

Shared and ethnic background site-specific dietary patterns in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL)

Roberta De Vito^{1*}, Briana Stephenson^{2*}, Daniela Sotres-Alvarez³, Anna-Maria Siega-Riz⁴, Josiemer Mattei⁵, Maria Parpinel⁶, Brandilyn A. Peters⁷, Sierra A. Bainter⁸, Martha L. Daviglius⁹, Linda Van Horn¹⁰, Valeria Edefonti^{11,12}

Original Article

¹ Department of Biostatistics and Data Science Initiative, Brown University, Providence, RI, USA.

² Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

³ Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

⁴ Departments of Nutrition and Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts, Amherst, MA, USA.

⁵ Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

⁶ Department of Medicine - DAME, Università degli Studi di Udine, Udine, Italy.

⁷ Department of Epidemiology & Population Health, Albert Einstein College of Medicine, Bronx, NY, USA.

⁸ Department of Psychology, University of Miami, Coral Gables, FL, USA.

⁹ Institute for Minority Health Research, University of Illinois at Chicago, Chicago, IL, USA.

¹⁰ Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA.

¹¹ Branch of Medical Statistics, Biometry and Epidemiology "G. A. Maccacaro", Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milano, Italy.

¹² Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milano, Italy.

*: equally contributed to the paper.

Corresponding Author: Valeria Edefonti, Branch of Medical Statistics, Biometry and Epidemiology "G. A. Maccacaro", Department of Clinical Sciences and Community Health, Università degli Studi di Milano and Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, via Celoria 22, 20133 Milano, Italy; telephone: 0039 02-50320853; fax: 0039 02-50320866; email: valeria.edefonti@unimi.it.

Abstract

Dietary patterns (DPs) synthesize multiple related dietary components in one or more combined variables. A drawback of DPs is their limited reproducibility across subpopulations, especially adopting *a posteriori* DPs, derived using standard multivariate methods [e.g., factor analysis (FA)]. Standard approaches assessing reproducibility of FA-based DPs mostly rely on correlation coefficients/agreement measures between pairs of factors and do not consider any statistical model. Multi-study factor analysis builds upon standard FA model to identify DPs shared across all subpopulations and those specific to some subpopulations. Pattern reproducibility is investigated from a different perspective: a shared DP identified within multi-study factor analysis is "reproducible" since it is common to all subpopulations. Bayesian multi-study factor analysis (BMSFA) has been developed to improve DP retention and identification, two critical issues as the number of subpopulations analyzed increases.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Using baseline (2008-2011) 24-hour dietary recalls from the Hispanic Community Health Study/Study of Latinos (n=16,415), we applied the BMSFA on 42 common nutrients to identify shared and subpopulation-specific DPs where subpopulations were defined as the cross-classification of ethnic background and study site (EBS).

Overall, 4 shared DPs were identified: *Plant-based foods*, *Processed foods*, *Dairy products*, and *Seafood*. At the subpopulation level, we identified 12 EBS-specific DPs, one for each EBS category, primarily representing variants of foods from animal sources. Different nuances were expressed by subsets of fairly similar EBS-specific DPs, including an *Animal vs. vegetable source*, an *Animal source only*, and a *Poultry vs. dairy products* overarching DPs. Shared DPs from BMSFA were similar to their counterparts from standard FA and frequentist multi-study factor analysis; EBS-specific DPs from BMSFA were better characterized than those from frequentist multi-study factor analysis.

In conclusion, the BMSFA successfully captured sources of both dietary homogeneity and heterogeneity in a large well-characterized study of US Hispanics/Latino adults by ethnic background and site.

Keywords: Bayesian analysis; Bayesian multi-study factor analysis; dietary patterns; factor analysis; Hispanics/Latinos; Hispanic Community Health Study/Study of Latinos; multi-study factor analysis; reproducibility of dietary patterns.

Running title: Dietary patterns via Bayesian multi-study factor analysis.

Conflicts of interest: None declared

Sources of financial support: The Hispanic Community Health Study/Study of Latinos is a collaborative study supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (HHSN268201300001I / N01-HC-65233), University of Miami (HHSN268201300004I / N01-HC-65234), Albert Einstein College of Medicine (HHSN268201300002I / N01-HC-65235), University of Illinois at Chicago (HHSN268201300003I / N01-HC-65236 Northwestern Univ), and San Diego State University (HHSN268201300005I / N01-HC-65237). The following Institutes/Centers/Offices have contributed to the HCHS/SOL through a transfer of funds to the NHLBI: National Institute on Minority Health and Health Disparities, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Neurological Disorders and Stroke, NIH Institution-Office of Dietary Supplements.

Author contributions: RDV, BJS, and VE designed overall research plan; DS-A: assisted with the research plan and writing of the manuscript; AMS-R, JM, and MP provided advice on nutritional issues and a critical review of the manuscript; RDV performed most of the statistical analyses, with contributions from BS; RDV prepared all tables and figures except for Supplementary Figure 1; RDV assisted in writing the Methods section and provided critical review of the manuscript; BS assisted in writing the Results section and provided critical review of the manuscript; BP, GCC, SAB, DGF, MLD, LVH assisted with interpretation of results and provided critical review of the manuscript; VE wrote the paper; VE and RDV had primary responsibility for final content. AMS-R, MLD, and LVH were involved in obtaining the cohort funding; AMS-R, MLD, LVH, and DS-A were involved in conducting the study including dietary data collection. All authors reviewed and approved the final manuscript. The authors have declared no conflicts of interest.

Data sharing: Data used in this study are not publicly available, but can be obtained through an internal process of the study, based on a Data and Materials Distribution Agreement (DMDA). Supporting code to implement the methods performed in this study is publicly and freely available. In detail, we finalized an R package, named *MSFA*, to implement the methods presented in De Vito et al. [De Vito R, Bellio R, Trippa L, and Parmigiani G, “Multi-study factor analysis” and “Bayesian multi-study factor analysis for high-throughput biological data”] and their application in nutritional epidemiology. The package is available on GitHub: <https://github.com/rdevito/MSFA>

Acknowledgements: The authors thank the staff and participants of HCHS/SOL for their important contributions. Investigators website - <http://www.csc.unc.edu/hchs/>.

Introduction

A posteriori dietary patterns (DPs) (1), defined by using multivariate statistical methods [i.e., principal component analysis (PCA), factor analysis (FA), and cluster analysis (2)], are advantageous in naturally reflecting actual study- or population-specific context (e.g., geography/climate, socioeconomic status, food supply, ethnic background, religion) (3), but their specificity limits generalizability, compared to the *a priori* (i.e., comparing subjects' diet against evidence-based benchmark diets) options (4). The absence of a standardized approach to analysis has further hindered discovery of genuinely reproducible *a posteriori* DPs across subpopulations within the same study or other studies from the same or similar populations (i.e., *cross-study* reproducibility (5)). The two issues have traditionally limited comparisons among sets of *a posteriori* DPs, thereby preventing firm conclusions about their health benefits (or risks) (6, 7).

To our knowledge, only one paper (8) has explored reproducibility of *a posteriori* DPs in subpopulations belonging to the same study. This study used stratified PCAs by US region, sex, and race to choose the optimal number of components to retain for a final analysis based on the overall sample. In a multi-cultural population, a full exploration of regional and ethnic diversity is important, but it requires efficient statistical methods to manage and compare the different level-specific solutions (9). Yet, different ethnic backgrounds pose additional research challenges, including identification of a proper food grouping scheme and coping with possible masked acculturation effects on dietary habits (10).

A few pioneering (11, 12) and more recent (13-19) papers have explored *cross-study reproducibility* of PCA-based or FA-based DPs across different studies. One of two major statistical approaches is applied on a common set of variables across studies (3): 1. a stratified approach, i.e., each study expresses its own set of DPs and a reproducibility analysis is carried out using congruence/correlation coefficients or agreement measures

(13, 14, 18); or 2. a pooled approach, i.e., studies are merged in a single dataset and forced to share common DPs (15, 16).

Multi-study factor analysis (20) (MSFA), recently applied in nutritional epidemiology (17), identifies shared DPs, common to all studies, as well as additional study-specific DPs for individual studies. Shared DPs identified within MSFA are “reproducible” across studies since they are common to all studies. Estimation of MSFA model parameters using Bayesian techniques [Bayesian multi-study factor analysis (BMSFA)] reportedly improve factor identification and choice of the number of factors to retain (21). These are two issues of great importance as the number of studies/subpopulations analyzed increases.

The Hispanic Community Health Study/Study of Latinos (HCHS/SOL) (22) is an ongoing US multi-site community-based cohort studying health and risk factors of cardiovascular and pulmonary outcomes of Hispanic/Latino adults, and differences across several ethnic backgrounds. This cohort provides the unique opportunity to define both shared and subpopulation-specific DPs by the cross-section of ethnic background and site (EBS), in support of future culturally tailored interventions. Two papers within HCHS/SOL have identified shared and subpopulation-specific *a posteriori* DPs using approaches conceptually similar to the MSFA (23, 24). Stephenson et al. (24) used robust profile clustering to cluster participants and food items (from the food propensity questionnaire) based on consumption behaviors shared amongst all participants and those specific to 9 EBS subpopulations (24). Compared to robust profile clustering, MSFA derives DPs analogous to those from FA or PCA, providing a more straightforward interpretation for nutrition researchers. In addition, robust profile clustering has limited functionality in handling continuous data like nutrient intakes. Maldonado et al. (23) conducted stratified principal FAs by 6 ethnic backgrounds on a common set of food groups from 24-hour recalls. Overarching DPs were defined based on food groups that shared high loadings in multiple ethnic backgrounds, but DP similarity was not statistically evaluated. Even though

investigator-consensus was used to reduce inherent subjectivity, overall factor selection is cumbersome to manage when several stratified FAs are performed, influencing the number and composition of the identified overarching DPs. In addition, associations of background-specific DPs with health outcomes are difficult to interpret. Our MSFA derives reproducible DPs directly, improving factor retention through BMSFA, and easing interpretability of associations with health outcomes of interest.

This paper aims to explore the use of BMSFA to jointly identify shared and EBS-specific DPs from a common set of nutrients derived from 24-hour recalls to address the following research questions: 1. Are there consistent and empirically estimable DPs shared across the HCHS/SOL? 2. Are there one or more additional DPs for some EBS-specific categories? 3. Are the identified shared and EBS-specific DPs interpretable in terms of food groups, and socio-demographic and lifestyle factors?

Materials and methods

Study design

The HCHS/SOL enrolled 16,415 adults aged 18–74 years residing in four US field sites (Bronx, Chicago, Miami, San Diego) from six Hispanic/Latino ethnic backgrounds (Cuban, Dominican Republic, Mexican, Puerto Rican, Central and South American) at baseline (2008-2011). A stratified 2-stage area probability design was adopted. Sampling design and cohort selection have been described previously (25). Participants provided written informed consent, and the Institutional Review Boards from each field center and the University of North Carolina at Chapel Hill (Coordinating Center) approved the study.

Protocol and Measurements

The baseline visit included questionnaires administered in Spanish or English based on the participant's language preference, anthropometry, and two 24-hour dietary recalls, among many other procedures. The first recall was administered in person and the 2nd one via

telephone, preferably ≤ 30 days from baseline in the participant's preferred language by trained interviewers using the Nutrition Data System for Research software (version 11) (26). Virtually all participants (99%) provided at least one recall.

Selection of participants

From the original cohort (n=16,415), participants were excluded who self-identified as belonging to other/multi/mixed (n=938) ethnic backgrounds, had recall data deemed unreliable by the interviewer (n=65), or provided extreme (i.e., <1st or >99th sex-specific percentiles) estimates of energy intake (n=171). After these exclusions, we removed 184 participants belonging to EBS categories with <200 participants – “fair” sample size to carry out FA (27) (i.e., South American - Bronx) – and 36 with missing Hispanic/Latino background. Overall, 15,021 participants were used for analysis from 12 EBS categories (**Supplemental Figure 1**).

Specification of variables and data preprocessing

From the Nutrition Data System for Research list of 139 available nutrients (26), we selected 42 that best represent the overall diet for Hispanics/Latinos; we expanded the fat profile to capture cardiovascular disease-related dietary habits (28, 29). Only nutrient intake from foods, not from supplements, was considered. We derived final intake from one or the mean of two available reliable recalls. Intakes were all log-transformed (base e) to improve the normality of the shared and EBS-specific factors, and of the EBS-specific errors (details in **Supplemental Methods**).

Statistical analysis

Factorability of the correlation matrices

We used Bartlett's test of sphericity, overall and individual measures of sampling adequacy (27), to assess whether the 12 EBS-specific correlation matrices and the overall correlation matrix of the log-transformed nutrient intakes were factorable.

Identification of nutrient-based dietary patterns

We carried out BMSFA (21) on the EBS-specific log-transformed data correlation matrices to estimate unobservable shared (K) and EBS-specific (J_s) factors, known as DPs. Compared with the frequentist approach (20) [here named frequentist multi-study factor analysis (FMSFA)], BMSFA offers two advantages: 1. a better-defined loading structure via the prior distribution that makes loadings extreme, and 2. a practical and useful approach (eigenvalue decomposition and 5% variance explained cut-off) to choose the number of shared and EBS-specific factors. After estimating the factor loadings, we applied the varimax rotation to the shared factor-loading matrix to obtain a better-defined loading structure. We named 'dominant nutrients' (27) those showing shared rotated (or EBS-specific unrotated) factor loadings $\geq|0.60|$ ($\geq|0.30|$). Factor scores estimate the degree to which each participant's diet is summarized by each identified DP. We calculated factor scores in BMSFA by following FA's standard Thurstone approach (30, 31) (correlation with Bartlett method ≥ 0.90 for the shared factors).

We evaluated the internal reliability of DPs in two ways. First, we assessed internal consistency with standardized Cronbach's alpha and *alpha-when-item-deleted* coefficients for those nutrients that loaded $>|0.40|$ on the shared ($>|0.20|$ on the EBS-specific) factors (27). Second, we assessed internal reproducibility by comparing BMSFA-based DPs with those from principal component FA (PCFA) on the overall sample (16) and FMSFA (17, 20). Throughout the paper, we referred to the congruence coefficient (CC) and its cut-offs proposed in (14) for DP comparison: "fair similarity" corresponds to $0.85 \geq CC \leq 0.94$ and "equivalence" to $CC > 0.95$ (**Supplemental Methods** for additional details).

Interpreting dietary patterns in terms of food groups, socio-demographic and lifestyle characteristics

We separately classified participants based on factor scores' quintiles (shared factors) or tertiles (EBS-specific factors) to validate DPs against selected 24-hour-recall-based food groups, socio-demographic and lifestyle factors. For the food-group-based validation, we estimated adjusted mean intakes of food groups (servings/day), stratified by quantile-based categories of DPs, using a linear regression model adjusted for age, sex, body mass index (BMI), EBS category (for shared factors only), and total energy intake. We expressed the deviation of the food group intake relative to its overall (i.e., based on the total final sample size) or EBS-specific mean as: $100\% \times [\text{adjusted mean within quintile-based category} / \text{overall HCHS/SOL mean}]$ for the shared factors and $100\% \times [\text{adjusted mean within tertile-based category} / \text{EBS-specific mean}]$ for the EBS-specific factors. Therefore, if the consumption of a food group exceeded 100%, individuals from that factor score quantile category were characterized by high consumption of that food group, compared with the HCHS/SOL overall (shared factor) or EBS-specific (EBS-specific factor) mean, and vice versa when the relative intake is below 100%.

For the validation of DPs against selected socio-demographic and lifestyle characteristics, we used the Pearson Chi-square test of independence (categorical characteristics) or the ANOVA (continuous characteristics); adjustment for multiple comparisons was carried out with the False Discovery Rate Method (32). To provide a comparison between the identified DPs and the Alternative Healthy Eating Index (AHEI-2010) – a measure of overall diet quality previously related to major chronic disease risk and mortality, with strong associations with cardiometabolic disease (11 components, score range: 0-110, lowest to highest quality)] (33) – we additionally estimated the adjusted mean AHEI-2010 score within quantile-based categories of factor scores using a linear regression model adjusted for age, sex, BMI, EBS combination (for shared factors only), and total energy intake. We calculated the p-values for linear trend based on the same models.

Except for the implementation of BMSFA, all statistical analyses accounted for HCHS/SOL complex survey design, including survey weights, stratification and clustering (34). All statistical tests were two-sided. Calculations were carried out using the open-source statistical computing environment R (35), with its libraries “statmod” (36), “psych” (37), “nFactors” (38), “ggplot2” (39), “survey” (40), and “MSFA” (41).

Results

Population characteristics

Socio-demographic and lifestyle characteristics by EBS categories were included in **Table 1**. The largest and smallest EBS categories were participants of Mexican background from San Diego (n=3775) and Bronx (n=205), respectively. Across all EBS categories, most individuals were between the ages of 18-44 years and reported an income of less than \$30,000; in most EBS categories, the majority of individuals were first-generation immigrants and non-consumers of supplements (all p-values<0.001). Differences between EBS categories were observed for sex, years living in the United States, age of immigration, marital status, employment status, education, physical activity, BMI, and energy intake. The highest percentage of individuals born in mainland US was identified for the Cuban background – Miami category, which also had the highest percentage of not meeting the 2008 Physical Activity Guidelines for Americans and the highest mean energy intake. Individuals of Mexican background from the Bronx had the highest percentage of being 18–44 years old and married, receiving an income <\$30,000, and following the 2008 Physical Activity Guidelines for Americans. They also showed almost the lowest mean energy intake. The Mexican background – Chicago individuals had the highest AHEI-2010 mean value, whereas individuals of Puerto Rican backgrounds from Chicago and the Bronx had the lowest AHEI-2010 scores, and the lowest mean age at arrival at mainland US. The Mexican background – San Diego category had the highest percentage of individuals in the >\$30,000 income category. The South American background - Miami category had the highest percentage of highly educated individuals and individuals who used dietary supplements.

Identification of nutrient-based dietary patterns

Factorability of correlation matrices was confirmed for each of the 12 EBS-specific and overall correlation matrices (**Supplemental Results** and **Supplemental Table 1**). The BMSFA estimated 4 shared DPs, common to all EBS categories (explaining 62.5% of the total variance, **Supplemental Table 2**) and one EBS-specific DP for each of the 12 categories (explaining a variance ranging from 10.7% for Central American background – Miami to 14.4% for Puerto Rican background – Chicago, **Supplemental Table 3**). A heatmap illustrates the factor loadings for shared and EBS-specific DPs (**Figure 1**).

Shared dietary patterns

Factor 1, namely *Plant-based foods*, showed high (i.e., dark blue in **Figure 1**, values in **Supplemental Table 2**) factor loadings on vegetable protein, phosphorus, magnesium, iron, zinc, copper, potassium, manganese, thiamin, niacin, pantothenic acid, vitamin B6, natural folate, soluble and insoluble dietary fiber. Factor 2, namely *Processed foods*, showed high loadings on long-chain saturated fatty acids, long-chain monounsaturated fatty acids, linoleic acid, linolenic acid, total trans fatty acids, and natural alpha-tocopherol. Factor 3, namely *Dairy products*, showed high loadings on short- and medium-chain saturated fatty acids, calcium, riboflavin, vitamin B12, retinol, and vitamin D. Factor 4, namely *Seafood*, showed high loadings on eicosapentaenoic acid (EPA), docosapentaenoic acid (DPA), and docosahexaenoic acid (DHA).

Ethnic background-site specific (EBS) dietary patterns

Ten of the identified EBS-specific DPs represented variants of an animal profile, as they were characterized by 2 or 3 nutrients among animal protein, arachidonic acid, and niacin (dark blue color in **Figure 1**). Based on CC (**Table 2**) and visual inspection of the factor loadings (**Supplementary Table 3**), additional profile similarities were related to common ethnic background or site and were expressed through the following three overarching EBS-specific DPs (i.e., DPs shared among some EBS-specific categories):

(1) *Animal vs. vegetable source*: animal protein/arachidonic acid vs. soluble and insoluble dietary fiber (CC reached 0.95, meaning equivalence, between DPs of Dominican background – Bronx and South American background – Chicago);

(2) Animal source only: animal protein/arachidonic acid/cholesterol with the same sign vs. no other dominant nutrients ($0.85 \geq CC \leq 0.94$, meaning fair similarity, for DPs of the following pairs: Puerto Rican background – Bronx, Central American background – Chicago, Cuban background – Miami, and South American background – Miami);

(3) Poultry vs. dairy products: less clearly identified and characterized by combinations of animal protein/arachidonic acid/niacin/vitamin B6/vitamin B12 vs. small- and medium-chain saturated fatty acids/soluble and insoluble fiber ($0.85 \geq CC \leq 0.94$, meaning fair similarity for 4 out of 6 possible pairs of EBS-specific DPs among the following: Central American background – Bronx, Central American background – Miami, Mexican background – Bronx, Mexican background – Chicago).

In contrast, one EBS-specific category, Puerto Rican background – Chicago, was characterized by a strikingly different DP, high on beta-carotene (dark blue color) and low on starch, iron, and thiamin (orange color) (**Figure 1**), which showed no fair similarity with any other EBS-specific DP. In addition, the Mexican background – San Diego DP loaded high on iron, niacin, vitamin B6, vitamin B12, as opposed to natural folate, soluble and insoluble fiber (**Figure 1**), showing fair similarity with the Mexican background – Chicago DP only (**Table 2**).

Finally, of the 24 CCs (36%) indicating fair similarity/equivalence between pairs of EBS-specific DPs in the upper triangular matrix (**Table 2**), 11 (underlined in the table) targeted similarities between DPs belonging to the same overarching structure and 13 targeted similarities between DPs belonging to different overarching structures, thus suggesting similarities also across overarching DPs.

Internal reproducibility and consistency of the identified patterns

Nutrient communalities and internal consistency of DPs, measured by Cronbach's alpha, were satisfactory, further supporting our selection of nutrients (**Supplementary Table 4** and **Supplemental Results**). The internal reproducibility of BMSFA-based DPs – as based on the comparison with FMSFA and PCFA – was reassuring (**Figure 2, Supplemental Results** and **Supplemental Tables 5-8**). Indeed, BMSFA estimated the same number of DPs and a similar total variance explained under FMSFA and PCFA, although DP order differed in

PCFA, with a more prominent role of the *Processed foods* DP (**Supplemental Table 5**). The BMSFA-based shared DPs were equivalent to their counterparts from FMSFA (all CCs ≥ 0.97) and PCFA (all CCs ≥ 0.95) (**Figure 2**, which shows a heatmap of CCs between the shared DPs identified with the BMSFA and those from PCFA and FMSFA). Visual inspection of the factor-loading matrix for the shared DPs (**Supplemental Table 5**) suggested that BMSFA was more effective in: 1. shrinking/increasing moderately low/high loadings in absolute value towards 0 or 1; 2. forcing potentially dominant nutrients to load on one DP instead of two.

While percentages of explained variances were similar to their BMSFA-specific counterparts, FMSFA-specific DPs generally opposed animal and vegetable sources of foods (animal protein/cholesterol/arachidonic acid vs. vegetable protein/folate/soluble and insoluble fiber in ≥ 8 DPs, **Supplemental Table 6**) and therefore showed fewer nuances than the corresponding BMSFA-based ones (fair similarity/equivalence for FMSFA: 92% vs. 36% for BMSFA, **Supplemental Table 7** vs. **Table 2**). This was confirmed in the one-to-one comparison between FMSFA- and BMSFA-based versions of the same EBS-specific DP, with 4 DPs only being fairly similar under the two approaches (**Supplemental Table 8**).

Top consumers of shared dietary patterns by ethnic background-site category

Ethnic background-site categories were well represented in the top quintile category of each shared DP, with a prevalence (i.e., weighted percentage) around the expected 20% (i.e., 18-22%) for most EBS categories (**Supplemental Table 9**). Major deviations from the 18-22% interval were observed for the *Seafood* DP, with percentages as low as 13.1% (Mexican background – Bronx) and as high as 24.3% (South American background – Chicago).

Food groups associated with the identified dietary patterns

Shared dietary patterns

Table 3 shows the deviation (%) of the adjusted food group mean for individuals in the top quintile of each shared DP, relative to the overall adjusted mean. Individuals in the top quintile category of the *Plant-based foods* DP were characterized by high (i.e., $>140\%$)

intakes of fruit (citrus and other) and vegetables (dark green, orange, tomatoes, beans, and others), whole grain, nuts and seeds, and alcohol, compared to the overall HCHS/SOL mean. Individuals in the top quintile category of *Processed foods* DP were characterized by high intakes of starchy vegetables, refined grain, red meat, deli meat, eggs, nuts and seeds, cheese, milk-based desserts, sugar sweetened beverages, alcohol, snacks, and added fats (regular and reduced). Individuals in the top quintile category of the *Dairy products* DP showed high consumption of deli meats, milk, cheese, yogurt, and milk-based desserts; they also showed a low (i.e., <60%) consumption of nuts and seeds. Finally, individuals in the top quintile category of the *Seafood* DP presented an extremely high fish intake, and high consumption of dark green vegetables, poultry, and alcohol.

Ethnic background-site-specific (EBS) dietary patterns

Compared with each EBS-specific mean, participants in the top tertile category of most EBS-specific DPs showed a higher-than-mean consumption of poultry and alcohol vs. a lower-than-mean consumption of nuts and seeds, thus confirming the general animal source of most EBS-specific DPs. Red meat was also highly consumed in the top tertile category of half of the EBS-specific DPs. Overarching DPs additionally showed:

1. *Animal vs. vegetable source*: higher-than-mean consumption of poultry vs. lower-than-mean consumption of non-citrus fruit (Dominican background – Bronx) or beans (South American background – Chicago);
2. *Animal source only*: higher-than-mean consumption of red meat (except for South American background – Miami, but 132% was close to 140%) vs. no clear lower-than-mean consumption of other food groups;
3. *Poultry vs. dairy products*: higher-than-mean consumption of poultry vs. lower-than-mean consumption of one or more among cheese, yogurt, and dairy dessert products, but still less clearly identified; lower-than-mean consumption of dairy products was common to individuals of Central American background from all sites and those of Cuban background from Miami, who belonged to the *Animal source only* overarching DP.

Individuals of Puerto Rican background – Chicago category in the top tertile showed a higher-than-mean consumption of yogurt, milk-based dessert products, (non-citrus) fruit, dark green and orange vegetables; this was different from any other EBS-specific DP,

including that of their Bronx counterparts of Puerto Rican background. Finally, individuals of Mexican background from Chicago and San Diego showed a similar pattern of consumption, including red meat, poultry, and alcohol, but fewer nuts and seeds, and beans. However, no similarities were found with the few Mexican background individuals from Bronx, whose top-consumers showed high consumption of citrus fruit and sugar-sweetened beverages (**Table 4**).

Alternative Healthy Eating Index associated with the Identified patterns

Shared dietary patterns

Higher quintile-based categories of *Plant-based foods* and *Seafood* DPs were consistently and significantly associated with higher mean AHEI-2010 scores (p for trend < 0.001, 5-point increment from lowest to highest category for both DPs), suggesting a higher quality of top-consumers' overall diet. *Dairy products* and *Processed foods* DPs showed the opposite trend (p for trend < 0.001): lower mean AHEI-2010 scores (i.e., lower quality) were consistently reported for increasing quintile-based categories of these DPs, with an overall ~3- and ~5-point decrease, respectively (**Figure 3**).

Ethnic background-site-specific (EBS) dietary patterns

Lower mean AHEI-2010 scores were consistently observed for increasing tertile-based categories of 11 EBS-specific combinations (all p for trend < 0.05, except for Central American background – Bronx category: p=0.12). In contrast, mean AHEI-2010 scores globally increased by ~2 points for increased consumption in participants of Puerto Rican background from Chicago. For the same tertile-based category, mean AHEI-2010 scores markedly differed across EBS-specific DPs, with Puerto Rican background participants from Bronx and Chicago showing the lowest mean scores [39.85 (SE: 0.20) and 42.15 (SE: 0.22), respectively] and Mexican background participants from Chicago and San Diego showing the highest mean scores [56.26 (SE: 0.15) and 54.27 (SE: 0.22), respectively] (**Figure 3**).

Socio-demographic and lifestyle factors associated with the Identified patterns

Shared dietary patterns

While top-consumers of all shared DPs were more likely to be younger and males, percentages of the youngest or male participants were even more extreme in top-consumers of the *Processed foods* DP. Notably, top-consumers of the *Plant-based foods* and *Seafood* DPs were less likely to have spent more than 10 years in the US, whereas those of the *Dairy products* and *Processed foods* DPs were more likely to have spent more than 10 years in the US. Top-consumers of the *Processed foods* DP showed a lower weighted mean age at immigration. While most of the overall population did not meet the 2008 Physical Activity Guidelines for Americans, top-consumers of the *Plant-based foods* and *Processed foods* DPs were more likely to be active, than the lowest consumers. Use of supplements was more likely in the top-consumers of the *Plant-based foods* and *Dairy products* DPs and less likely in top-consumers of the *Processed foods* DP (**Table 5**, with most adjusted p-values<0.05).

Ethnic background-site-specific (EBS) dietary patterns

The identified EBS-specific DPs were sparingly related to the selected socio-demographic and lifestyle factors. Age, sex, age of immigration, education, and use of supplements showed adjusted p-values<0.05 for a range of two (age of immigration) to seven (sex) DPs. The two DPs expressed by Chicago and San Diego individuals of Mexican background were significantly related to most (54% and 62%, respectively) of the selected variables in the same direction: top-consumers of these DPs were more likely to be younger, to be first generation immigrants, to have lived less than 10 years in the US, to be married/living with a partner, to report an income less than \$30,000, and to have less than a high school education (**Supplementary Table 10**).

Discussion

This paper explores the use of BMSFA to assess reproducibility of a *posteriori* nutrient-based DPs across the 12 EBS categories available in HCHS/SOL. Four DPs - *Plant-based foods*, *Processed foods*, *Dairy products*, and *Seafood* – were shared across all EBS categories and were therefore reproducible. One EBS-specific DP was additionally identified for each of the 12 categories. The 12 EBS-specific DPs represented variants of an animal

profile associated with consumption of poultry, alcohol, and sometimes red meat. Most EBS-specific DPs showed additional similarities in profile and were further grouped into overarching DPs, i.e., an *Animal vs. vegetable source*, an *Animal source only*, and a *Poultry vs. dairy products* DPs. Shared DPs from BMSFA were similar to their counterparts identified in PCFA and FMSFA, but EBS-specific DPs from BMSFA showed more nuances than those from FMSFA.

The shared DPs capture traits common to all EBS-specific populations in HCHS/SOL, independently of background and site of recruitment, and are meant to reflect key aspects of the overall diet of Hispanic/Latino adults that reside in the US. These DPs target vegetable (*Plant-based DP*) and animal sources (*Dairy foods* and *Seafood DPs*) of foods and the fat component of foods, as derived from vegetable and animal sources (*Processed foods DP*). Despite the central role of cereals and meat in Hispanic/Latino diets (42-44), we have not identified separate shared DPs for grains (i.e., a DP loading high on starch and/or total sugars and/or fiber) or meat (i.e., a DP loading high on animal protein and/or cholesterol). Grains were rather eaten together with vegetables or meats; meat was also consumed in combination with vegetables or dairy products or was included in baked or fast-food products. However, except for *Processed foods DP*, no other combinations including grains or meat were prevalent enough to emerge as clear separate DPs common to all EBS-specific categories. The identified shared DPs and the absence of separate DPs for grains and meat likely reflect a combination of a general Hispanic/Latino culinary tradition (45) and acculturation to more US-American diets.

The 12 EBS-specific DPs summarize variants of that animal profile captured only in part by the shared DPs. Overarching DPs likely capture a combination of background-specific (e.g., culinary traditions, cultures, beliefs) (45) and site-specific (e.g., food access and environment) (46) factors. Curiously, every overarching DP presented with EBS-specific categories that differed in both background and site. On the other hand, at the lowest AHEI-

2010 mean scores, individuals of Puerto Rican background from Bronx and Chicago showed strikingly different DPs; at the highest AHEI-2010 mean scores, similarities were found between DPs of Mexican background from San Diego and Chicago, but not with those from Bronx. This further suggests the need to conduct analyses at the cross-section of site and background levels and the importance of using *a posteriori* DPs, which revealed differences in background by site, unexpected from previous analyses based on the AHEI-2010 (47).

A few previous studies identified *a posteriori* food-group-based DPs with factor or cluster analyses conducted on overall populations of Hispanics/Latinos living in the US or on multi-cultural populations from US, including Hispanics/Latinos and Non-Hispanic Whites and Blacks (23, 24, 48-51). Although we reinterpreted nutrient-based DPs with food groups, the comparison of our results with existing literature is fraught with differences in dietary collection tools, age and race/ethnicity of participants, sample size, type and preprocessing of input variables, and DP identification method. Earlier than HCHS/SOL, at least 4 studies identified *a posteriori* DPs of Hispanic/Latino adults residing in the US (48-51). We restricted our comparison to those deriving DPs on Hispanics/Latinos only (48, 50). Dietary patterns of Mexican American adults from the US National Health and Nutrition Examination Survey 2001/2002 were examined in one study (48), where 4 shared clusters, named *Poultry and alcohol*, *Milk and baked products*, *Traditional Mexican*, and *Meat*, were identified. The *Traditional Mexican* and the *Meat* DPs confirm that grains (tacos/tortillas in this case) are consumed in combination with legumes (*Traditional Mexican* DP) or (red) meat (*Meat* DP) and load high on two separate DPs. The *Poultry and alcohol* DP confirms that poultry is frequently consumed with alcohol, as in most of our EBS-specific DPs. Dairy products loaded high on a separate animal-source *Milk and baked products* DP; the high loadings on cakes, cookies, and pizza suggest similarities with our *Processed foods* DP. Another study (50) reported on DPs of 45-75 years old Puerto Rican adults from the Boston area and identified 3 FA-based DPs named *Meat*, *processed meat*, and *French fries*, *Rice*, *beans*,

and oils, and Sweets, sugary beverages, and dairy desserts. The Rice, beans, and oils confirms the rice and legumes' combination; it also reveals the major role of oils. The Meat, processed meat, and French fries and the Sweets, sugary beverages, and dairy desserts DPs share similarities with our Processed foods DP; however, while our DP likely reflects acculturation to the US-American diet, the previous DPs likely capture modern industrialized diets related to nutrition transition, as they were shown not to be related to acculturation (52).

Within HCHS/SOL, two previous papers derived *a posteriori* DPs using food items/groups at either ethnic background-specific (23) or EBS-specific (24) levels. Maldonado et al. (23) described one fully and 4 partially reproducible DPs derived using FA on 34 food groups from 24-hour recalls stratifying by 6 ethnic background-specific categories. Their fully reproducible Burgers, Fries, & Soft drinks overarching DP is fairly similar to our shared Processed foods DP. Consistent with our Seafood DP, Maldonado's Fish DP (23) loaded high on fish and to less extent poultry, but it showed opposing loadings for poultry in Dominican background participants (23); we observed the same behavior of fish and poultry in the EBS-specific Dominican background – Bronx DP. Our Dairy products DP shares similarities with Maldonado's Egg & Cheese DP (23); however, starchy vegetables and processed meats (in Dominican and Puerto Rican participants only) suggest partial overlapping with our Processed foods DP. Maldonado's White Rice, Beans, & Red Meats partially overlaps with our Plant-based foods DP, which showed 7 additional vegetable and fruit groups other than beans and was more oriented towards consumption of whole grains compared to refined grains. Whereas Maldonado et al. derived DPs separately by ethnic background using standard FA (23), Stephenson et al. (24) used robust profile clustering to jointly classify individuals and 129 food propensity questionnaire items into global or local food patterns, based on 9 out of our 12 EBS categories. Stephenson's global profiles favored a more frequent consumption of fruits, vegetables, poultry, and fish in *Global Profile*

1, and foods with oils, added sugars, and eggs in *Global Profile 2* (24). This indicates that *Global Profile 1* contains a mix of elements from our *Plant-based foods* and *Seafood* DPs, whereas *Global Profile 2* contains many elements of our *Processed foods* DP. Like our BMSFA results, chicken was commonly consumed across all EBS categories, but with different levels of consumption frequency.

The association between the *Processed foods* DP and younger age at immigration/more years spent in the US aligns with results suggesting dietary practice outcomes worsen with acculturation (e.g., (51)). Similarly, top-consumers of the *Plant-based foods* and *Seafood* DPs were less likely to have spent more than 10 years in the US. Currently, there is still little agreement on how acculturation should be measured (52-54) and regression models are needed to consider the more complex patterns including acculturation, socio-demographic factors, diet and/or health outcomes, likely within a mediation analysis approach (55).

The current analysis has strengths and limitations. The HCHS/SOL is the largest cohort of individuals of Hispanic and Latino background in the United States; the study design and probability sampling in urban areas with large ethnically diverse Hispanic and Latino populations provide adequate representation. The identification of shared and EBS-specific DPs via MSFA were carried out in a single step. Also, the introduction of prior distributions (which act like rotations) in BMSFA has provided the most straightforward interpretation of the EBS-specific DPs observed, compared to FMSFA. Improved interpretation of the EBS-specific DPs under the BMSFA is not at the expense of the shared DP interpretation, as there is equivalence with shared FMSFA- or PCFA-based DPs. Based on 24-hour recall data, our identified DPs may not represent the participant's usual diet and may have failed to capture episodically consumed foods well. In general, self-reported dietary assessment tools are prone to measurement error and 24-hour recalls have been shown to underestimate total dietary intake (56). Systematic under-reporting of energy and protein intake was observed in a biomarker calibration study in HCHS/SOL that varied by ethnic

background (57), which may explain some of the lower mean energy values in certain EBS categories.

In conclusion, this application of BMSFA within HCHS/SOL reveals shared DPs that clearly document dietary habits of Hispanics/Latinos who live in the US. Using single EBS-specific DPs grouped as overarching DPs showed additional similarities in animal food sources related to the cross-section between background and site of recruitment. Although our results on the association between shared DPs and acculturation suggest dietary practice outcomes worsen with acculturation, implications for potential preventive strategies require further specific efforts to assess the more complex patterns of acculturation, socio-demographic factors, diet and/or health outcomes.

Table 1. Baseline socio-demographic and lifestyle characteristics (weighted percentages and standard error in parenthesis) by ethnic background and study site (EBS). Hispanic Community Health Study/Study of Latinos, 2008-2011.

Characteristic	Total (n=15,021)	D BX (n=1362)	CA BX (n=217)	CA CHI (n=417)	CA MIA (n=1012)	CU MIA (n=2241)	M BX (n=205)	M CHI (n=2403)	M SD (n=3775)	PR BX (n=1790)	PR CHI (n=761)	SA CHI (n=373)	SA MIA (n=465)	p- value ^a
Age (y) (%)														
18-44	59.3 (0.8)	63.3 (2.2)	66.1 (4.2)	66.8 (3.4)	63.6 (2.3)	45.2 (1.5)	85.2 (4.1)	72.7 (1.2)	62.7 (1.6)	52.7 (2.0)	55.6 (2.8)	63.9 (4.5)	56.3 (3.2)	
45-55	19.0 (0.5)	18.5 (1.4)	19.0 (3.4)	15.6 (2.3)	18.0 (1.5)	21.6 (0.9)	6.2 (1.6)	15.3 (1.0)	19.2 (1.0)	20.9 (1.3)	22.6 (2.4)	17.7 (2.5)	20.3 (2.0)	
55-74	21.6 (0.6)	18.2 (1.4)	14.9 (2.8)	17.5 (2.3)	18.4 (1.6)	33.2 (1.4)	8.6 (4.0)	12.0 (0.8)	18.1 (1.2)	26.4 (1.7)	21.8 (1.4)	18.4 (3.1)	23.4 (2.7)	< .001
Sex (%)														
Female	52.5 (0.6)	60.1 (1.9)	52.3 (4.5)	45.0 (3.1)	54.9 (2.2)	48.0 (1.1)	53.4 (4.2)	49.6 (1.3)	54.8 (1.3)	51.2 (1.8)	48.0 (2.1)	47.1 (3.4)	60.1 (2.5)	< .001
Immigrant Generation (%)														
First	77.2 (0.8)	82.7 (2.1)	81.3 (4.2)	93.9 (2.3)	96.3 (0.8)	93.5 (0.9)	94.8 (2.1)	84.9 (1.4)	66.6 (1.6)	49.2 (1.8)	46.1 (2.6)	95.1 (2.2)	96.7 (0.9)	< .001
Years lived in mainland United States (50 states and DC) (%)														
Born in mainland United States	28.3 (1.0)	24.1 (1.9)	25.1 (5.2)	36.7 (5.5)	41.5 (2.4)	50.2 (1.9)	34.0 (4.8)	25.4 (1.6)	21.7 (1.7)	6.6 (1.0)	3.9 (1.0)	40.3 (4.1)	46.3 (3.5)	
< 10 y	50.2 (0.8)	58.4 (2.2)	56.2 (5.0)	57.7 (5.1)	55.3 (2.2)	43.4 (1.7)	60.8 (4.9)	59.7 (1.5)	48.4 (1.6)	45.6 (1.9)	43.4 (2.6)	54.9 (4.3)	50.4 (3.4)	
≥10 y	21.5 (0.8)	17.5 (2.1)	18.7 (4.2)	5.7 (2.3)	3.3 (0.8)	6.4 (0.9)	5.2 (2.1)	14.8 (1.4)	29.9 (1.4)	47.8 (2.0)	52.7 (2.5)	4.9 (2.2)	3.3 (0.9)	< .001
Age of immigration^b (y)	26.9 (0.3)	25.6 (0.5)	24.9 (1.0)	24.8 (0.7)	27.7 (0.5)	35.3 (0.5)	22.9 (1.0)	23.0 (0.3)	25.1 (0.6)	16.8 (0.8)	17.1 (0.7)	26.5 (1.0)	31.0 (0.9)	< .001
Employment (%)														
Retired and not currently employed	8.2 (0.4)	8.3 (0.9)	6.6 (1.5)	5.2 (1.2)	3.2 (0.7)	11.5 (0.9)	0.8 (0.5)	3.1 (0.4)	5.5 (0.6)	16.9 (1.6)	14.2 (1.7)	3.4 (0.8)	4.0 (1.2)	
Not retired and not currently employed	41.1 (0.7)	42.6 (2.3)	34.9 (4.4)	31.5 (3.6)	38.3 (1.9)	46.7 (1.4)	37.7 (4.8)	31.5 (1.1)	41.7 (1.5)	45.2 (2.1)	38.8 (2.5)	23.1 (2.8)	31.5 (2.6)	
Part-time (≤35 hrs)	16.8 (0.5)	17.2 (1.5)	21.0 (3.5)	22.3 (2.8)	23.3 (1.6)	11.1 (0.7)	18.8 (3.3)	19.4 (1.1)	19.7 (1.1)	11.6 (1.2)	13.4 (2.2)	28.7 (2.7)	24.1 (2.1)	<
Full-time (35+hrs)	33.9 (0.7)	31.9 (2.0)	37.5 (4.5)	41.0 (2.8)	35.2 (1.7)	30.7 (1.2)	42.8 (4.6)	46.0 (1.3)	33.0 (1.5)	26.2 (1.8)	33.6 (2.6)	44.8 (2.7)	40.4 (2.9)	.001
Marital Status (%)														
Single	33.8 (0.7)	47.6 (1.9)	46.3 (4.4)	27.7 (3.3)	39.9 (1.8)	25.5 (1.3)	27.3 (4.7)	26.2 (1.5)	28.9 (1.4)	51.5 (1.9)	37.8 (2.7)	30.3 (4.6)	30.8 (2.7)	
Married/living with partner	49.9 (0.8)	36.7 (1.8)	42.2 (4.4)	58.7 (3.9)	44.0 (1.8)	52.3 (1.5)	65.8 (4.9)	63.1 (1.5)	57.6 (1.7)	29.9 (1.7)	41.5 (2.9)	56.7 (4.3)	46.5 (3.1)	
Separated/divorced/widowed	16.3 (0.5)	15.7 (1.3)	11.5 (2.2)	13.6 (2.0)	16.1 (1.3)	22.2 (1.2)	6.9 (1.8)	10.7 (0.6)	13.4 (1.0)	18.5 (1.6)	20.7 (2.1)	13.0 (2.1)	22.7 (2.9)	< .001
Yearly household income (%)														
< \$30k	61.7 (1.0)	66.4 (1.9)	62.5 (4.8)	66.2 (4.0)	69.3 (2.4)	64.5 (1.6)	75.0 (4.1)	65.4 (1.4)	52.6 (2.5)	64.6 (1.9)	55.4 (2.5)	59.2 (4.3)	61.3 (3.2)	< .001

Characteristic	Total (n=15,021)	D BX (n=1362)	CA BX (n=217)	CA CHI (n=417)	CA MIA (n=1012)	CU MIA (n=2241)	M BX (n=205)	M CHI (n=2403)	M SD (n=3775)	PR BX (n=1790)	PR CHI (n=761)	SA CHI (n=373)	SA MIA (n=465)	p- value ^a
≥ \$30k	31.8 (1.0)	25.9 (1.6)	27.7 (4.2)	30.5 (3.3)	21.4 (2.0)	23.8 (1.4)	17.7 (4.0)	29.7 (1.3)	44.8 (2.5)	29.8 (1.9)	41.6 (2.6)	38.1 (4.4)	33.1 (3.1)	
Not reported	6.4 (0.4)	7.7 (1.0)	9.8 (2.5)	3.3 (1.2)	9.3 (1.1)	11.7 (1.1)	7.4 (1.9)	4.9 (0.6)	2.6 (0.4)	5.6 (0.7)	3.0 (0.8)	2.7 (1.1)	5.7 (1.2)	
Education status (%)														
Less than high school	33.0 (0.7)	37.2 (1.9)	36.0 (4.2)	39.6 (4.2)	39.3 (2.0)	21.6 (1.1)	48.0 (5.3)	49.8 (1.7)	29.2 (1.6)	39.4 (2.1)	31.3 (2.3)	33.0 (4.3)	15.0 (2.3)	
High school or equivalent	28.4 (0.6)	23.5 (1.9)	30.0 (4.6)	19.8 (2.6)	26.3 (1.6)	30.0 (1.4)	36.1 (4.5)	30.5 (1.6)	28.5 (1.3)	26.8 (1.5)	30.4 (2.1)	27.3 (3.2)	25.3 (2.4)	
Greater than high school	38.6 (0.8)	39.3 (1.8)	34.1 (4.6)	40.6 (4.9)	34.3 (1.8)	48.5 (1.5)	16.0 (3.7)	19.8 (1.2)	42.3 (2.0)	33.7 (2.0)	38.2 (2.2)	39.7 (3.7)	60.0 (2.8)	< .001
Met 2008 Physical Activity Guidelines for Americans (%)														
No	57.6 (0.9)	42.9 (2.1)	40.5 (2.2)	53.1 (4.2)	63.1 (2.4)	78.5 (1.5)	28.4 (5.1)	54.9 (1.3)	58.9 (1.6)	43.3 (2.0)	63.0 (3.0)	50.6 (3.3)	65.0 (2.9)	< .001
Dietary Supplement Use (%)														
No	59.3 (0.7)	63.8 (2.1)	66.5 (4.1)	56.6 (4.0)	57.5 (2.1)	57.6 (1.5)	69.6 (5.2)	67.7 (1.3)	52.2 (1.6)	65.5 (1.8)	61.4 (2.0)	62.6 (4.7)	46.7 (2.6)	< .001
Body Mass Index^c (%)														
Underweight	1.2 (0.1)	1.4 (0.4)	1.6 (1.3)	0.1 (0.1)	1.1 (0.4)	1.5 (0.3)		0.8 (0.3)	1.4 (0.4)	0.9 (0.3)	0.4 (0.3)	2.5 (1.7)	0.6 (0.4)	
Normal	21.6 (0.5)	20.9 (1.5)	21.3 (3.8)	24.2 (2.6)	22.5 (2.0)	23.2 (1.1)	18.0 (3.2)	20.6 (1.0)	21.7 (1.1)	17.5 (1.5)	22.2 (2.5)	25.0 (3.4)	32.2 (3.4)	
Overweight	37.4 (0.6)	36.5 (1.9)	36.9 (4.3)	43.4 (3.3)	38.2 (2.3)	37.4 (1.2)	39.8 (5.0)	37.9 (1.3)	39.4 (1.6)	33.3 (1.8)	31.0 (2.1)	42.7 (2.8)	38.8 (3.1)	
Obesity	39.8 (0.7)	41.2 (2.0)	40.2 (4.6)	32.3 (2.7)	38.2 (1.6)	37.9 (1.3)	42.3 (5.3)	40.7 (1.5)	37.5 (1.7)	48.3 (2.0)	46.4 (2.5)	29.8 (3.3)	28.4 (2.7)	< .001
Energy intake^b (kcal/day)														
	1910.1 (10.8)	1620.0 (28.6)	1634.6 (70.4)	1929.6 (44.9)	1910.7 (37.0)	2093.3 (19.9)	1652.1 (61.6)	2000.4 (17.5)	1937.8 (19.5)	1756.1 (24.7)	1982.8 (37.3)	1919.1 (52.5)	2028.8 (39.0)	< .001

^aP-values are from Chi-square test of independence for categorical characteristics and from ANOVA for continuous characteristics. P-values were adjusted for multiple comparisons using the False Discovery Rate method.

^bValues are survey weighted means with standard errors indicated in parenthesis.

^cWeight categories were as follows: Underweight (<18.5 kg/m²), Normal (18.5–24.9 kg/m²), Overweight (25–29.9 kg/m²), and Obesity (≥30 kg/m²).

ABBREVIATIONS: AHEI: Alternate Healthy Eating Index; BX: BRONX; CA: Central American; CU: Cuban; CHI: Chicago; D: Dominican; M: Mexican; MIA: Miami; PR: Puerto Rican; SA: South American; SD: San Diego.

Table 2. Factor congruence coefficients^{a,b} between pairs of ethnic-background-site-specific dietary patterns. Hispanic Community Health Study/Study of Latinos, 2008-2011.

	D BX	CA BX	CA CHI	CA MIA	Cu MIA	M BX	M CHI	M SD	PR BX	PR CHI	SA CHI	SA MIA
D BX	1	0.91	0.75	0.91	0.83	0.76	0.78	0.82	0.91	0.49	<i>0.95</i>	0.82
CA BX		1	0.74	0.91	0.83	0.81	0.69	0.78	0.88	0.51	0.88	0.76
CA CHI			1	0.84	0.86	0.84	0.67	0.51	0.90	0.01	0.75	0.92
CA MIA				1	0.93	0.92	0.87	0.83	0.91	0.35	0.93	0.85
Cu MIA					1	0.83	0.75	0.64	0.87	0.25	0.88	0.89
M BX						1	0.85	0.76	0.86	0.14	0.78	0.79
M CHI							1	0.86	0.71	0.36	0.84	0.69
M SD								1	0.68	0.61	0.84	0.51
PR BX									1	0.18	0.85	0.93
PR CHI										1	0.50	0.05
SA CHI											1	0.79
SA MIA												1

^aThe congruence coefficient matrix is symmetric with respect to the main diagonal (all 1's); 66 (i.e., number of ethnic-background-site-specific categories*(number of ethnic-background-site-specific categories -1)/2) is the total number of (unique) congruence coefficients between the ethnic background-site-specific dietary patterns.

^bCongruence coefficients range between 0 and 1 in absolute value. Values between 0.85 and 0.94 indicate fair similarity between corresponding dietary patterns and were shown in bold typeface in the upper triangular matrix; values ≥ 0.95 indicate equivalence of the corresponding dietary patterns and were shown in bold and italics typeface in the upper triangular matrix. Of the 24 congruence coefficients indicating fair similarity/equivalence between pairs of dietary patterns (in bold typeface in the upper triangular matrix), 11 were underlined in the upper triangular matrix to target similarities between DPs belonging to the same overarching structures.

ABBREVIATIONS: BX: BRONX; CA: Central American; Cu: Cuban; CHI: Chicago; D: Dominican; M: Mexican; MIA: Miami; PR: Puerto Rican; SA: South American; SD: San Diego.

Table 3. Deviation (%) of the adjusted food group mean for individuals in the top quintile of each shared dietary pattern, relative to the overall adjusted mean. Hispanic Community Health Study/Study of Latinos, 2008-2011. ^{a,b}

Food group	<i>Shared dietary patterns</i>			
	<i>Plant-based foods</i>	<i>Processed foods</i>	<i>Dairy products</i>	<i>Seafood</i>
Fruit - Citrus	163	88	120	115
Fruit - Others ^c	159	86	98	108
Vegetables - Dark green	159	90	104	152
Vegetables - Orange	153	83	89	121
Vegetables - Tomato	149	121	109	102
White Potatoes	130	118	100	123
Vegetables - Starchy	140	164	102	95
Vegetables - Beans	223	118	78	89
Vegetables - Others ^d	149	133	99	130
Refined Grain	127	153	122	101
Whole Grain	149	90	96	91
Red Meat	124	169	111	69
Deli Meat	107	185	141	101
Poultry	127	135	92	151
Fish	133	120	103	405
Eggs	98	148	127	126
Nuts and Seeds	235	203	59	102
Milk	135	76	215	100
Cheese	100	147	214	90
Yogurt	125	68	183	110
Milk-based Dessert	102	169	246	100
Sugar	112	96	123	99
Dessert	90	139	138	99
Sugar Sweetened Beverage	110	158	112	95
Diet Beverage	110	96	89	106
Alcohol	173	153	101	152
Snack - Overall	135	180	102	99
Regular fat	125	181	99	113
Reduced fat	118	184	120	116

^a Mean intakes of food groups (servings/day) were adjusted for age, sex, body mass index, total energy intake, and ethnic-background-site category.

^b If the relative consumption of a food group exceeded 100%, it indicates that individuals belonging to the top quintile category of factor score were characterized by a relatively high consumption of that food group, compared with the reference HCHS/SOL overall mean, and vice versa when the relative intake is below 100%. Percentages below 60% or above 140% were indicated in bold typeface.

^cFruit – Others group included fruit juice (excluding citrus juice), fruit (excluding citrus fruit), avocado and similar, fried fruits, and fruit-based savory snack.

^dVegetables – Others group included other vegetables, fried vegetables, vegetable juice, and pickled foods.

Table 4. Deviation (%) of the adjusted food group mean for individuals in the top tertile of each ethnic-background-site-specific dietary pattern, relative to the ethnic-background-site-specific adjusted mean. Hispanic Community Health Study/Study of Latinos, 2008-2011.^{a,b}

<i>Ethnic background-site-specific (EBS) dietary patterns</i>												
Food group	D BX	CA BX	CA CHI	CA MIA	Cu MIA	M BX	M CHI	M SD	PR BX	PR CHI	SA CHI	SA MIA
Fruit - Citrus	82	79	108	96	93	175	94	64	66	135	76	134
Fruit - Other^c	56	60	74	83	84	117	77	71	96	151	77	89
Vegetables - Dark green	66	42	129	72	91	73	91	60	71	205	65	125
Vegetables - Orange	81	80	116	105	93	103	105	69	81	147	83	127
Vegetables - Tomato	83	80	106	105	110	124	94	89	79	110	91	97
White Potatoes	88	113	157	119	123	128	82	91	105	101	91	110
Vegetables - Starchy	86	106	80	82	90	72	114	91	90	126	94	104
Vegetables - Beans	100	107	82	62	81	64	43	44	85	98	56	59
Vegetables – Other^d	81	78	112	111	106	123	101	82	88	118	81	101
Refined Grain	121	99	93	106	102	94	119	109	93	65	116	98
Whole Grain	79	84	108	93	94	44	73	86	81	115	81	100
Red Meat	157	134	150	149	160	138	141	138	152	94	137	132
Deli - Meat	124	67	92	114	113	93	114	100	113	74	145	119
Poultry	156	178	190	144	142	142	150	129	153	105	143	167
Fish	80	105	106	64	77	137	86	75	55	72	77	63
Eggs	111	124	93	119	126	94	98	82	125	104	95	150
Nuts and Seeds	18	16	20	53	47	5	55	62	65	105	28	36
Milk	76	78	66	95	93	96	90	103	81	105	91	79

Ethnic background-site-specific (EBS) dietary patterns

Food group	D BX	CA BX	CA CHI	CA MIA	Cu MIA	M BX	M CHI	M SD	PR BX	PR CHI	SA CHI	SA MIA
Cheese	84	58	57	62	76	70	80	69	87	84	47	93
Yogurt	67	0	90	139	80	109	69	93	46	151	27	85
Milk-based Dessert	65	47	59	43	57	80	177	67	100	150	75	80
Sugar Dessert	86	87	87	74	82	78	83	87	93	108	82	98
Sugar Sweetened Beverage	88	67	100	42	68	83	135	91	67	94	64	97
Diet Beverage	120	113	113	111	97	153	127	128	106	76	114	104
Alcohol	103	102	95	88	96	100	107	82	90	100	100	100
Snack - Overall	116	90	228	196	143	282	173	150	160	90	158	106
Regular fat	63	71	65	65	65	58	87	76	64	88	71	76
Reduced fat	103	82	106	94	105	89	102	76	91	97	110	101
	100	116	87	105	101	83	129	104	94	102	111	103

^a Mean intakes of selected food groups (servings/day) were adjusted for age, sex, body mass index, and total energy intake.

^b If the relative consumption of a food group exceeded 100%, it indicates that individuals belonging to top tertile category of factor score were characterized by a relatively high consumption of that food group, compared with the ethnic background site specific overall mean, and vice versa when the relative intake is below 100%. Percentages below 60% or above 140% were indicated in bold typeface.

^c Fruit – Others group included fruit juice (excluding citrus juice), fruit (excluding citrus fruit), avocado and similar, fried fruits, and fruit-based savory snack.

^d Vegetables – Others group included other vegetables, fried vegetables, vegetable juice, and pickled foods.

Table 5. Socio-demographic and lifestyle characteristics (weighted percentages and standard error in parenthesis) for participants by shared quintile-based (Q1, Q5) categories. Hispanic Community Health Study/Study of Latinos, 2008-2011.

	<i>Plant-based foods</i>			<i>Processed foods</i>			<i>Dairy products</i>			<i>Seafood</i>		
	Q1	Q5	p-value ^a	Q1	Q5	p-value ^a	Q1	Q5	p-value ^a	Q1	Q5	p-value ^a
Age (y) (%)												
18-44	38.3 (1.9)	42.4 (2.2)		30.8 (1.6)	55.9 (2.3)		35.0 (1.8)	41.5 (2.2)		39.6 (2.1)	39.8 (1.7)	
45-55	25.8 (1.5)	24.6 (1.3)	< 0.001	23.8 (1.4)	26.8 (1.7)	< 0.001	28.2 (1.5)	25.0 (1.6)	0.002	27.0 (1.5)	26.5 (1.5)	< 0.001
55-74	36.0 (1.7)	33.0 (2.0)		45.3 (1.7)	17.2 (1.7)		36.7 (1.6)	33.5 (2.0)		33.4 (1.9)	33.7 (1.7)	
Sex (%)												
Female	69.9 (1.6)	41.9 (1.9)	< 0.001	79.5 (1.4)	24.1 (1.5)	< 0.001	56.2 (1.4)	49.1 (2.0)	< 0.001	55.8 (1.9)	47.8 (1.7)	< 0.001
Immigrant Generation (%)												
First	98.9 (0.3)	97.9 (0.7)	< 0.001	98.5 (0.4)	97.8 (0.7)	< 0.001	98.5 (0.4)	98.0 (0.7)	< 0.001	97.3 (0.8)	97.9 (0.7)	< 0.001
Years lived in mainland United States (50 states and DC) (%)												
Born in mainland United States	28.2 (1.4)	27.8 (1.7)		25.7 (1.3)	25.9 (1.4)		28.2 (1.7)	27.5 (1.6)		25.3 (1.4)	28.6 (1.5)	
< 10 y	45.9 (1.3)	53.6 (1.7)	< 0.001	61.6 (1.3)	42.1 (1.4)	< 0.001	53.7 (1.6)	48.4 (1.5)	0.003	46.6 (1.6)	54.5 (1.4)	< 0.001
≥10 y	25.9 (1.4)	18.6 (1.3)		12.7 (0.9)	32.0 (1.6)		18.1 (1.3)	24.1 (1.4)		28.1 (1.5)	16.9 (1.1)	
Age of immigration^b (y)	30.6 (1.6)	30.1 (1.6)	0.007	31.7 (0.6)	26.5 (0.6)	< 0.001	31.1 (0.6)	30.7 (0.7)	0.748	30.5 (0.6)	30.2 (0.6)	0.016
Employment (%)												
Retired and not currently employed	13.2 (1.2)	11.4 (1.2)		18.4 (1.3)	4.3 (0.8)		12.2 (1.1)	12.9 (1.3)		11.7 (1.2)	12.1 (1.2)	
Not retired and not currently employed	40.9 (1.7)	35.2 (1.9)	< 0.001	41.8 (1.7)	29.1 (1.8)	< 0.001	38.4 (1.6)	36.3 (2.2)	0.47	38.6 (1.8)	33.9 (1.5)	0.025
Part-time (≤35 hrs)	15.6 (1.4)	15.8 (1.3)		15.8 (1.1)	15.8 (1.5)		16.1 (1.1)	15.2 (1.3)		15.1 (1.4)	16.9 (1.2)	
Full-time (35+hrs)	30.3 (1.6)	37.6 (1.9)		24.1 (1.4)	50.9 (2.3)		33.3 (1.7)	35.7 (1.9)		34.6 (1.7)	37.0 (2.0)	
Marital Status (%)												
Single	17.9 (1.3)	16.0 (1.4)		16.6 (1.4)	22.2 (1.8)		18.5 (1.3)	17.4 (1.5)		19.2 (1.6)	17.1 (1.2)	
Married/living with partner	57.4 (1.8)	64.3 (1.8)	0.011	58.2 (1.7)	62.0 (2.2)	< 0.001	57.6 (1.7)	63.6 (2.1)	0.028	60.2 (1.9)	61.2 (1.9)	< 0.001

	<i>Plant-based foods</i>			<i>Processed foods</i>			<i>Dairy products</i>			<i>Seafood</i>		
Separated/divorced/widowed	24.6 (1.5)	19.7 (1.4)		25.1 (1.4)	15.8 (1.5)		23.9 (1.3)	18.9 (1.4)		20.6 (1.4)	21.7 (1.5)	
Yearly household income (%)												
< \$30k	71.8 (1.8)	64.8 (2.1)		69.6 (1.8)	62.9 (2.1)		67.9 (1.8)	63.4 (2.6)		60.7 (2.0)	64.1 (2.1)	
≥ \$30k	19.9 (1.6)	31.8 (2.1)	< 0.001	24.4 (1.6)	32.9 (2.0)	< 0.001	25.6 (1.8)	31.0 (2.7)	0.012	32.6 (1.9)	31.7 (2.3)	< 0.001
Missing	8.3 (1.0)	3.3 (0.7)		6.0 (0.7)	4.2 (0.8)		6.5 (0.9)	5.6 (1.1)		6.6 (1.0)	4.1 (0.7)	
Education status (%)												
Less than high school education	42.2 (1.7)	36.1 (2.0)		41.6 (1.8)	36.3 (2.3)		44.9 (1.7)	35.0 (2.1)		39.6 (1.9)	36.3 (1.8)	
High school education or equivalent	25.4 (1.5)	25.6 (1.5)	0.027	21.7 (1.2)	27.4 (1.9)	0.003	22.9 (1.3)	26.1 (1.9)	< 0.001	27.4 (2.0)	25.9 (1.8)	0.783
Greater than high school education	32.4 (1.7)	38.3 (1.9)		36.6 (1.8)	36.3 (2.2)		32.2 (1.6)	38.9 (2.2)		33.0 (1.8)	37.8 (2.1)	
Met 2008 Physical Activity Guidelines for Americans (%)												
No	69.6 (1.6)	56.9 (2.1)	< 0.001	66.8 (1.7)	52.4 (2.2)	< 0.001	64.8 (1.7)	64.1 (2.1)	0.07	66.0 (1.9)	61.2 (2.1)	0.706
Supplemental Use (%)												
No	59.3 (1.8)	47.6 (2.0)	< 0.001	45.0 (1.6)	61.6 (2.1)	< 0.001	57.5 (1.7)	46.6 (2.1)	< 0.001	50.2 (2.0)	50.3 (1.6)	0.053
Body Mass Index^c (%)												
Underweight	1.2 (0.3)	1.1 (0.3)		0.4 (0.2)	1.9 (0.5)		1.1 (0.2)	1.3 (0.4)		1.6 (0.4)	0.9 (0.4)	
Normal	20.0 (1.2)	23.0 (1.2)	0.086	18.6 (1.0)	23.6 (1.1)	0.002	19.9 (1.0)	24.6 (1.3)	0.349	23.0 (1.2)	18.8 (1.0)	0.079
Overweight	35.6 (1.3)	39.5 (1.4)		37.6 (1.3)	34.9 (1.4)		37.9 (1.4)	35.6 (1.4)		35.1 (1.4)	39.0 (1.5)	
Obesity	43.3 (1.6)	36.6 (1.5)		43.3 (1.4)	39.6 (1.5)		41.0 (1.4)	38.5 (1.6)		40.3 (1.4)	41.3 (1.5)	

	<i>Plant-based foods</i>			<i>Processed foods</i>			<i>Dairy products</i>			<i>Seafood</i>		
Energy intake^b (kcal/day)	1385 (15.2)	2443 (26.7)	< 0.001	1286 (13.3)	2621 (24.5)	< 0.001	1615 (19.9)	2259 (22.6)	< 0.001	1951 (21.5)	1941 (23.0)	0.047

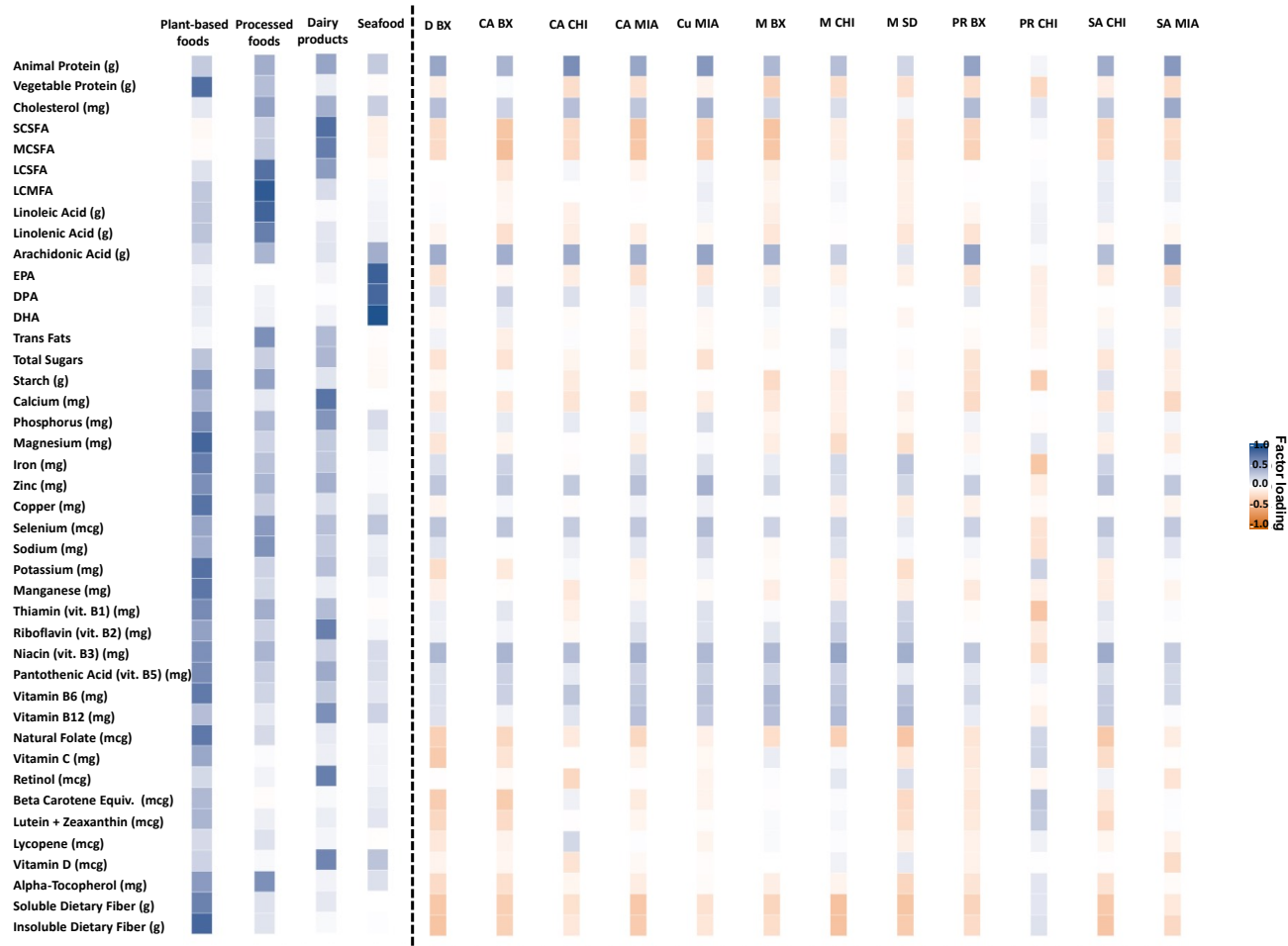
^aP-values are from Chi-square test of independence for categorical characteristics and from ANOVA for continuous characteristics. P-values were adjusted for multiple comparisons using the False Discovery Rate method. The adjustment was carried out considering all dietary patterns simultaneously.

^bValues are weighted means with standard errors indicated in parenthesis.

^cWeight categories were as follows: Underweight (<18.5 kg/m²), Normal (18.5–24.9 kg/m²), Overweight (25–29.9 kg/m²), and Obesity (≥30 kg/m²).

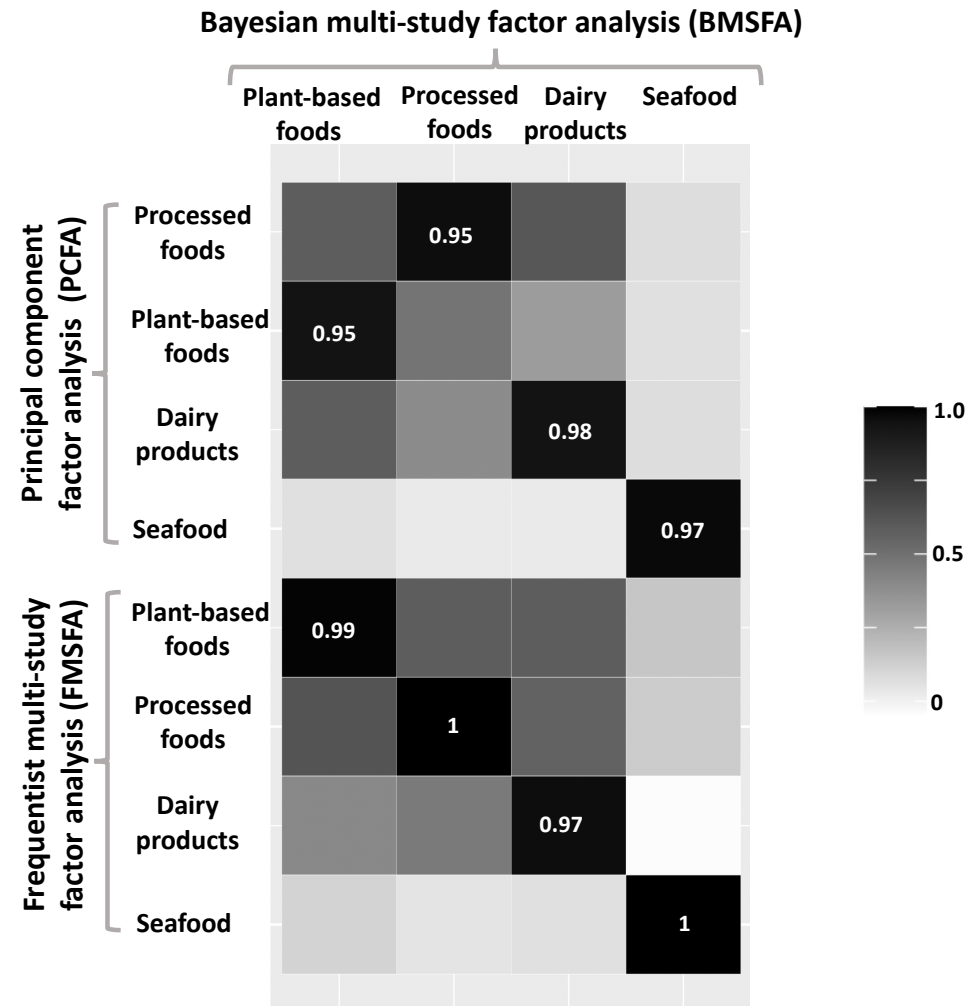
ABBREVIATIONS: BX: BRONX; CA: Central American; Cu: Cuban; CHI: Chicago; D: Dominican; M: Mexican; MIA: Miami; PR: Puerto Rican; SA: South American; SD: San Diego.

Figure 1. Heatmap of the estimated factor-loading matrix for the shared and ethnic-background-site specific dietary patterns identified with the BMFA. Dashed line indicates the division between the shared and ethnic-background-site-specific dietary patterns. Hispanic Community Health Study/Study of Latinos, 2008-2011.



ABBREVIATIONS: BMSFA: Bayesian multi-study factor analysis; BX: BRONX; CA: Central American; Cu: Cuban; CHI: Chicago; D: Dominican; DHA: docosahexaenoic acid; DPA: docosapentaenoic acid; EPA: eicosapentaenoic acid; M: Mexican; MCSFA: medium-chain saturated fatty acids; LCMFA: long-chain monounsaturated fatty acids; LCSFA: long-chain saturated fatty acids; MIA: Miami; PR: Puerto Rican; SA: South American; SCSFA: short-chain saturated fatty acids; SD: San Diego.

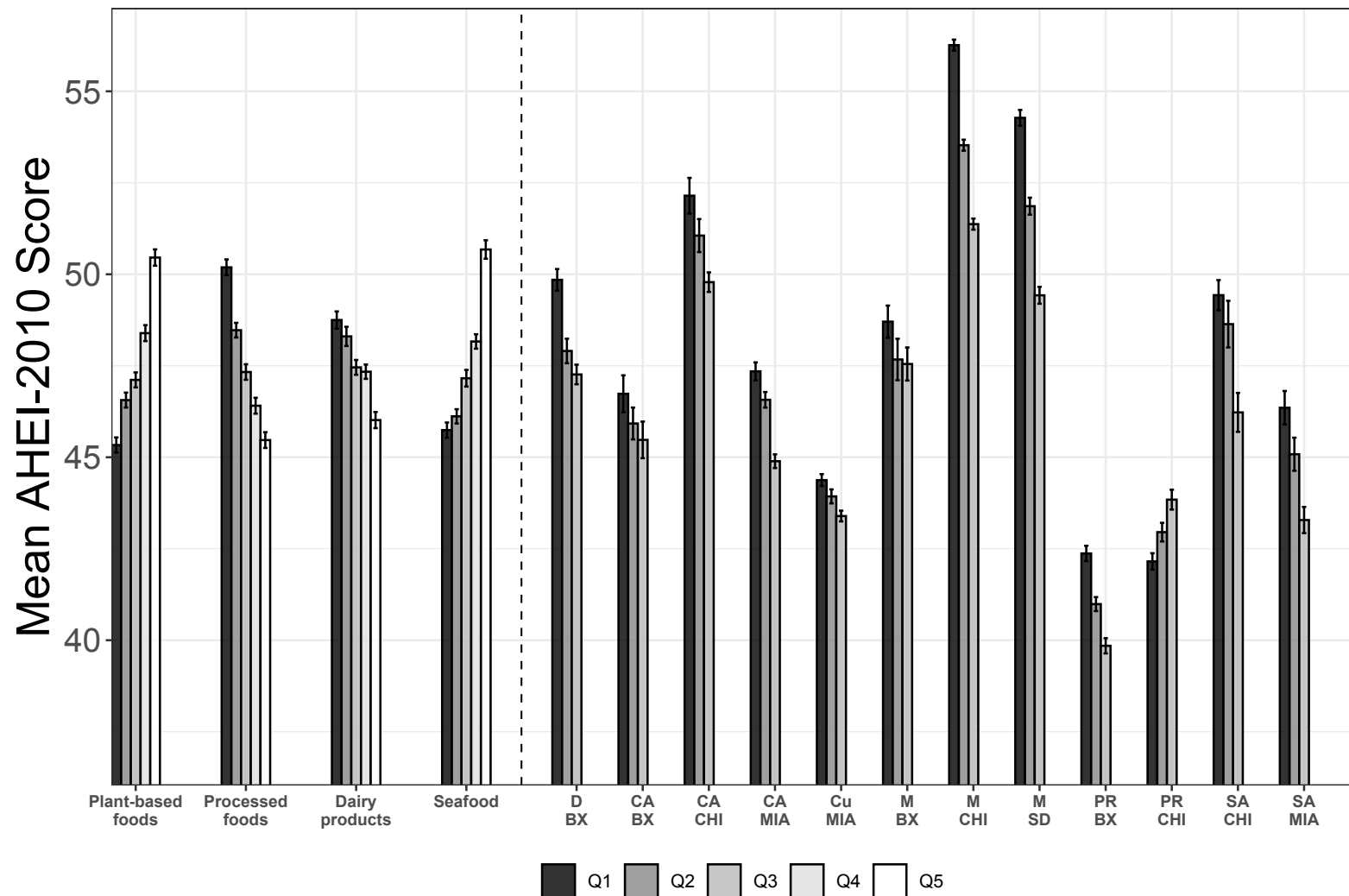
Figure 2. Heatmap of the factor congruence coefficients^{a,b} between the shared dietary patterns identified with the BMSFA and FMSFA and the overall-sample dietary patterns from PCFA. Hispanic Community Health Study/Study of Latinos, 2008-2011.



^aCongruence coefficients range between 0 and 1 (in absolute value), with values between 0.85 and 0.94 indicating fairly similarity and values ≥ 0.95 indicating equivalence of the corresponding dietary patterns.

^bDietary patterns were ordered in terms of proportion of total variance explained in each solution (see **Supplemental Table 5**).

Figure 3. Mean AHEI-2010 scores (standard errors) by quintiles of shared dietary patterns and by tertiles of ethnic-background-site-specific dietary patterns^{a,b,c}. Hispanic Community Health Study/Study of Latinos, 2008-2011.



^a Weighted mean AHEI-2010 scores were adjusted for age, sex, body mass index, and total energy intake; models for the shared dietary patterns were further adjusted for ethnic-background-site combinations and for the other shared dietary patterns (quintiles). Asterisks indicate significant p for trend: ***<0.001, **<0.01, and *<0.05, respectively.

^b The AHEI-2010 measures overall diet quality in terms of adherence to the Dietary Guidelines for Americans 2010; compared to the healthy Eating Index 2010, it also incorporates additional components that focus on foods and nutrients to predict the risk of chronic disease. The total AHEI-2010 score ranges from 0 to 110, with higher scores indicating a healthier diet.

^c The same color scale was adopted for the first three quintile-based categories and for the three tertile-based categories.

ABBREVIATIONS: AHEI: Alternative Healthy Eating Index; BX: BRONX; CA: Central American; Cu: Cuban; CHI: Chicago; D: Dominican; M: Mexican; MIA: Miami; PR: Puerto Rican; SA: South American; SD: San Diego.

References

1. Zhao J, Li Z, Gao Q, Zhao H, Chen S, Huang L, Wang W, Wang T. A review of statistical methods for dietary pattern analysis. *Nutr J* 2021;20(1):37. doi: 10.1186/s12937-021-00692-7.
2. Gleason PM, Boushey CJ, Harris JE, Zoellner J. Publishing nutrition research: a review of multivariate techniques--part 3: data reduction methods. *J Acad Nutr Diet* 2015;115(7):1072-82. doi: 10.1016/j.jand.2015.03.011.
3. Jannasch F, Riordan F, Andersen LF, Schulze MB. Exploratory dietary patterns: a systematic review of methods applied in pan-European studies and of validation studies. *Br J Nutr* 2018;120(6):601-11. doi: 10.1017/S0007114518001800.
4. Schulze MB, Martinez-Gonzalez MA, Fung TT, Lichtenstein AH, Forouhi NG. Food based dietary patterns and chronic disease prevention. *BMJ* 2018;361:k2396. doi: 10.1136/bmj.k2396.
5. Edefonti V, De Vito R, Salvatori A, Bravi F, Patel L, Dalmartello M, Ferraroni M. Reproducibility of A Posteriori Dietary Patterns across Time and Studies: A Scoping Review. *Adv Nutr* 2020;11(5):1255-81. doi: 10.1093/advances/nmaa032.
6. Krebs-Smith SM, Subar AF, Reedy J. Examining Dietary Patterns in Relation to Chronic Disease: Matching Measures and Methods to Questions of Interest. *Circulation* 2015;132(9):790-3. doi: 10.1161/CIRCULATIONAHA.115.018010.
7. Liese AD, Krebs-Smith SM, Subar AF, George SM, Harmon BE, Neuhauser ML, Boushey CJ, Schap TE, Reedy J. The Dietary Patterns Methods Project: synthesis of findings across cohorts and relevance to dietary guidance. *J Nutr* 2015;145(3):393-402. doi: 10.3945/jn.114.205336.
8. Judd SE, Letter AJ, Shikany JM, Roth DL, Newby PK. Dietary Patterns Derived Using Exploratory and Confirmatory Factor Analysis are Stable and Generalizable Across Race, Region, and Gender Subgroups in the REGARDS Study. *Front Nutr* 2014;1:29. doi: 10.3389/fnut.2014.00029.
9. Tucker KL. Dietary patterns, approaches, and multicultural perspective. *Appl Physiol Nutr Metab* 2010;35(2):211-8. doi: 10.1139/H10-010.
10. de Souza RJ, Zulyniak MA, Desai D, Shaikh MR, Campbell NC, Lefebvre DL, Gupta M, Wilson J, Wahi G, Atkinson SA, et al. Harmonization of Food-Frequency Questionnaires and Dietary Pattern Analysis in 4 Ethnically Diverse Birth Cohorts. *J Nutr* 2016;146(11):2343-50. doi: 10.3945/jn.116.236729.
11. Balder HF, Virtanen M, Brants HA, Krogh V, Dixon LB, Tan F, Mannisto S, Bellocco R, Pietinen P, Wolk A, et al. Common and country-specific dietary patterns in four European cohort studies. *J Nutr* 2003;133(12):4246-51. doi: 10.1093/jn/133.12.4246.
12. Mannisto S, Dixon LB, Balder HF, Virtanen MJ, Krogh V, Khani BR, Berrino F, van den Brandt PA, Hartman AM, Pietinen P, et al. Dietary patterns and breast cancer risk: results from three cohort studies in the DIETSCAN project. *Cancer Causes Control* 2005;16(6):725-33. doi: 10.1007/s10552-005-1763-7.
13. Castello A, Lope V, Vioque J, Santamarina C, Pedraz-Pingarron C, Abad S, Ederra M, Salas-Trejo D, Vidal C, Sanchez-Contador C, et al. Reproducibility of data-driven dietary patterns in two groups of adult Spanish women from different studies. *Br J Nutr* 2016;116(4):734-42. doi: 10.1017/S000711451600252X.
14. Castello A, Buijsse B, Martin M, Ruiz A, Casas AM, Baena-Canada JM, Pastor-Barriuso R, Antolin S, Ramos M, Munoz M, et al. Evaluating the Applicability of Data-Driven Dietary Patterns to Independent Samples with a Focus on Measurement Tools for Pattern Similarity. *J Acad Nutr Diet* 2016;116(12):1914-24 e6. doi: 10.1016/j.jand.2016.05.008.
15. Moskal A, Pisa PT, Ferrari P, Byrnes G, Freisling H, Boutron-Ruault MC, Cadeau C, Nailler L, Wendt A, Kuhn T, et al. Nutrient patterns and their food sources in an

- International Study Setting: report from the EPIC study. *PLoS One* 2014;9(6):e98647. doi: 10.1371/journal.pone.0098647.
16. Edefonti V, Hashibe M, Ambrogi F, Parpinel M, Bravi F, Talamini R, Levi F, Yu G, Morgenstern H, Kelsey K, et al. Nutrient-based dietary patterns and the risk of head and neck cancer: a pooled analysis in the International Head and Neck Cancer Epidemiology consortium. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2012;23(7):1869-80. doi: 10.1093/annonc/mdr548.
 17. De Vito R, Lee YCA, Parpinel M, Serraino D, Olshan AF, Zevallos JP, Levi F, Zhang ZF, Morgenstern H, Garavello W, et al. Shared and Study-specific Dietary Patterns and Head and Neck Cancer Risk in an International Consortium. *Epidemiology* 2019;30(1):93-102. doi: 10.1097/EDE.0000000000000902.
 18. Murakami K, Shinozaki N, Fujiwara A, Yuan X, Hashimoto A, Fujihashi H, Wang HC, Livingstone MBE, Sasaki S. A Systematic Review of Principal Component Analysis-Derived Dietary Patterns in Japanese Adults: Are Major Dietary Patterns Reproducible Within a Country? *Adv Nutr* 2019;10(2):237-49. doi: 10.1093/advances/nmy079.
 19. Bakolis I, Hooper R, Bachert C, Lange B, Haahtela T, Keil T, Hofmaier S, Fokkens W, Rymarczyk B, Janson C, et al. Dietary patterns and respiratory health in adults from nine European countries-Evidence from the GA(2) LEN study. *Clin Exp Allergy* 2018;48(11):1474-82. doi: 10.1111/cea.13243.
 20. De Vito R, Bellio R, Trippa L, Parmigiani G. Multi-study factor analysis. *Biometrics* 2019;75(1):337-46. doi: 10.1111/biom.12974.
 21. De Vito R, Bellio R, Trippa L, Parmigiani G. Bayesian multi-study factor analysis for high-throughput biological data. *The Annals of Applied Statistics* 2021;15(4):1723-41. doi: DOI: 10.1214/21-AOAS1456.
 22. Sorlie PD, Aviles-Santa LM, Wassertheil-Smoller S, Kaplan RC, Daviglius ML, Giachello AL, Schneiderman N, Raji L, Talavera G, Allison M, et al. Design and implementation of the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol* 2010;20(8):629-41. doi: 10.1016/j.annepidem.2010.03.015.
 23. Maldonado LE, Adair LS, Sotres-Alvarez D, Mattei J, Mossavar-Rahmani Y, Perreira KM, Daviglius ML, Van Horn LV, Gallo LC, Isasi CR, et al. Dietary Patterns and Years Living in the United States by Hispanic/Latino Heritage in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). *J Nutr* 2021;151(9):2749-59. doi: 10.1093/jn/nxab165.
 24. Stephenson BJK, Sotres-Alvarez D, Siega-Riz AM, Mossavar-Rahmani Y, Daviglius ML, Van Horn L, Herring AH, Cai J. Empirically Derived Dietary Patterns Using Robust Profile Clustering in the Hispanic Community Health Study/Study of Latinos. *J Nutr* 2020;150(10):2825-34. doi: 10.1093/jn/nxaa208.
 25. Lavange LM, Kalsbeek WD, Sorlie PD, Aviles-Santa LM, Kaplan RC, Barnhart J, Liu K, Giachello A, Lee DJ, Ryan J, et al. Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol* 2010;20(8):642-9. doi: 10.1016/j.annepidem.2010.05.006.
 26. Nutrition Data System for Research (NDS-R) version 2011. NDS- R software, 1998–1999 Regents of the University of Minnesota. Minneapolis (MN): Nutrition Coordinating Center, University of Minnesota.
 27. Pett MA, Lackey NR, Sullivan JJ. Making sense of factor analysis: the use of factor analysis for instrument development in health care research: Thousand Oaks, CA: Sage, 2003.
 28. Islam MA, Amin MN, Siddiqui SA, Hossain MP, Sultana F, Kabir MR. Trans fatty acids and lipid profile: A serious risk factor to cardiovascular disease, cancer and diabetes. *Diabetes Metab Syndr* 2019;13(2):1643-7. doi: 10.1016/j.dsx.2019.03.033.

29. Lichtenstein AH. Dietary fat and cardiovascular disease risk: quantity or quality? *J Womens Health (Larchmt)* 2003;12(2):109-14. doi: 10.1089/154099903321576493.
30. Johnson RA, Wichern DW. *Applied multivariate statistical analysis*. 5th ed. Upper Saddle River, NJ: Prentice Hall, 2002.
31. DiStefano C, Zhu M, Mindrila D. Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation* 2009;14(20):1-11.
32. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* 1995;57(1):289-300.
33. Chiuve SE, Fung TT, Rimm EB, Hu FB, McCullough ML, Wang M, Stampfer MJ, Willett WC. Alternative dietary indices both strongly predict risk of chronic disease. *J Nutr* 2012;142(6):1009-18. doi: 10.3945/jn.111.157222.
34. Lumley T. *Complex surveys: a guide to analysis using R*: John Wiley & Sons, 2011.
35. R Development Core Team. *R: A Language and Environment for Statistical Computing 2022* Vienna, Austria: R Foundation for Statistical Computing, Available at: [http:// www.R-project.org](http://www.R-project.org) (23 June 2022, date last accessed)
36. Giner G, Smyth GK. *Statmod: probability calculations for the Inverse Gaussian distribution*. arXiv preprint arXiv:160306687 2016.
37. Revelle W. *Psych: procedures for psychological, psychometric, and personality research*. Version 2.1.6. 2021, Available at: <https://CRAN.R-project.org/package=psych> (23 June 2022, date last accessed), Evanston, IL: Northwestern University.
38. Raïche G, Walls TA, Magis D, Riopel M, Blais J-G. Non-graphical solutions for Cattell's scree test. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences* 2013;9(1):23.
39. Wickham H. *Ggplot2: elegant graphics for data analysis*: Springer-Verlag New York}, 2016.
40. Lumley T. *Survey: analysis of complex survey samples*. R package version 2020;4.0.
41. De Vito R, Bellio R. *MSFA*. GitHub repository <https://github.com/rdevito/MSFA> 2020.
42. Dixon LB, Sundquist J, Winkleby M. Differences in energy, nutrient, and food intakes in a US sample of Mexican-American women and men: findings from the Third National Health and Nutrition Examination Survey, 1988-1994. *Am J Epidemiol* 2000;152(6):548-57. doi: 10.1093/aje/152.6.548.
43. Siega-Riz AM, Sotres-Alvarez D, Ayala GX, Ginsberg M, Himes JH, Liu K, Loria CM, Mossavar-Rahmani Y, Rock CL, Rodriguez B, et al. Food-group and nutrient-density intakes by Hispanic and Latino backgrounds in the Hispanic Community Health Study/Study of Latinos. *Am J Clin Nutr* 2014;99(6):1487-98. doi: 10.3945/ajcn.113.082685.
44. Bermudez OI, Falcon LM, Tucker KL. Intake and food sources of macronutrients among older Hispanic adults: association with ethnicity, acculturation, and length of residence in the United States. *J Am Diet Assoc* 2000;100(6):665-73. doi: 10.1016/s0002-8223(00)00195-4.
45. Ellithorpe ME, Takahashi B, Alumit Zeldes G, Dorrance-Hall E, Chavez M, Plasencia J. Family and Cultural Perceptions About Meat Consumption among Hispanic/Latino and White Adults in the United States. *Ecol Food Nutr* 2021:1-14. doi: 10.1080/03670244.2021.2018309.
46. Sanchez-Flack JC, Anderson CAM, Arredondo EM, Belch G, Martinez ME, Ayala GX. Fruit and Vegetable Intake of US Hispanics by Food Store Type: Findings from NHANES. *J Racial Ethn Health Disparities* 2019;6(1):220-9. doi: 10.1007/s40615-018-0517-7.

47. Mattei J, Sotres-Alvarez D, Daviglius ML, Gallo LC, Gellman M, Hu FB, Tucker KL, Willett WC, Siega-Riz AM, Van Horn L, et al. Diet Quality and Its Association with Cardiometabolic Risk Factors Vary by Hispanic and Latino Ethnic Background in the Hispanic Community Health Study/Study of Latinos. *J Nutr* 2016;146(10):2035-44. doi: 10.3945/jn.116.231209.
48. Carrera PM, Gao X, Tucker KL. A study of dietary patterns in the Mexican-American population and their association with obesity. *J Am Diet Assoc* 2007;107(10):1735-42. doi: 10.1016/j.jada.2007.07.016.
49. Lin H, Bermudez OI, Tucker KL. Dietary patterns of Hispanic elders are associated with acculturation and obesity. *J Nutr* 2003;133(11):3651-7. doi: 10.1093/jn/133.11.3651.
50. Noel SE, Newby PK, Ordovas JM, Tucker KL. A traditional rice and beans pattern is associated with metabolic syndrome in Puerto Rican older adults. *J Nutr* 2009;139(7):1360-7. doi: 10.3945/jn.109.105874.
51. Sofianou A, Fung TT, Tucker KL. Differences in diet pattern adherence by nativity and duration of US residence in the Mexican-American population. *J Am Diet Assoc* 2011;111(10):1563-9 e2. doi: 10.1016/j.jada.2011.07.005.
52. Mattei J, McClain AC, Falcon LM, Noel SE, Tucker KL. Dietary Acculturation among Puerto Rican Adults Varies by Acculturation Construct and Dietary Measure. *J Nutr* 2018;148(11):1804-13. doi: 10.1093/jn/nxy174.
53. Arandia G, Nalty C, Sharkey JR, Dean WR. Diet and acculturation among Hispanic/Latino older adults in the United States: a review of literature and recommendations. *J Nutr Gerontol Geriatr* 2012;31(1):16-37. doi: 10.1080/21551197.2012.647553.
54. Ayala GX, Baquero B, Klinger S. A systematic review of the relationship between acculturation and diet among Latinos in the United States: implications for future research. *J Am Diet Assoc* 2008;108(8):1330-44. doi: 10.1016/j.jada.2008.05.009.
55. Liu J, Fulda KG, Tao MH. Association between acculturation and metabolic syndrome in Hispanic adults mediated by fruits intake. *Public Health Nutr* 2021;24(18):6472-6. doi: 10.1017/S1368980021003530.
56. Jonnalagadda SS, Mitchell DC, Smiciklas-Wright H, Meaker KB, Van Heel N, Karmally W, Ershow AG, Kris-Etherton PM. Accuracy of energy intake data estimated by a multiple-pass, 24-hour dietary recall technique. *J Am Diet Assoc* 2000;100(3):303-8; quiz 9-11. doi: 10.1016/s0002-8223(00)00095-x.
57. Mossavar-Rahmani Y, Shaw PA, Wong WW, Sotres-Alvarez D, Gellman MD, Van Horn L, Stoutenberg M, Daviglius ML, Wylie-Rosett J, Siega-Riz AM, et al. Applying Recovery Biomarkers to Calibrate Self-Report Measures of Energy and Protein in the Hispanic Community Health Study/Study of Latinos. *Am J Epidemiol* 2015;181(12):996-1007. doi: 10.1093/aje/kwu468.