. Estimated associations by including the self-reported cardiometabolic diseases a or cancer**

a. The self-reported cardiometabolic diseases include hypertension, diabetes, hyperlipidemia, coronary heart disease, and stroke.

All models adjusted for age, sex, ethnic group, urbanicity, education, household income, profession, marital status, smoking, physical activity, total energy intake, regular intake of soft drinks, dietary supplements, spicy food, and peppery food, insomnia symptoms, anxiety symptoms, depression symptoms, menopause status, and family history of cardiometabolic diseases using logistic regression with inverse probability of exposure weighting (IPEW).



**Figure S5.** **Estimated associations by using the strict definition a of metabolic health**

a. The strict definition of metabolic health was absence of all following criteria: 1) triglycerides (TGs) ≥ 1.7 mmol/L; 2) high-density lipoprotein cholesterol (HDL-C) < 1.0 mmol/L for males and < 1.3 mmol/L for females; 3) SBP ≥ 130 mmHg or DBP ≥ 85 mmHg; 4) fasting blood glucose (FBG) ≥ 5.6 mmol/L.

All models adjusted for age, sex, ethnic group, urbanicity, education, household income, profession, marital status, smoking, physical activity, total energy intake, regular intake of soft drinks, dietary supplements, spicy food, and peppery food, insomnia symptoms, anxiety symptoms, depression symptoms, menopause status, and family history of cardiometabolic diseases using logistic regression with inverse probability of exposure weighting (IPEW).



**Figure S6.** **Estimated associations by using waist-hip ratio (WHR) a criteria instead of waist circumference (WC) criteria**

a. The WHR criteria was defined as ≥ 0.90 for men and ≥ 0.85 for women.

All models adjusted for age, sex, ethnic group, urbanicity, education, household income, profession, marital status, smoking, physical activity, total energy intake, regular intake of soft drinks, dietary supplements, spicy food, and peppery food, insomnia symptoms, anxiety symptoms, depression symptoms, menopause status, and family history of cardiometabolic diseases using logistic regression with inverse probability of exposure weighting (IPEW).



**Figure S7.** **Estimated associations by additionally adjusting for drinking status a**

a. The drinking status was classified as current drinker or not.

All models for aMED adjusted for age, sex, ethnic group, urbanicity, education, household income, profession, marital status, smoking, physical activity, total energy intake, regular intake of soft drinks, dietary supplements, spicy food, and peppery food, insomnia symptoms, anxiety symptoms, depression symptoms, menopause status, and family history of cardiometabolic diseases using logistic regression with inverse probability of exposure weighting (IPEW) (identical to the main analysis). All models for DASH further adjusted for drinking status.