MS FREIDA BLOSTEIN (Orcid ID : 0000-0002-5734-4969)
PROFESSOR BETSY FOXMAN (Orcid ID : 0000-0001-6682-238X)

Article type : Original Manuscript

Dietary patterns associated with dental caries in adults in the United States

Authors: Freida A. Blostein[1], Erica C. Jansen[2], Andrew D. Jones[2], Teresa A. Marshall[3], Betsy Foxman[1]

[1] University of Michigan School of Public Health Department of Epidemiology
[2] University of Michigan School of Public Health Department of Nutritional Sciences
[3] University of Iowa College of Dentistry Department of Preventive and Community Dentistry

Address Correspondence to: Betsy Foxman, Center for Molecular and Clinical Epidemiology of Infectious Diseases, Office M5108 SPH II University of Michigan School of Public Health, Ann Arbor, Michigan. bfoxman@umich.edu

Author Contributions

Betsy Foxman and Andrew D. Jones conceived of the study. Freida Blostein ran the analysis and wrote the initial draft of the manuscript. Freida Blostein, Erica C. Jansen, Andrew D. Jones, Teresa A. Marshall & Betsy Foxman contributed to the conception, interpretation and critical revision of the manuscript. Data collection and study design was by the National Health and Nutrition Survey.

Abstract

Objectives: Dental caries experience, which affects 91% of US adults, is a consequence of a carious process influenced by diet. Although individual foods have been implicated, we hypothesized that dietary patterns might be important predictors of caries presence.

Methods: We analyzed data from 4467 people ≥18 years old participating in the 2013-2014 National Health & Nutrition Examination Survey, a nationally representative sample of the US population. Data from 24-hour dietary recalls were classified into standard food categories and reduced to 3 dietary patterns using principal components analysis. We used regression to model the log transformed decayed, missing, and filled teeth (DMFT) score and the prevalence of any caries experience by quartiles of principal component scores, controlling for potential confounders. Dietary patterns differed by age with respect to dental caries so 18-30-year-olds (n=1074) and >30-year-olds (n=3393) were analyzed separately.

Results: Similar dietary patterns existed among individuals aged 18-30 years and >30 years, but the prevalence of DMFT score>0 and the median of DMFT was greater in those >30: 78.7% (95% CI: 76.1%, 81.3%) vs 92.6% (95% CI: 91.4%, 93.7%) and 4 (95% CI: 4, 5); vs 12 DMFT (95% CI: 11, 13), respectively. In those 18-30, no dietary pattern was associated with greater prevalence or severity of dental caries experience. Among those >30, the prevalence of DMFT>)was higher by 2% for those in each subsequent quartile of a diet high in sugar-sweetened beverages and sandwiches (adjusted PR: 1.02, 95% CI: 1.001, 1.03) - thus, the prevalence of dental caries experience was 6% higher among those in the uppermost quartile than in the lowest quartile. For every subsequent quartile in the same pattern there was a 1.98% higher (95% CI: (0.15%, 3.85%)) DMFT score. However, analysis using the two strongest loading food groups from any of the principal components did not identify any predictors of caries experience.

Conclusions: Dietary patterns were associated with the prevalence of dental caries experience, with differing findings by age. Although effect sizes were small, the population impact may be substantial. While food groups high in sugar were associated with caries prevalence and severity, associations were more apparent in the context of overall diet. Prospective studies are needed to confirm whether particular dietary patterns are causally related to the development of dental caries.

Keywords: Dental decay, diet, dietary sugars, NHANES, cariogenic agent

Introduction

In 2011, 91% of American adults aged 20-64 experienced dental caries.[1] Untreated dental decay adversely affects quality of life, social relations and health.[2,3] Further, direct oral health care expenditures in the US are significant, exceeding $105 billion.[4] Coronal caries forms throughout the lifespan,[5,6] and is the primary cause of tooth loss in older adults, accounting for most of older adults' oral health expenditures.[5] Therefore, prevention of dental caries throughout life is important.

Teeth are constantly demineralized and remineralized; when demineralization outpaces remineralization, caries results.[7,8] Demineralization occurs when acidogenic bacterial species feed on dietary carbohydrates and produce weak organic acids,[7–9] thus, a cariogenic diet is an important risk factor.[2] Sugar is considered a cariogenic dietary component, and a robust literature has linked sugar consumption with dental caries.[10–12] In a longitudinal study of 533 American men, the frequency of sugar-sweetened beverage consumption was positively associated with root caries increment[11]; similarly, a longitudinal study of 939 Finnish adults found a positive association between the frequency of sugar-sweetened beverages and the net increment of decayed, missing and filled teeth (DMFT) over a 4-year study period.[12] Nonetheless, some contrasting findings exist. For example a study of 3,212 Danish adults found no association between sugar consumption and root caries.[13] Discrepancies between studies may be attributable to differences in exposure measurement – for example, measuring frequency of sugar consumption as opposed to amount of sugar consumed – or to variability in fluoride exposure. A study of 1,702 Finnish adults found associations between DMFT and the amount (but not frequency) of sugar consumed.[14] The same study found controlling for the use of fluoridated toothpaste reduced the strength of the association between amount of sugar consumed and DMFT.[14]

In contrast to sugary foods, dairy products may decrease caries risk.[2,15,16] Dairy products contain calcium, which may encourage enamel remineralization.[15] Two longitudinal studies, among 600 Japanese[17] and 432 Danish[18] adults respectively, found an inverse association between milk and caries incidence. Yet, in an analysis of 31,571 Swedish adults whose diet information was collected 0 to 5 years prior to a dental examination, mean decayed, missing and filled surface (DMFS) scores were higher among those with more frequent milk consumption. Among these individuals, higher frequency of milk consumption was associated with higher frequency of sweet snack consumption, leading the authors to postulate that the protective effects of milk may be modulated by patterns of consumption of other food groups.[19] This highlights the need to understand the possible antagonisms and synergies in cariogenic potential which may arise when foods are eaten in combination.[2,20–24]

Although certain foods are often consumed together, only a few studies have examined the association of dietary patterns with caries experience, and the findings are inconsistent. We found two studies of the association of dietary patterns with caries experience in adults. One, a longitudinal study in male veterans aged 47 to 90, found that individuals with better adherence to a high-quality anti-hypertensive diet experienced lower root caries increments than those with poor adherence.[11] However, in a Detroit cross-sectional study of 1,021 low-income adults, patterns of liquid (excluding soft drinks) and food consumption derived from factor analysis were not associated with caries after adjustment for age, education, income, frequency of tooth brushing and gingival plaque score and soft drink consumption, but soft drink consumption was associated with dental caries.[24] In response to the paucity of studies examining dietary patterns in relation to dental caries in adults, we investigated associations of dietary patterns with dental caries experience within the 2013-2014 National Health and Nutrition Examination Survey, a nationally representative survey of the United States. A secondary aim was to examine the associations between dental caries experience and individual foods found within the dietary patterns.

Methods

Study population: We analyzed data from the 2013-2014 National Health and Nutrition Examination Survey (NHANES).[25] NHANES uses a complex, multistage survey design to sample from the non-institutionalized, civilian population of the United States of America. NHANES III collected demographic and 24-hour dietary recall data and conducted dental health examinations.[26] We included all participants 18 years of age and older with complete dental examinations and two 24-hour dietary recalls. NHANES top-codes all individuals over 80 to 80 years of age to protect individuals' privacy. Edentulous individuals, defined as those with all teeth marked "Tooth not present" in the dental examination, were excluded. NHANES data are public use; thus, the University of Michigan institutional review board deems this work exempt from human subjects' regulations.

Exposure data: The first of two dietary interviews were conducted in person at the time of the health examination by a trained interviewer using the Automated Multiple-Pass Method, a computer-assisted dietary interview software developed by the United States Department of Agriculture (USDA).[27] Visual media were provided to respondents to assist in quantifying the amount of foods and beverages consumed. For each food/beverage, the respondent reported the day of the week of intake, whether the food/beverage was eaten in combination with other foods, the time and name of the eating occasion, where the item was obtained, whether the item was eaten at home, and the amount consumed. The second 24-hour recall interview was conducted by telephone 3-10 days after the first. The USDA Food and Nutrient Database for Dietary Studies (USDA FNDDS) was used to calculate respondents' intakes of energy, macronutrients, and 60 additional micronutrients and dietary components.[28] NHANES dietary survey weights take into account the day of the week used for reporting. The survey weights calculated based on individuals who completed both days of recall were used in the principal component analysis (PCA) and all subsequent statistical analysis.

A sensitivity analysis included all individuals who completed at least one day of dietary recall (Appendix Figure 1; Appendix Section 2). In the sensitivity analysis, dietary weights based upon only the first day of recall were used.

We collapsed the individual food items from the 24-hour recall into 153 mutually exclusive "What We Eat in America" (WWEIA) food categories developed for the dietary portion of NHANES by the United States Department of Agriculture.[28] These food categories were further collapsed into 48 broader food groups based on similarities in nutritional composition and usage (Supplemental Table 2). These food groups were generally coherent with respect to the cariogenicity of the included food categories, with the exception of cereals, for which both "high-sugar" and "low-sugar" cereals were grouped in the "Ready to eat cereals" food group. However, a sensitivity analysis in which high- and low-sugar cereals were not collapsed into the broader food group showed that the overall findings were unaltered when using the combined "Ready to eat cereals" food group (Appendix Section 4). For each food group, the grams consumed over all days of recall were summed and taken as a percentage of total grams consumed over the total period of recall (gram percentages).

Identification of dietary patterns: Gram percentages were centered, scaled to unit variance and used in a principal component analysis (PCA) to identify linear combinations of food groups which explained the greatest variance, i.e. dietary patterns. If no individual within the age group reported consumption of a food group, that group was not used in the PCA. This led to the exclusion of two food groups (human milk and infant formulas) for both the 18-30-year-olds and over-30-year-olds. The resulting factors were rotated orthogonally for interpretability. Initial analyses revealed important differences in dietary patterns by age; accordingly, PCA and all subsequent statistical analyses were performed separately by age group. The number of principal components (PCs) retained was based on eigenvalues>1, inspection of the scree plot (Appendix Figure 2), and interpretability. Loadings of food category variables > |0.25| were used to characterize PCs as dietary patterns. PC scores were categorized into quartiles and, after examining linearity, used as ordinal predictors in the statistical models.

Outcome data: Data from the examination by licensed dentists were compiled into a decayed, missing and filled tooth (DMFT) count. We analyzed DMFT scores as a dichotomous variable for presence of any DMFT (DMFT>0 vs DMFT=0) to assess prevalence of DMFT, and as a log-transformed continuous variable after excluding those with DMFT=0 to assess severity of DMFT.

Additional covariates: Potential confounders were identified based upon the prior literature. Summed kilocalorie counts were averaged over available days of dietary recall to create a mean daily energy intake variable which was used to account for confounding by energy consumption. Breakfast consumption and frequency of snack consumption were also examined as potential confounders. Counts of different independent eating occasions identified in Spanish or English as breakfast were averaged across days of recall to create an average breakfast consumption frequency variable. Counts of different independent eating occasions identified as snacks in Spanish or English were averaged over available days of recall to create an average snack frequency per day variable (Appendix Table 1). The average breakfast and snack variables take into account separate occasions of eating regardless of the number of food items consumed at each occasion. Body mass index (BMI) was calculated by NHANES, using weight in kilograms divided by height in meters squared (kg/m2), rounded to one decimal place. Gender, age, head of household education level and ratio of family income to poverty level were included in the demographic survey portion. For this analysis, the household head's education level was dichotomized into high school education or less and > high school education.

Statistical modeling: To investigate dietary patterns' associations with DMFT prevalence we used a modified Poisson approach.[29] The modified Poisson approach allowed us to estimate the prevalence ratio and is better suited for non-rare outcomes. To investigate dietary patterns' associations with greater severity of DMFT among those with DMFT score >0, we fit a linear regression predicting log-transformed DMFT score. Convergence issues when including the continuous parameterization of the average energy consumption variable prohibited use of a negative binomial model (which allows the inclusion of participants with DMFT=0). A sensitivity analysis using an alternative parameterization of the average energy consumption variable as a dichotomous variable in a negative binomial model is included in the Appendix (Appendix Section 5).

As a post-hoc test, the average number of snacks and average report of breakfast consumption per day were included in the models to see whether these variables explained observed associations between dietary patterns and dental outcomes.

To investigate whether highly-loading foods explained associations with dietary patterns, we fit the same models using gram percentages of the top two loading food groups from each PC as predictors in place of the quartile-ranked PC scores, including all previously described covariates.

We descriptively investigated associations of combinations of dietary patterns with dental decay. We dichotomized PC scores using the medians within each age group ("high" vs "low") and looked at all possible combinations of the resulting variables. This resulted in eight combinations of high and low for the 3 PCs. We visually inspected the distribution of DMFT within each combination and age group using violin plots with captive boxplots.

PCA and subsequent statistical models were performed in R, with complex sampling design accounted for using the svyprcomp and svyglm functions in the SRVY package.[30]

Results

After applying inclusion and exclusion criteria, 5043 individuals were eligible, of whom 4467 completed both days of recall and were eligible for the main analysis. Of these, 1074 were aged 18-30 years and 3393 were aged over 30 years (Supplemental Figure 1). Table 1 compares the distribution of sociodemographic and dietary variables between those with (DMFT>0) and without (DMFT=0) any caries experience within each of the age groups.

As expected, individuals with any caries experience tended to be older. Over-30-year-olds had higher median counts of DMFT (median (95% CI): 12 (11, 13)) than 18-30-year-olds (median (95% CI): 4 (4, 5)). The number of filled teeth was the largest contributor to the DMFT score in both age groups (mean filled tooth count (95% CI) age 18-30: 3.75 (3.37, 4.13), age>30: 8.97 (8.5, 9.45)). Over-30-year-olds had higher mean counts of missing teeth (1.92 (1.72, 2.12)) than 18-30-year-olds (0.28 (0.15, 0.42)) and slightly lower mean counts of decayed teeth (1.17 (0.97, 1.37)) than 18-30-year-olds (1.19 (0.93, 1.45)).

Over-30-year-olds had higher median ratios of family income to poverty level than those 30 or younger (median ratio of family income to poverty (95% CI) 18-30: 1.9 (1.6, 2.3) vs >30: 3.3 (2.9, 3.8). In both age groups, median count of DMFT was slightly higher among those living in poverty (ratio of family income to poverty <1). Among 18-30-year-olds, the median DMFT of those in poverty was 5 (4, 7) as compared to median DMFT of 4 (3, 5) for those not in poverty. Among those aged over 30 years, those in poverty had a median DMFT of 13 (12, 14) while those not in poverty had a median DMFT of 12 (11, 13)). Similarly, those with a High school education or less had slightly higher median DMFT counts than those with more than a High school education (18-30: <High school 5 (3, 6) vs > High school 4 (3, 5);>30: <High school 13 (12, 14) vs >High school 12 (11, 13)). Lower socioeconomic status individuals had more decayed and missing but fewer filled teeth than higher socioeconomic status individuals. (Appendix Table 5 and 6).

Over-30-year-olds had higher median BMIs on average (median BMI (95% CI) 28.2 (27.9, 28.6)) than 18-30-year-olds (median BMI (95% CI) 26.1 (25.1, 26.9)) but reported lower mean energy consumption in kcal (mean kcal (95% CI) 2050 (2002, 2098)) than the younger age group (mean kcal (95% CI) 18-30 2225 (2128, 2322)).The mean energy consumption (kcal) did not significantly differ between those with DMFT>0 vs DMFT=0 in either age group (18-30: DMFT>0: 2239 95% CI (2125, 2354) versus DMFT=0: 2171, 95% CI (2004, 2338); >30: DMFT>0 DMFT: 2047.99, 95% CI (1997, 2099); DMFT=0: 2077, 95% CI (1949, 2204)).

Principal Components (Table 2): Three dietary patterns with similar food loadings were identified in each age group. The first PC loaded positively on breads and high-fat foods such as cheese, fats and oils ("diet high in breads & fats") in the 18-30-year-olds; these items loaded negatively for the >30-year-olds, so we reverse-coded pattern scores for comparability. The second PC loaded positively on sugar-sweetened beverages and sandwiches, and negatively on fruit and water consumption ("diet high in sugar-sweetened beverages & sandwiches"). Again, the directions of the loadings were reversed in >30-year-olds, and thus were reverse-coded. The third PC captured variation in breakfast foods, with high loadings on milk and cereal consumption contrasted with tea and coffee consumption ("diet high in milk & cereal"). The first three PCs together explained approximately 16% of the total variation in the dietary recall among 18-30-year-olds (PC1: 7%, PC2: 5%, PC3: 4%) and 15% among those over 30 (PC1: 8%, PC2: 4%, PC3: 3%).

Poisson model (Table 3): Among 18-30-year-olds, no principal component was associated with prevalence of any DMFT. Of the 18-30-year-olds, 63.9% (95% CI: 58.6%, 69.2%) reported eating breakfast both days and the mean number of snacks per day was 1.7 (95% CI 1.6, 1.8). The addition of these variables did not affect the estimates. Among those over 30, every subsequent quartile of the diet high in sugar-sweetened beverages & sandwiches pattern was associated with a 2% higher prevalence of DMFT (95% CI: (0.14%, 3%)). For those over 30, 82.9% (95% CI: 80.4%, 85.4%) reported eating breakfast both days and the mean number of snacks per day was 2.02 (95% CI 1.9, 2.1). The addition of these variables did not alter the precision nor effect estimate. None of the two highest loading food groups from each PC were significant independent predictors of DMFT prevalence. In the final model, age was positively associated with prevalence of DMFT in both 18-30-year-olds and > 30-year-olds.

Linear model (Table 4): Among those 18-30 with any DMFT, no PC was associated with DMFT. By contrast, among those >30 with any DMFT, every subsequent quartile of the diet high in breads & fats pattern was associated with a 2.19% higher (95% CI: (0.48%, 3.93%)) DMFT score. In addition, every subsequent quartile of the diet high in sugar-sweetened beverages & sandwiches pattern was associated with a 1.98% higher (95% CI: (0.15%, 3.85%))

DMFT score. Inclusion of breakfast and snack variables made the estimates less precise. When we examined whether individual foods predicted higher DMFT, only sugar-sweetened beverages were marginally associated with greater DMFT: every one percent higher gram percentage of total grams consumed was associated with 24.42% higher DMFT, although the difference was not statistically significant (95% CI: (-1.01%, 56.37%). Age and ratio of family income to poverty were the only other variables in the final model associated with higher DMFT, with higher age positively associated and higher ratio of family income-to-poverty inversely associated in 18-30-year-olds. In >30-year-olds, higher age was positively associated with higher DMFT and being male was inversely associated with DMFT.

Sensitivity analyses: When individuals who completed only one day of dietary recall were included in the analysis, the directions of the associations did not change, and effect estimates changed only slightly in magnitude (see Appendix Section 2). Results were insensitive to the grouping of low and high sugar cereals (see Appendix Section 4). Results of the negative-binomial modeling approach were consistent with those from the log-linear approach for the dietary patterns, although additional dietary patterns and individual foods demonstrated associations with the DMFT count in the negative binomial model (see Appendix Section 5).

Patterns of Principal Components differed by age group (Figure 1): For 18-30-year-olds, those low in “diet high in sugar-sweetened beverages & sandwiches” and high in “diet high in breads & fats” and “diet high in milk and cereal” had the lowest median DMFT, while for those >30, those low in all PCs had the lowest median DMFT. Those high in all three patterns had the highest median DMFT score in both age groups. Among those >30, scoring highly on “diet high in milk & cereal” and “diet high in sugar-sweetened beverages & sandwiches” resulted in a lower median DMFT than scoring highly on “diet high in sugar-sweetened beverages & sandwiches” alone. Further, scoring highly on “diet high in breads & fats” and on “diet high in sugar-sweetened beverages & sandwiches” was associated with a higher median DMFT than scoring highly on “diet high in sugar-sweetened beverages & sandwiches” alone.

Discussion

In this nationally representative sample of US adults, we identified three dietary patterns among 18-to-30- and >30-year-olds. No pattern was associated with prevalence or severity of DMFT in those aged 18-30 years. However, a diet "high in sugar-sweetened beverages & sandwiches" was associated with DMFT prevalence and severity in >30-year-olds, and a "diet high in breads & fat" was associated with severity of decay. Intake of individual foods loading heavily on these dietary patterns did not strongly predict dental caries. In line with current literature, we observed a strong and consistent positive association between age and dental decay and a more moderate association between lower socioeconomic status and dental decay.[31,32]

Strengths of our analysis include the large sample size, nationally representative data, and high-quality outcome data. An additional strength was the use of data from two 24-hour recall assessments. A single day of dietary recall can be a random, non-representative snapshot of an individual's true diet, and therefore including individuals with only one day of dietary recall can affect the precision of exposure measurement.[33] Despite this strength, the exposure measurement used in our study has several weaknesses. Twenty-four-hour recalls may not be an accurate measure of usual intake and are memory dependent. Additionally, frequency of consumption may be more relevant to caries experience than the amount of food consumed.[2] Unfortunately, a food frequency questionnaire was not used in the more recent cycles of NHANES.[2,26] Our use of WWEIA food groupings allowed a higher-level exposure categorization but may have obscured differences in relevant nutritional components, such as free sugars, between individual food items of the same food group. The cross-sectional design of NHANES is also a major limitation of our study as it prohibits causal inference and parsing of age, period and cohort effects. The complex survey design limited the number of residual degrees of freedom available restricting our ability to more fully explore interactions between dietary patterns and age and interactions among the dietary patterns.

Our findings are consistent with the literature on the cariogenicity of the underlying food groups including sugar-sweetened beverages.[2,11,14,34–36] Notably, individual food groups underlying the dietary patterns of our study did not always associate with dental caries experience, indicating the possible role of food interactions in caries risk. This is consistent with one study in American adults[11] as well as with a study of dietary patterns in 504 Australian adolescents, where high-starch dietary patterns predicted caries experience but no significant correlations between individual foods and caries increment were identified.[21] However a cross-sectional study of dietary patterns and caries risk among Detroit adults found an association between sugar-sweetened beverages and caries, but did not find associations between patterns of liquid and food consumption and caries after multivariate adjustment.[24] That study was conducted among 821 low-income African-American individuals, while our study population is larger and nationally representative. Additionally, that study grouped solid and liquid food frequency data into separate patterns. We allowed solid and liquid food groups to be grouped together and used a measure of food amount rather than food frequency. We believe allowing liquid and solid foods to be grouped together more realistically reflects dietary patterns and that this is a strength of our study. However, as discussed above, our use of gram consumption as opposed to food frequency is a potential limitation which may explain these differences.[2]

A notable finding from our study was the age-specificity of the associations between dietary patterns and caries. Differences in associations may indicate mechanistic changes in dental decay due to aging, such as changes in cariogenic microbiota or calcium absorption.[5,37] It is possible associations are only revealed in older adults because the lifelong, cumulative exposure to a cariogenic diet leads to dental decay. Consistent with the literature, adults >30 years had more dental decay than younger adults.[39] Alternatively, slight differences in food exposures by age groups may explain age-specific associations; for example, a diet high in breads and fats loaded strongly on cheese in the younger age group but not in the older age group. Cheese and other dairy foods have a cariostatic effect, potentially explaining why a diet high in breads and fats was only associated with severity of dental decay in the older adults. Cohort and period effects could also explain this finding: cumulative fluoride exposure differences by birth cohort or a period effect related to the introduction of fluoride products could modify relationships between food intake and caries outcomes.[35,38] Alternatively, these differences in associations could reflect reverse causation, with changes in eating habits resulting from age-related tooth loss.[37] Owing to the cross-sectional nature of NHANES, it was not possible to tease out age, period and cohort effects or to exclude non-causal explanations for age-specific differences in associations between dietary patterns and dental decay; longitudinal study designs are needed.

Our study is one of only a few to have examined the impact of dietary patterns on dental caries, and to explore the effects of these patterns on caries in adults. Although effect estimates using principal components were small, preventing even a small amount of tooth decay through dietary interventions could have large health benefits and cost-savings at a national scale. While food groups high in sugar were associated with caries prevalence and severity, associations were more apparent in the context of overall diet. Policy recommendations pertaining to total diet, rather than single foods or individual nutritional components, may be relevant for lowering caries risk. Moreover, as noted in the Global Burden of Disease Study, a suboptimal diet can have broader negative impacts on health beyond adversely affecting oral health.[40]

References

1. Dye BA, Thornton-Evans G, Li X, Iafolla TJ. *Dental caries and tooth loss in adults in the United States, 2011–2012. NCHS Data Brief, no 197*. Hyattsville, MD; 2015.
2. Moynihan P, Petersen PE. Diet, nutrition and the prevention of dental diseases. *Public Health Nutr*. 2004;7(1a):201-226.
3. Peres MA, Macpherson LMD, Weyant RJ, et al. Oral diseases: a global public health challenge. *Lancet*. 2019;394(10194):249-260.
4. Meier T, Deumelandt P, Christen O, Stangl GI, Riedel K, Langer M. Global burden of sugar-related dental diseases in 168 countries and corresponding health care costs. *J Dent Res*. 2017;96(8):845-854.
5. Saunders RH, Meyerowitz C. Dental caries in older adults. *Dent Clin North Am*. 2005;49(2):293-308.
6. Thomson WM. Dental caries experience in older people over time : what can the large cohort studies tell us ? *Br Dent J*. 2004;196(2):89-92.
7. Pitts NB, Zero DT, Marsh PD, et al. Dental caries. *Nat Rev Dis Prim*. 2017;3:17030.
8. Selwitz RH, Ismail AI, Pitts NB. Dental caries. *Lancet*. 2007;369(9555):51-59.
9. Bradshaw DJ, Lynch RJM. Diet and the microbial aetiology of dental caries: new paradigms. *Int Dent J*. 2013;63:64-72.
10. Sheiham A, James WPT. A reappraisal of the quantitative relationship between sugar

intake and dental caries: the need for new criteria for developing goals for sugar intake. *BMC Public Health*. 2014;14(1):863.

11. Kaye EK, Heaton B, Woosung S, Rich SE, Avron S, Garcia RI. The dietary approaches to stop hypertension diet and new and recurrent root caries events in men. *J Am Geriatr Soc*. 2015;63(9):1812-1819.
12. Bernabé E, Vehkalahti MM, Sheiham A, Aromaa A, Suominen AL. Sugar-sweetened beverages and dental caries in adults: A 4-year prospective study. *J Dent*. 2014;42(8):952-958.
13. Christensen LB, Bardow A, Ekstrand K, et al. Root caries, root surface restorations and lifestyle factors in adult Danes. *Acta Odontol Scand*. 2014;73(6):467-473.
14. Bernabé E, Vehkalahti MM, Sheiham A, Lundqvist A, Suominen AL. The shape of the dose-response relationship between sugars and caries in adults. *J Dent Res*. 2016;95(2):167-172.
15. Moynihan P. Foods and factors that protect against dental caries. 2000:281-286.
16. Dror DK, Allen LH. Dairy product intake in children and adolescents in developed countries: trends, nutritional contribution, and a review of association with health outcomes. *Nutr Rev*. 2014;72(2):68-81.
17. Yoshihara A, Watanabe R, Hanada N, Miyazaki H. A longitudinal study of the relationship between diet intake and dental caries and periodontal disease in elderly Japanese subjects. *Gerodontology*. 2009;26(2):130-136.
18. Adegboye ARA, Twetman S, Christensen LB, Heitmann BL. Intake of dairy calcium and tooth loss among adult Danish men and women. *Nutrition*. 2012;28(7/8):779-784.
19. Johansson I, Esberg A, Eriksson L, Haworth S, Lif Holgerson P. Self-reported bovine milk intake is associated with oral microbiota composition. *PLoS One*. 2018;13(3):e0193504.
20. Hu FB. Dietary pattern analysis: A new direction in nutritional epidemiology. *Curr Opin Lipidol*. 2002;12(3):3-9.
21. Campain CA, Morgan VM, Evans WR, et al. Sugar–starch combinations in food and the relationship to dental caries in low-risk adolescents. *Eur J Oral Sci*. 2003;111(4):316-325.
22. Lim S, Sohn W, Burt BA, et al. Cariogenicity of soft drinks, milk and fruit juice in low-income African-American children. *J Am Dent Assoc*. 2008;139(7):959-967.

23. Sohn W, Burt BA, Sowers MR. Carbonated soft drinks and dental caries in the primary dentition. *J Dent Res*. 2006;85(3):262-266.
24. Burt BA, Kolker L, Sandretto AM. Dietary Patterns Related to Caries in a Low-Income Adult Population. 2006;2029:473-480.
25. Centers for Disease Control and Prevention (CDC) ., National Center for Health Statistics (NCHS). *National Health and Nutrition Examination Survey Data.* Hyattsville, MD
26. Centers for Disease Control and Prevention (CDC) ., National Center for Health Statistics (NCHS) . *National Health and Nutrition Examination Survey Examination Protocols for Disease Control and Prevention*. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention https://wwwn.cdc.gov/nchs/nhanes/ContinuousNhanes/Default.aspx?BeginYear=2013.
27. Raper N, Perloff B, Ingwersen L, Steinfeldt L, Anand J. An overview of USDA's Dietary Intake Data System. *J Food Compos Anal*. 2004;17(3):545-555.
28. U. S. Department of Agriculture Agricultural Research Service. *What We Eat in America Food Categories 2013-2014.*; 2016. www.ars.usda.gov/nea/bhnrc/fsrg.
29. Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol*. 2004;159(7):702-706.
30. Lumley T. survey: analysis of complex survey samples. *R Packag version 332*. 2017.
31. Shaffer JR, Polk DE, Feingold E, et al. Demographic, socioeconomic, and behavioral factors affecting patterns of tooth decay in the permanent dentition: principal components and factor analyses. *Community Dent Oral Epidemiol*. 2013;41(4):364-373.
32. Liu Y, Li Z, Walker MP. Social disparities in dentition status among American adults. *Int Dent J*. 2014;64(1):52-57.
33. Willett W. *Nutritional Epidemiology*. 3rd ed. Oxford University Press; 2012.
34. Rugg-Gunn AJ, Hackett AF, Appleton DR, Jenkins GN, Eastoe JE. Relationship between dietary habits and caries increment assessed over two years in 405 English adolescent school children. *Arch Oral Biol*. 1984;29(12):983-992.
35. Heller KE, Burt BA, Eklund SA. Sugared soda consumption and dental caries in the United States. *J Dent Res*. 2001;80(10):1949-1953.
36. Moynihan PJ, Kelly SAM. Effect on caries of restricting sugars intake: Systematic review to inform WHO guidelines. *J Dent Res*. 2014;93(1):8-18.

37. Carvalho TS, Lussi A. Age-related morphological, histological and functional changes in teeth. *J Oral Rehabil*. 2017;44(4):291-298.

38. Jones S, Brian B, Poul E, Lennon M. The effective use of fluorides in public health. *Bull World Health Organ*. 2005;83(9):670.

39. Dye BA, Tan S, Smith V, et al. Trends in oral health status: United States, 1988-1994 and 1999-2004. *Vital Health Stat 11*. 2007;(248):1-92.

40. Afshin A, Sur P, Fay K, et al. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet (London, England)*. 2019;393(10184):1958-1972.

Author Contributions

Betsy Foxman and Andrew D. Jones conceived of the study. Freida Blostein ran the analysis and wrote the initial draft of the manuscript. Freida Blostein, Erica C. Jansen, Andrew D. Jones, Teresa A. Marshall & Betsy Foxman contributed to the conception, interpretation and critical revision of the manuscript. Data collection and study design was by the National Health and Nutrition Survey.

Acknowledgements

FB, BF, TM, & AJ were funded by NIH grant R01DE014899 (BF) and EJ was funded by 5T32DK071212-12 (EJ). The authors declare no potential conflicts of interest.

Main Paper Tables and Figures

Table 1: Associations between sociodemographic characteristics and number of Decayed Missing and Filled teeth (DMFT). Individuals 18 and over in the National Health and Nutrition Examination Survey (NHANES) III 2013-2014.

	18 to 30 Years of Age n=1074			31 to 80 & Over Years of Age[1] n=3393		
	DMFT score = 0 % (95% CI)	>0 DMFT score % (95% CI)	(p)	DMFT score = 0 % (95% CI)	>0 DMFT score % (95% CI)	(p)
Percent of Sample	21.3% (18.7%, 23.9%)	78.7% (76.1%, 81.3%)		7.4% (6.3%, 8.6%)	92.6% (91.4%, 93.7%)	
Age[1]*			<0.01			<0.01
18-24 Years	73.3% (65.6%, 81%)	53.8% (48.5%, 59.1%)		NA	NA	
25-30 Years	26.7% (19%, 34.4%)	46.2% (40.9%, 51.5%)		NA	NA	
31-60 Years	NA	NA		81.0% (75.5%, 86.5%)	70.0% (67.9%, 72.2%)	
Over 60 Years	NA	NA		19.0% (13.5%, 24.5%)	30.0% (27.8%, 32.1%)	
Gender			0.39			0.08
Male	53.4% (43.4%, 63.4%)	48.2% (44%, 52.3%)		41.4% (35.5%, 47.3%)	47.9% (45.1%, 50.6%)	
Head of Household Education			0.66			<0.01
Greater than High school	63% (51.9%, 74.2%)	60.5% (56.5%, 64.5%)		52.1% (42.0%, 62.2%)	66.7% (62.0%, 71.3%)	
Ratio of Family Income to Poverty[2]*			0.26			0.25
In poverty	20.9% (13.2%, 28.6%)	25.1% (19.5%, 30.8%)		14.7% (9.8%, 19.6%)	11.7% (8.4%, 15.0%)	
Average Daily Energy Consumption*			0.74			0.62
Equal or Over 2000 kilocalories	55.1% (47.1%, 63.0%)	53.5% (46.6%, 60.5%)		49.7% (40.9%, 58.4%)	47.6% (44.6%, 50.7%)	
BMI[3]			0.52			0.74
Normal	45.3% (35.3%, 55.3%)	39.3% (34.3%, 44.4%)		24.1% (16.3%, 32.0%)	25.2% (22.7%, 27.8%)	
Overweight	23.5% (16.0%, 31.1%)	28.9% (25.0%, 32.7%)		36.1% (29.6%, 42.6%)	33.7% (30.9%, 36.6%)	
Obese (Class 1-3)	28.5% (21.0%, 36.0%)	28.8% (24.9%, 32.8%)		39.3% (32.6%, 46.0%)	40% (36.3%, 43.7%)	

Table Note: Results weighted to represent the United States.

[1]NHANES participants over 80 years of age are top-coded at 80 years of age.

[2]A ratio of family income to poverty <1 indicates a family that is living in poverty.[26]

[3]BMI categories based on Center for Disease Control BMI categories for adults: Underweight: <18.5, Normal: 18.5-24.9, Overweight: 25-29.9, Obese: ≥30. Categories not adjusted for adolescents; adult categories used throughout.

*Variables thus marked are categorized for display purposes in this table only and were parameterized as continuous variables in multivariable models.

Table 2: Characterization of Principal Components: Food group variables with loadings > |0.25|from principal component analysis on by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014

		18 to 30 Years of Age	31 to 80 & Over Years of Age*
Dietary Pattern	Food Group (Number of WWEIA categories)	Loading	Loading
"Diet high in breads & fats"			
	Breads, Rolls & Tortillas (4)	0.34	-0.27
	Cheese (2)	0.29	NA
	Fats/Oils (6)	0.25	-0.27
	Cured Meats/Poultry (4)	0.25	NA
"Diet high in sugar-sweetened beverages & sandwiches"			
	Sweetened Beverages (5)	0.35	-0.38
	Vegetables, excluding potatoes (11)	NA	0.35
	Sandwiches (5)	0.25	-0.33

	Fruits (9)	-0.30	0.29
	Cooked Grains (2)	NA	0.26
	Plain Water (2)	-0.26	0.25
	Protein & Nutritional Powders (1)	-0.26	NA
"Diet high in milk & cereal"			
	Ready-To-Eat Cereals (2)	0.41	0.41
	Milk (4)	0.33	0.36
	Flavored Milk (4)	0.26	NA
	Fruits (9)	NA	0.25
	100% Juice (4)	0.25	NA
	Coffee & Tea (2)	-0.28	-0.27

*In over-30-year-olds, "Breads & fats" and "Sugar-Sweetened beverages & sandwiches" were recoded to reverse directionality for future analysis, however, original loadings are presented in this table.

NA indicates a food category for which the absolute value of the loading was not above 0.25 for the age group despite being above 0.25 in the other age group.

Table 3: Associations of Principal Components (PC) with Any Decayed, Missing and Filled Teeth (DMFT>0) for Subsequent Quartiles in Principal Component (PC) Scores by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014

Principal Component Models			
18 to 30 Years of Age		Over 30 Years of Age	
	"Sugar-Sweetened		"Sugar-Sweetened beverages & sandwiches"

	"Breads & fats" Prevalence Ratio e^{β}, (95% CI)	beverages & sandwiches" Prevalence Ratio e^{β}, (95% CI)	"Milk & cereal" Prevalence Ratio e^{β}, (95% CI)	"Breads & fats" Prevalence Ratio e^{β}, (95% CI)	Prevalence Ratio e^{β}, (95% CI)	"Milk & cereal" Prevalence Ratio e^{β}, (95% CI)
Model[1]	0.99 (0.96, 1.02)	1.02 (0.98, 1.06)	0.98 (0.94, 1.02)	1.01 (0.997, 1.03)$^{+}$	1.01 (0.99, 1.02)	0.99 (0.98, 1.002)$^{+}$
Model 2[2]	0.99 (0.94, 1.03)	1.02 (0.97, 1.07)	0.99 (0.95, 1.04)	1 (0.99, 1.02)	1.02 (1.003, 1.03)*	0.99 (0.97, 1.003)
Model 3[3]	0.99 (0.94, 1.04)	1.02 (0.97, 1.07)	0.99 (0.94, 1.04)	1 (0.99, 1.02)	1.02 (1.002, 1.03)*	0.99 (0.97, 1.004)
Model 4[4]	0.98 (0.93, 1.04)	1.03 (0.97, 1.09)	0.99 (0.94, 1.04)	1 (0.98, 1.03)	1.02 (1.001, 1.03)*	0.99 (0.97, 1.01)

Individual Food Models

	18 to 30 Years of Age			Over 30 Years of Age		
Highest Loading Food Group	Bread Prevalence Ratio	Sweetened Beverages Prevalence Ratio	Cereals Prevalence Ratio	Fats/Oils Prevalence Ratio	Sweetened Beverages Prevalence Ratio	Cereals Prevalence Ratio
Food Categories Model[5]	1.14 (0.01, 100.99)	1.11 (0.7, 1.76)	0.95 (0.86, 1.04)	2.47 (0.63, 9.72)	1.04 (0.83, 1.3)	0.99 (0.95, 1.03)
Second Highest Loading Food Group	Cheese Prevalence Ratio	Fruits (Negative Loading) Prevalence Ratio	Milk Prevalence Ratio	Bread Prevalence Ratio	Vegetables (Negative Loading) Prevalence Ratio	Milk Prevalence Ratio
Food Categories Model[5]	0.97 (0.86, 1.09)	0.21 (0.01, 8.06)	2.18 (0.74, 6.46)$^{+}$	0.92 (0.32, 2.62)	1.02 (0.49, 2.13)	0.85 (0.56, 1.28)

*$p<0.05$ $^{+}$ $p<0.10$

[1] Model 1 included all three Principal Component (PC) variables as quartile ranking variables modeled ordinally. Each prevalence ratio corresponds to the change from one quartile to the next subsequent quartile.

[2] Model 2 contained the three PC variables described above and the following sociodemographic variables: gender, age (continuous), head of household education indicator variable for ≥ high school education and ratio of family income to poverty (continuous).

[3] Model 3 contained all the same variables as Model 2 and the following dietary variables: mean daily energy (continuous, kilocalories), body mass index (continuous variable).

[4] Model 4 was contained the same variables as Model 2 with the addition of the average snacking occasions per day and average breakfast per day variables

[5] The highest and second highest loading food group (based on absolute value) from each principal component were included as predictors in models as percents' gram consumption / total gram consumption averaged over two days. Low gram percentage consumption of cereals and cheese resulted

in very wide confidence intervals. These models included the sociodemographic and dietary variables listed in footnotes 2 and 3.

Table 4: Percent Change in Number of Decayed, Missing and Filled Teeth (DMFT) for each Subsequent Quartile in Identified Principal Components or Top Loading Food Groups from Principal Components Among those with DMFT > 0, by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014

Principal Component Models						
	18 to 30 Years of Age			Over 30 Years of Age		
	"Breads & fats" Percent Change $(e^{\beta} - 1) * 100$ (95% CI)	"Sugar-Sweetened beverages & sandwiches" Percent Change $(e^{\beta} - 1) * 100$ (95% CI)	"Milk & cereal" Percent Change $(e^{\beta} - 1) * 100$ (95% CI)	"Breads & fats" Percent Change $(e^{\beta} - 1) * 100$ (95% CI)	"Sugar-Sweetened beverages & sandwiches" Percent Change $(e^{\beta} - 1) * 100$ (95% CI)	"Milk & cereal" Percent Change $(e^{\beta} - 1) * 100$ (95% CI)
Model [1]	-2.08% (-5.74%, 1.72%)	3.66% (-0.65%, 8.15%)	-0.06% (-3.8%, 3.83%)	2.4%** (1.06%, 3.76%)	1.34%* (0.06%, 2.62%)	0.22% (-1.71%, 2.19%)
Model 2[2]	-2.49% (-6.23%, 1.41%)	3.09% (-2.01%, 8.45%)	-0.64% (-4.58%, 3.47%)	2.19%** (0.74%, 3.66%)	1.9%* (0.31%, 3.5%)	-0.28% (-1.78%, 1.25%)
Model 3[3]	-2.71% (-6.5%, 1.24%)	2.78% (-2.62%, 8.49%)	-0.53% (-4.65%, 3.77%)	2.24%* (0.76%, 3.75%)	1.8%* (0.19%, 3.43%)	-0.26% (-1.78%, 1.28%)
Model 4[4]	-3.16% (-7.7%, 1.62%)	3.92% (-2.18%, 10.41%)	-0.65% (-5.22%, 4.14%)	2.19%* (0.48%, 3.93%)	1.98%* (0.15%, 3.85%)	-0.48% (-2.17%, 1.25%)
Individual Food Models						
	18 to 30 Years of Age			Over 30 Years of Age		
Highest Loading Food Group	Bread Percent Change	Sweetened Beverages Percent Change	Cereals Percent Change	Fats/Oils Percent Change	Sweetened Beverages Percent Change	Cereals Percent Change
Food Categories Model[5]	-92.06% (-99.87%, 388.64%)	8.77% (-29.43%, 67.65%)	3.82% (-7.02%, 15.93%)	5.17% (-2.49%, 13.42%)	24.42%$^{+}$ (-1.01%, 56.37%)	-2.69% (-6.23%, 0.99%)
Second Highest	Cheese	Fruit	Milk	Bread	Vegetables	Milk

Loading Food Group	Percent Change	(Negative Loading) Percent Change	Percent Change	Percent Change	(Negative Loading) Percent Change	Percent Change
Food Categories Model[5]	1.19% (-9.85%, 13.58%)	-19.57% (-84.32%, 312.49%)	-4.12% (-69.24%, 198.82%)	-55.03% (-84.8%, 33.05%)	-8.54% (-48.83%, 63.46%)	13.35% (-26.84%, 75.63%)

$^{**}p<0.01$ $^{*}p<0.05$ $^{+}p<0.10$

[1] Model 1 included all three Principal Component (PC) variables as quartile ranking variables modeled ordinally. Each coefficient corresponds to the change from one quartile to the next subsequent quartile.

[2] Model 2 contained the three PC variables described above and the following sociodemographic variables: gender, age (continuous), head of household education indicator variable for ≥ high school education and ratio of family income to poverty (continuous).

[3]Model 3 contained all the same variables as Model 2 and the following dietary variables: mean daily energy (continuous, kilocalories), body mass index (continuous variable).

[4] Model 4 was contained the same variables as Model 2 with the addition of the average snacking occasions per day and average breakfast per day variables

[5]The highest and second highest loading food group (based on absolute value) from each principal component were included as predictors in models as percents' gram consumption / total gram consumption averaged over two days. Low gram percentage consumption of cereals and cheese resulted in very wide confidence intervals. These models included the sociodemographic and dietary variables listed in footnotes 2 and 3.

Text only Figure 1:

Figure 1: Distributions of Decayed, Missing and Filled Teeth (DMFT) by Combinations of Dichotomized ("High" versus "Low") Principal Components by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014

Figure Note: Within each age subset, violin plots with captured box plots are ordered from lowest to highest median DMFT score by pattern of principal components. Across age subsets, patterns retain the same color. Median DMFT score within each pattern is displayed above each violin plot and was estimated taking into account sample weights. High and Low in Principal Component refer respectively to being in the upper half or lower half of the principal component scores. PC1 - Diet high in cheese, bread, oils & fats. PC2 - Diet high in sandwiches and sugar sweetened beverages, low in water and fruit. PC3 - High in cereal and milk, low in coffee and tea.

Dietary patterns associate with dental caries in adolescents and adults

Authors: Freida A. Blostein[1], Erica C. Jansen[2], Andrew D. Jones[2], Teresa A. Marshall[3], Betsy Foxman[1]
[1] University of Michigan School of Public Health Department of Epidemiology
[2] University of Michigan School of Public Health Department of Nutritional Sciences
[3] University of Iowa College of Dentistry Department of Preventive and Community Dentistry

Address Correspondence to Betsy Foxman, Center for Molecular and Clinical Epidemiology of Infectious Diseases, University of Michigan School of Public Health, Ann Arbor, Michigan. bfoxman@umich.edu

Appendix Table of Contents

- Appendix Table 9: Table 4: Percent Change in Number of Decayed, Missing and Filled Teeth (DMFT) for each Quartile Increase in Identified Principal Components or Top Loading Food Groups from Principal Components Among those with DMFT>0, by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014 (Sensitivity Analysis)

Appendix Section 3: Alternative exposure variable parameterization for statistical models in main analysis

- Appendix Table 10: Associations of Principal Components (PC) with Any Decayed, Missing and Filled Teeth (DMFT >0) for Continuous Principal Component Coordinates by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014 (Main Analysis)
- Appendix Table 11: Percent Change in Number of Decayed, Missing and Filled Teeth (DMFT) for Continuous Principal Component Coordinates Among those with DMFT>0, by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014 (Main Analysis)
- Appendix Figure 5: Partial Regression Plots of Continuous Principal Component coordinates in Log Linear Models for Over 30 Year Olds from National Health and Nutrition Examination Survey 2013-14 (Main Analysis)

Appendix Section 4: Methods and results of sensitivity analysis: Separation of high- and low-sugar cereals

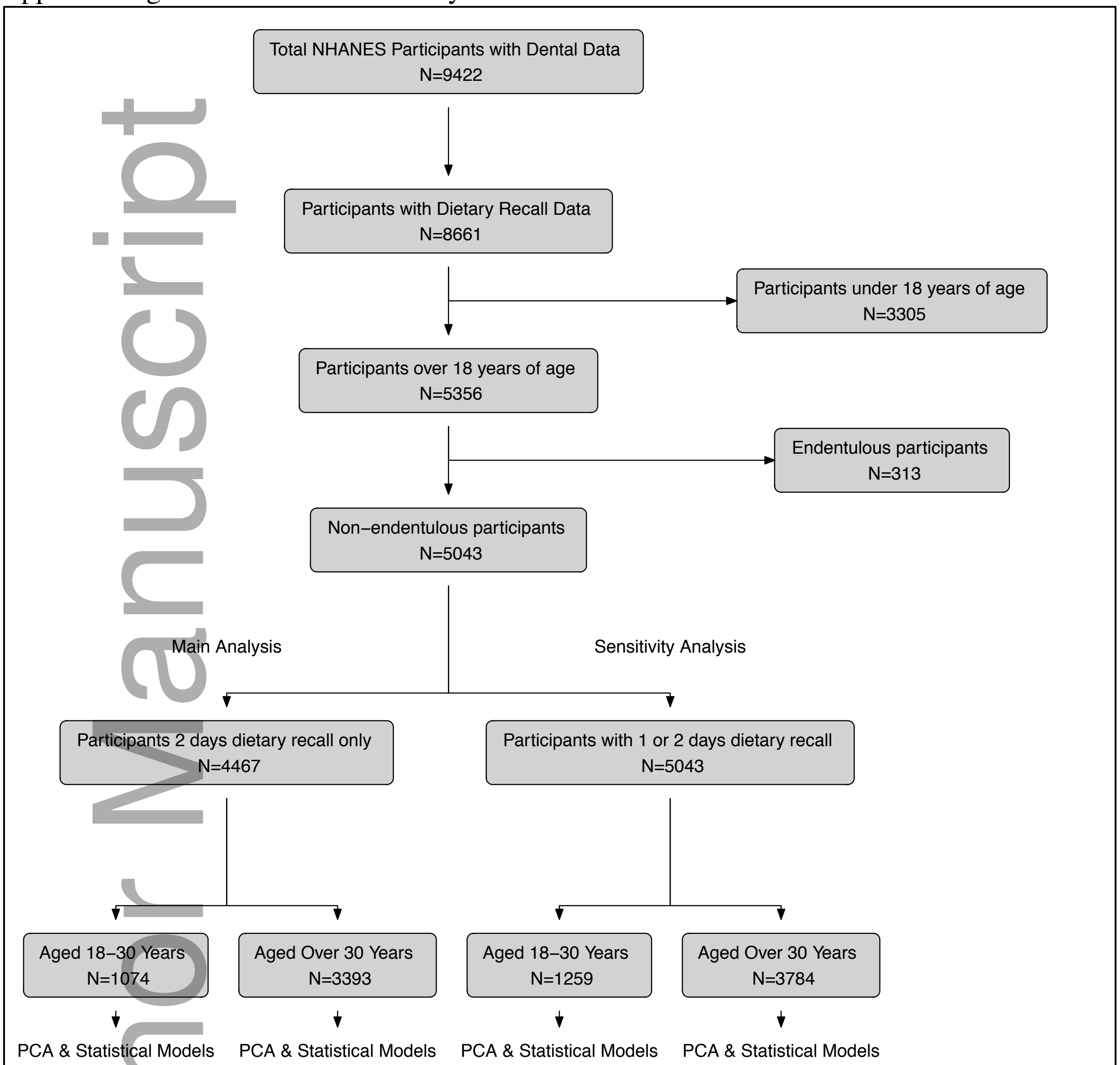
- Appendix Table 12 Loadings from principal component analysis; 18-30 year olds in the National Health and Nutrition Examination Survey 2013-2014 (Sensitivity Analysis: Separation of cereals)
- Appendix Table 13: Loadings from principal component analysis; over 30 year olds in the National Health and Nutrition Examination Survey 2013-2014 (Sensitivity Analysis: Separation of cereals)
- Appendix Table 14: Associations of Principal Components (PC) with Any Decayed, Missing and Filled Teeth (DMFT >0) for quartile rankings of Principal Component Coordinates by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014 (Main Analysis)
- Appendix Table 15: Percent Change in Number of Decayed, Missing and Filled Teeth (DMFT) for quartile rankings of Principal Component coordinates Among those with DMFT>0, by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014 (Main Analysis)

Appendix Section 5: Methods and results of sensitivity analysis: Utilization of a negative binomial model with dichotomous parameterization of average energy variable.

- Appendix Table 16: Change in mean number of decayed, missing and filled teeth (DMFT) for quartile rankings of Principal Component coordinates Among those with DMFT>0, by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014 (Main Analysis)

Methods & Additional Results of Main Analysis

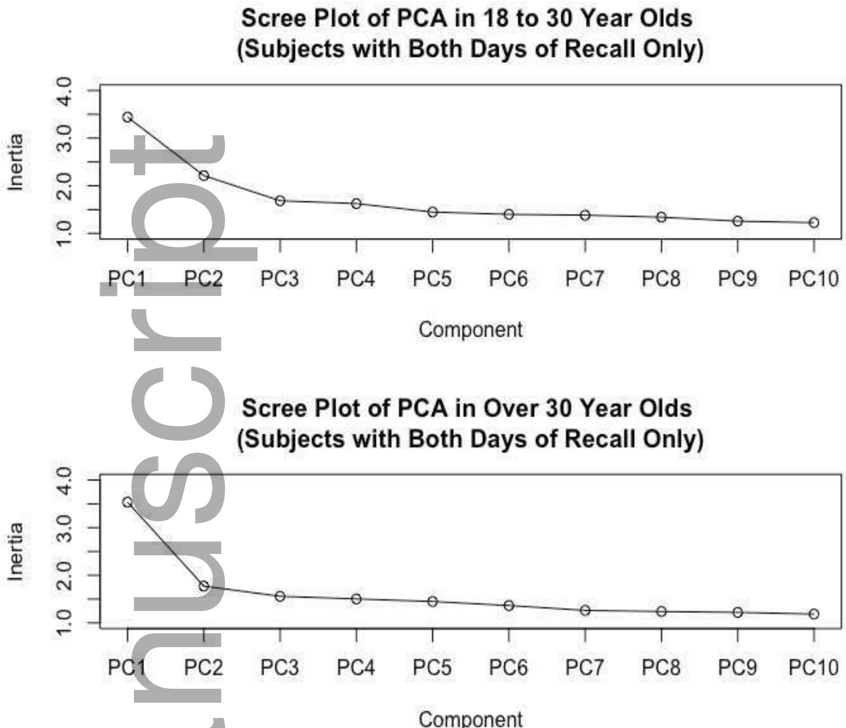
Appendix Figure 1: Flow chart of analysis subsets



Appendix Table 1: Breakfast and Snack Identifiers for Eating Occasion Descriptor in National Health and Nutrition Examination Survey 2013-14

Breakfast Identifiers	Snack Identifiers
"breakfast"	"snack"
"desayuno"	"extended consumption"
	"comida"
	"merienda"
	"entre comida"
	"botana"
	"bocadillo"
	"tentempie"

Appendix Figure 2: Scree plots used in principal component selection (Main Analysis)



Appendix Table 2: Summary of What We Eat in America food categories contained within each larger food group

Food groups (used in downstream analysis)	What We Eat in America Food Categories	Number of WWEIA Food categories in each food group
Milk	Milk, reduced fat	4
	Milk, whole	
	Milk, low-fat	
	Milk, nonfat	
Flavored Milk	Flavored milk, whole	4
	Flavored milk, nonfat	
	Flavored milk, reduced fat	
	Flavored milk, low-fat	
Dairy Drinks & Substitutes	Milk substitutes	2
	Milk shakes and other dairy drinks	
Cheese	Cheese	2
	Cottage/ricotta cheese	
Yogurt	Yogurt, regular	2
	Yogurt, Greek	
Meats	Beef, excludes ground	5
	Ground beef	

	Pork	
	Lamb, goat, game	
	Liver and organ meats	
Poultry	Turkey, duck, other poultry	3
	Chicken, whole pieces	
	Chicken patties, nuggets and tenders	
Seafood	Fish	2
	Shellfish	
Eggs	Eggs and omelets	1
Cured Meats & Poultry	Cold cuts and cured meats	4
	Bacon	
	Frankfurters	
	Sausages	
Plant-based Protein	Beans, peas, legumes	3
	Processed soy products	
	Nuts and seeds	
Mixed Meat, Poultry or Seafood Dishes	Meat mixed dishes	3
	Poultry mixed dishes	
	Seafood mixed dishes	
Mixed Grain-based Dishes	Rice mixed dishes	4
	Turnovers and other grain-based items	
	Pasta mixed dishes, excludes macaroni and cheese	
	Macaroni and cheese	
Asian Dishes	Stir-fry and soy-based sauce mixtures	3
	Fried rice and lo/chow mein	
	Egg rolls, dumplings, sushi	
Mixed Mexican Dishes	Burritos and tacos	3
	Other Mexican mixed dishes	
	Nachos	
Pizza	Pizza	1
Mixed Sandwiches	Other sandwiches (single code)	5
	Burgers (single code)	
	Chicken/turkey sandwiches (single code)	
	Egg/breakfast sandwiches (single code)	
	Frankfurter sandwiches (single code)	
Soups	Soups	1
Cooked Grains	Pasta, noodles, cooked grains	2
	Rice	

Breads, Rolls, & Tortillas	Yeast breads	4
	Rolls and buns	
	Bagels and English muffins	
	Tortillas	
Quick Breads and Products	Biscuits, muffins, quick breads	2
	Pancakes, waffles, French toast	
Ready-to-eat Cereals	Ready-to-eat cereal, lower sugar (=<21.2g/100g)	2
	Ready-to-eat cereal, higher sugar (>21.2g/100g)	
Cooked Cereals	Grits and other cooked cereals	2
	Oatmeal	
Savory Snacks	Tortilla, corn, other chips	4
	Pretzels/snack mix	
	Potato chips	
	Popcorn	
Crackers	Crackers, excludes saltines	2
	Saltine crackers	
Snack/Meal Bars	Cereal bars	2
	Nutrition bars	
Sweet Bakery Products	Cakes and pies	3
	Doughnuts, sweet rolls, pastries	
	Cookies and brownies	
Candy	Candy not containing chocolate	2
	Candy containing chocolate	
Other Desserts	Ice cream and frozen dairy desserts	3
	Pudding	
	Gelatins, ices, sorbets	
Fruits	Citrus fruits	9
	Dried fruits	
	Other fruits and fruit salads	
	Apples	
	Bananas	
	Melons	
	Berries	
	Grapes	
	Peaches and nectarines	
Vegetables, excluding Potatoes	Vegetable mixed dishes	11
	Other vegetables and combinations	
	Other starchy vegetables	

	Dark green vegetables, excludes lettuce	
	Lettuce and lettuce salads	
	Carrots	
	Other red and orange vegetables	
	Tomatoes	
	String beans	
	Corn	
	Onions	
White Potatoes	White potatoes, baked or boiled	3
	Mashed potatoes and white potato mixtures	
	French fries and other fried white potatoes	
100% Juice	Citrus juice	4
	Other fruit juice	
	Apple juice	
	Vegetable juice	
Diet Beverages	Diet soft drinks	3
	Other diet drinks	
	Diet sport and energy drinks	
Sweetened Beverages	Smoothies and grain drinks	5
	Fruit drinks	
	Soft drinks	
	Nutritional beverages	
	Sport and energy drinks	
Coffee & Tea	Coffee	2
	Tea	
Alcoholic Beverages	Beer	3
	Liquor and cocktails	
	Wine	
Plain Water	Tap water	2
	Bottled water	
Flavored or Enhanced Water	Flavored or carbonated water	2
	Enhanced or fortified water	
Fats & Oils	Cream and cream substitutes	6
	Cream cheese, sour cream, whipped cream	
	Butter and animal fats	
	Margarine	
	Mayonnaise	
	Salad dressings and vegetable oils	
Condiments & Sauces	Dips, gravies, other sauces	6

	Pasta sauces, tomato-based	
	Soy-based condiments	
	Mustard and other condiments	
	Olives, pickles, pickled vegetables	
	Tomato-based condiments	
Sugars	Jams, syrups, toppings	3
	Sugars and honey	
	Sugar substitutes	
Baby Foods	Baby food: yogurt	6
	Baby food: snacks and sweets	
	Baby food: meat and dinners	
	Baby food: cereals	
	Baby food: fruit	
	Baby food: vegetable	
Baby Beverages	Baby juice	2
	Baby water	
Infant Formulas	Formula, ready-to-feed	3
	Formula, prepared from powder	
	Formula, prepared from concentrate	
Human Milk	Human milk	1
Protein & Nutritional Powders	Protein and nutritional powders	1
Not included in a food category	Not included in a food category	1

Appendix Table 3: Loadings from principal component analysis; 18-30-year-olds in the National Health and Nutrition Examination Survey 2013-14 (Main Analysis)

What We Eat in America Food Categories	PC1	PC2	PC3
Milk	0.18	-0.11	0.33
Flavored Milk	0.08	0.06	0.26
Dairy Drinks & Substitutes	0.06	-0.22	0.01
Cheese	0.29	0.08	0.07
Yogurt	0.08	-0.19	0.12
Meats	0.16	-0.02	-0.11
Poultry	0.14	-0.03	-0.18
Seafood	0.03	-0.05	-0.04
Eggs	0.17	-0.18	-0.22
Cured Meats & Poultry	0.25	0.07	0.14
Plant-based Protein	0.14	-0.21	-0.07
Mixed Meat, Poultry or Seafood Dishes	0.06	0.01	0.11
Mixed Grain-based Dishes	0.09	0.13	0.14
Asian Dishes	0.02	-0.05	0.09
Mixed Mexican Dishes	0.01	0.08	0.06
Pizza	0.09	0.2	0
Mixed Sandwiches	0.03	0.25	-0.08

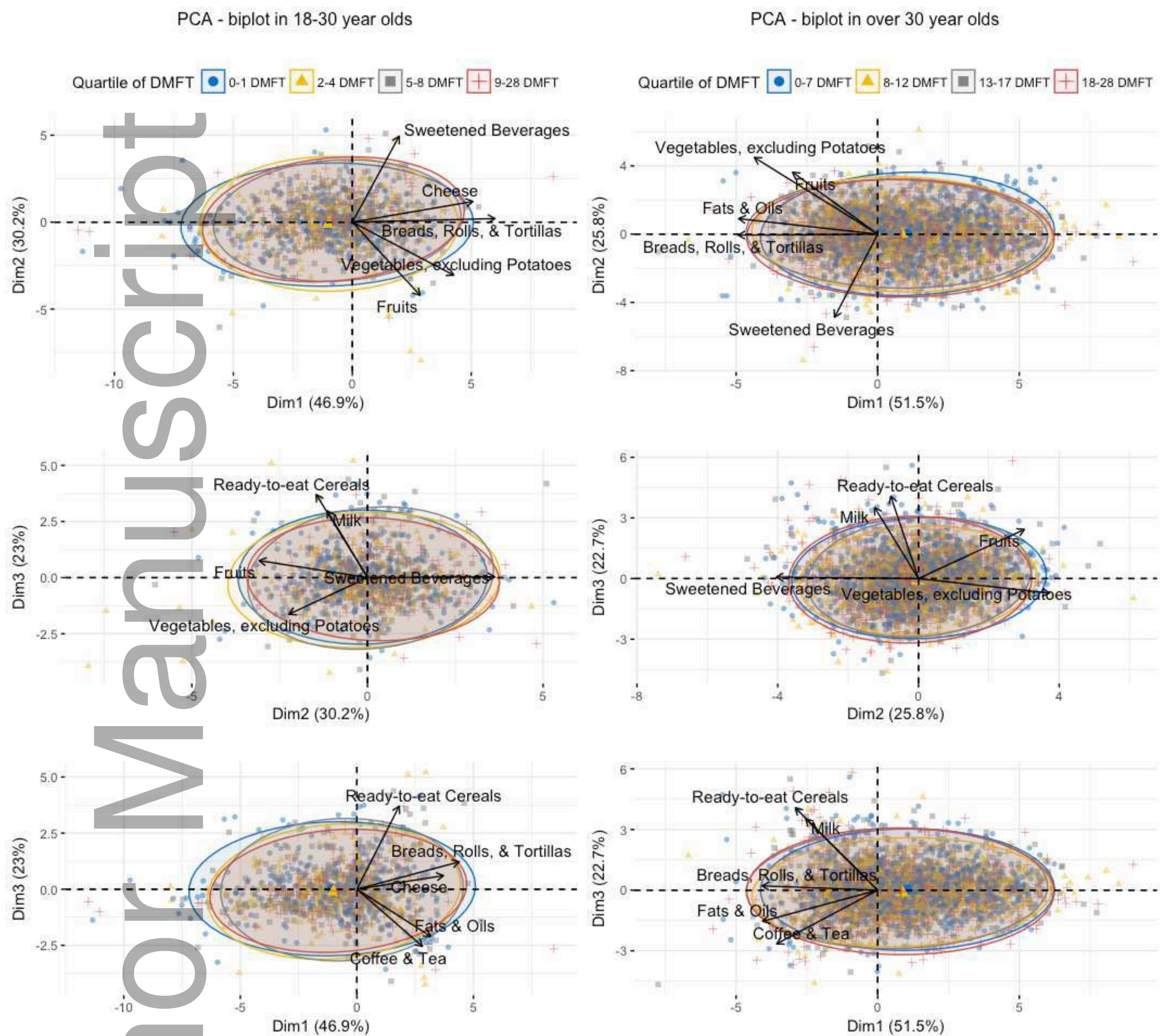
Soups	0.05	0.04	0.07
Cooked Grains	0.05	-0.23	-0.11
Breads, Rolls, & Tortillas	0.34	0.01	0.14
Quick Breads and Products	0.2	0.17	-0.13
Ready-to-eat Cereals	0.14	-0.14	0.41
Cooked Cereals	0.07	-0.11	0.09
What We Eat in America Food Categories (Continued)	PC1	PC2	PC3
Savory Snacks	0.19	0.13	0.12
Crackers	0.2	0.01	-0.11
Snack/Meal Bars	0.07	-0.09	-0.02
Sweet Bakery Products	0.15	0.07	0.09
Candy	0.14	0.17	0.11
Other Desserts	0.11	0.03	0.18
Fruits	0.16	-0.30	0.08
Vegetables, excluding Potatoes	0.24	-0.22	-0.18
White Potatoes	0.16	0.17	-0.07
100% Juice	0.14	0.03	0.25
Diet Beverages	0.03	0.08	0.02
Sweetened Beverages	0.11	0.35	0
Coffee & Tea	0.22	0.02	-0.28
Alcoholic Beverages	0.05	0.05	-0.07
Plain Water	0.13	-0.26	0.03
Flavored or Enhanced Water	0.04	0.03	-0.05
Fats & Oils	0.25	0.06	-0.23
Condiments & Sauces	0.21	0	-0.16
Sugars	0.21	0.15	-0.16
Baby Foods	-0.02	0	0
Baby Beverages	0.01	0.1	-0.04
Protein & Nutritional Powders	0.06	-0.26	-0.06
Not included in a food category	0.03	-0.18	-0.08

Appendix Table 4: Loadings from principal component analysis; over-30-year-olds in the National Health and Nutrition Examination Survey 2013-14 (Main Analysis)

What We Eat in America Food Categories	PC1	PC2	PC3
Milk	-0.17	-0.12	0.36
Flavored Milk	-0.03	-0.04	0.12
Dairy Drinks & Substitutes	-0.05	0.11	0.24
Cheese	-0.24	0.06	-0.05
Yogurt	-0.14	0.24	0.05
Meats	-0.17	-0.03	-0.23
Poultry	-0.11	0.07	-0.11
Seafood	-0.07	0.14	-0.08
Eggs	-0.14	0.07	-0.15
Cured Meats & Poultry	-0.2	-0.14	-0.04
Plant-based Protein	-0.11	0.18	0
Mixed Meat, Poultry or Seafood Dishes	-0.12	-0.08	0.04
Mixed Grain-based Dishes	-0.1	-0.06	-0.05
Asian Dishes	-0.04	0.1	-0.01
Mixed Mexican Dishes	-0.04	-0.12	0.07
Pizza	-0.09	-0.18	0.14
Mixed Sandwiches	-0.08	-0.33	0.08
Soups	-0.04	0.13	0.11

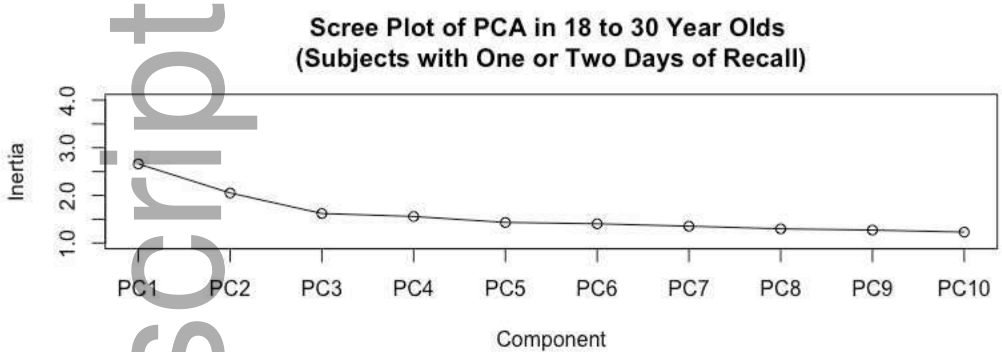
Cooked Grains	-0.01	0.26	0
Breads, Rolls, & Tortillas	-0.27	0	0.02
Quick Breads and Products	-0.12	0	-0.16
Ready-to-eat Cereals	-0.2	-0.07	0.41
Cooked Cereals	0	0.1	-0.07
Savory Snacks	-0.21	-0.14	0.05
What We Eat in America Food Categories (Continued)	PC1	PC2	PC3
Crackers	-0.2	-0.04	0.13
Snack/Meal Bars	-0.14	0.02	0.2
Sweet Bakery Products	-0.2	-0.11	-0.1
Candy	-0.17	-0.15	0.09
Other Desserts	-0.18	-0.08	0.1
Fruits	-0.17	0.29	0.25
Vegetables, excluding Potatoes	-0.24	0.35	-0.07
White Potatoes	-0.18	-0.17	-0.17
100% Juice	-0.09	0.06	0.08
Diet Beverages	-0.15	-0.12	0.03
Sweetened Beverages	-0.08	-0.38	0.01
Coffee & Tea	-0.24	-0.01	-0.27
Alcoholic Beverages	-0.08	-0.02	-0.24
Plain Water	-0.15	0.25	0.13
Flavored or Enhanced Water	-0.1	0.07	0.06
Fats & Oils	-0.27	0.07	-0.16
Condiments & Sauces	-0.21	0.04	-0.14
Sugars	-0.18	-0.1	-0.21
Protein & Nutritional Powders	-0.03	0.09	0.14
Not included in a food category	0	0.04	0.02

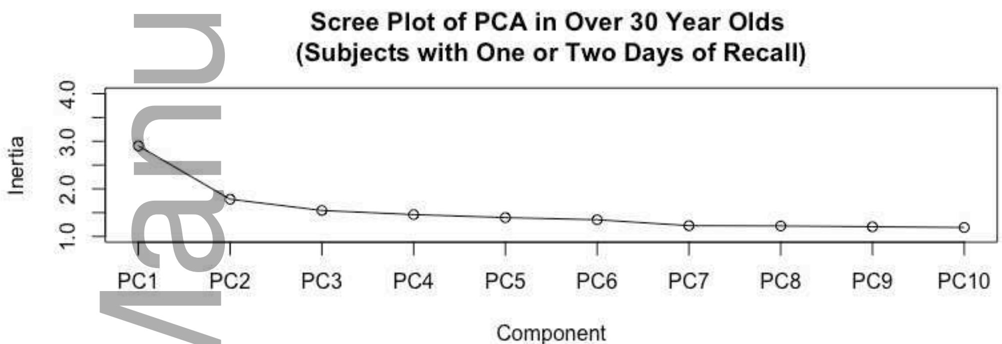
Appendix Figure 3: PCA Biplots



Methods & Results of Sensitivity Analysis: Using individuals with one and two days of available recall

Appendix Figure 4: Scree plots used in principal component selection (Sensitivity Analysis)





Appendix Table 5: Loadings from principal component analysis; 18-30-year-olds in the National Health and Nutrition Examination Survey 2013-14 (Sensitivity Analysis)

What We Eat in America Food Categories	PC1	PC2	PC3
Milk	0.2	0.03	-0.33
Flavored Milk	0.04	-0.01	-0.16
Dairy Drinks & Substitutes	0.04	0.21	-0.07
Cheese	0.31	-0.15	-0.08
Yogurt	0.11	0.16	-0.18
Meats	0.2	0	0.02
Poultry	0.15	0.01	0.24
Seafood	-0.01	0.03	0.04
Eggs	0.18	0.12	0.32
Cured Meats & Poultry	0.27	-0.08	-0.01
Plant-based Protein	0.12	0.16	-0.04
Mixed Meat, Poultry or Seafood Dishes	0.02	-0.01	-0.09
Mixed Grain-based Dishes	0.06	-0.15	-0.04
Asian Dishes	-0.02	0.09	-0.13
Mixed Mexican Dishes	-0.01	-0.1	-0.09
Pizza	-0.01	-0.2	-0.08
Mixed Sandwiches	-0.05	-0.26	-0.04
Soups	0.01	-0.01	-0.09
Cooked Grains	0.06	0.22	0.03

Breads, Rolls, & Tortillas	0.36	-0.01	-0.06
Quick Breads and Products	0.16	-0.17	0.2
What We Eat in America Food Categories (Continued)			
Ready-to-eat Cereals	0.18	0.12	-0.41
Cooked Cereals	0.08	0.08	-0.03
Savory Snacks	0.14	-0.15	-0.17
Crackers	0.18	-0.05	0
Snack/Meal Bars	0.08	0	-0.16
Sweet Bakery Products	0.14	-0.09	-0.23
Candy	0.1	-0.19	-0.18
Other Desserts	0.07	0.05	-0.16
Fruits	0.18	0.29	-0.08
Vegetables, excluding Potatoes	0.3	0.19	0.19
White Potatoes	0.13	-0.19	0.08
100% Juice	0.14	-0.01	-0.11
Diet Beverages	0.03	-0.08	-0.06
Sweetened Beverages	0.02	-0.42	-0.08
Coffee & Tea	0.19	-0.01	0.25
Alcoholic Beverages	0.01	-0.08	0.04
Plain Water	0.09	0.31	0
Flavored or Enhanced Water	0.03	-0.04	0.05
Fats & Oils	0.3	-0.08	0.18
Condiments & Sauces	0.18	-0.08	0.01
Sugars	0.18	-0.17	0.24
Baby Foods	-0.03	0.01	-0.01
Baby Beverages	-0.03	-0.04	-0.01
Protein & Nutritional Powders	0.05	0.22	0.04
Not included in a food category	0	0.14	-0.02

Appendix Table 6: Loadings from principal component analysis; over 30 year olds in the National Health and Nutrition Examination Survey 2013-14 (Sensitivity Analysis)

What We Eat in America Food Categories	PC1	PC2	PC3
Milk	-0.16	0.1	-0.5
Flavored Milk	-0.03	0.08	-0.05
Dairy Drinks & Substitutes	-0.04	-0.02	-0.17
Cheese	-0.26	-0.07	0.06
Yogurt	-0.16	-0.25	0.01
Meats	-0.17	0.06	0.19
Poultry	-0.08	-0.04	0.15
Seafood	-0.06	-0.14	0.05
Eggs	-0.13	-0.06	0.15
Cured Meats & Poultry	-0.23	0.12	0.07
Plant-based Protein	-0.09	-0.16	-0.02
Mixed Meat, Poultry or Seafood Dishes	-0.09	0.05	-0.03
Mixed Grain-based Dishes	-0.08	0.09	0.02
Asian Dishes	-0.01	-0.06	-0.03
Mixed Mexican Dishes	0	0.16	-0.11
Pizza	-0.04	0.15	0.01
Mixed Sandwiches	-0.03	0.32	-0.08

Soups	0.01	-0.11	-0.16
Cooked Grains	0.05	-0.19	0.01
What We Eat in America Food Categories (Continued)			
Breads, Rolls, & Tortillas	-0.27	-0.01	-0.02
Quick Breads and Products	-0.11	0.06	0.02
Ready-to-eat Cereals	-0.19	0.07	-0.52
Cooked Cereals	0	-0.13	0.03
Savory Snacks	-0.22	0.14	0.01
Crackers	-0.19	0	-0.08
Snack/Meal Bars	-0.11	-0.04	-0.08
Sweet Bakery Products	-0.2	0.13	0.03
Candy	-0.16	0.11	-0.03
Other Desserts	-0.2	0.03	-0.18
Fruits	-0.16	-0.32	-0.22
Vegetables, excluding Potatoes	-0.26	-0.35	0.09
White Potatoes	-0.18	0.19	0.14
100% Juice	-0.07	-0.05	-0.13
Diet Beverages	-0.16	0.05	0.07
Sweetened Beverages	-0.03	0.41	-0.03
Coffee & Tea	-0.26	0.03	0.21
Alcoholic Beverages	-0.05	0.07	0.27
Plain Water	-0.06	-0.29	-0.08
Flavored or Enhanced Water	-0.1	-0.01	0.03
Fats & Oils	-0.3	-0.07	0.15
Condiments & Sauces	-0.22	-0.01	0.09
Sugars	-0.18	0.15	0.02
Protein & Nutritional Powders	-0.04	-0.09	-0.07
Not included in a food category	0.01	-0.05	-0.03

Appendix Table 7: Distribution of Socio-demographic Characteristics by Age and Number of Decayed Missing and Filled Teeth (DMFT).* Individuals 18 and Over in the National Health and Nutrition Examination Survey (NHANES) III 2013-2014.

	18 to 30 Years of Age n=1259			31 Years of Age and Over[1] n=3784		
	No Caries (<1 DMFT score)	Any Caries (≥1 DMFT score)	(p)	No DMFT (<1 DMFT score)	Any DMFT (≥1 DMFT score)	(p)
Percent of Sample	22.3% (19.6%, 25%)	77.7% (75%, 80.4%)		7.6% (6.6%, 8.5%)	92.4% (91.5%, 93.4%)	
Age*			<0.01			<0.01
18-24 Years	70% (62.3%, 77.7%)	54.3% (50.3%, 58.3%)		NA	NA	
25-30 Years	30% (22.3%, 37.7%)	45.7% (41.7%, 49.7%)		NA	NA	
31-60 Years	NA	NA		78.6% (74.3%, 82.9%)	70.1% (67.9%, 72.3%)	
Over 60 Years	NA	NA		21.4% (17.1%, 25.7%)	29.9% (27.7%, 32.1%)	
Gender			0.13			0.26
Male	57.7% (50.9%, 64.5%)	50% (46%, 54%)		43.5% (36.9%, 50.1%)	47.6% (45.6%, 49.7%)	
Head of Household Education			0.05			0.02
Greater than High school	67.3% (59%, 75.5%)	58.7% (54.5%, 63%)		57.2% (48.5%, 65.8%)	65.4% (60.8%, 70%)	
Ratio of Family Income to Poverty*[2]			0.52			0.03
In poverty (Ratio<1)	23.2% (15.3%, 31%)	25.4% (20.7%, 30.1%)		17% (12.2%, 21.8%)	12.2% (8.9%, 15.4%)	
Average Daily Energy Consumption*			0.86			
Equal or Over 2000 kilocalories	53.7% (45.2%, 62.2%)	52.9% (47.9%, 57.8%)		46.3% (39.2%, 53.5%)	47.2% (44.8%, 49.6%)	0.8
BMI*[3]			0.85			
Normal	44% (35%, 53%)	40.6% (36.5%, 44.7%)		23% (15.2%, 30.8%)	25.3% (23.4%, 27.2%)	0.79
Overweight	25.6% (18.9%, 32.3%)	27.9% (24.2%, 31.6%)		36.3% (29.2%, 43.4%)	33.5% (31.2%, 35.8%)	
Obese (Class 1-3)	26.6% (18.4%, 34.8%)	28.3% (25.2%, 31.4%)		40% (32.7%, 47.2%)	40.4% (37.7%, 43.1%)	

Table Note: Results weighted to represent the United States.

[1]NHANES participants over 80 years of age are top-coded at 80 years of age.

[2]A ratio of family income to poverty <1 indicates a family that is living in poverty.

[3]BMI categories based on Center for Disease Control BMI categories for adults: Underweight: <18.5, Normal: 18.5-24.9, Overweight: 25-29.9, Obese: ≥30. Categories not adjusted for adolescents; adult categories used throughout.

*Variable thus marked are categorized for display purposes in this table only and were parameterized as continuous variables in multivariate models.

Appendix Table 8: Associations of Principal Components (PC) with Any Decayed, Missing and Filled Teeth (DMFT >0) for Quartile Increases in Principal Component (PC) Scores by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014 (Sensitivity Analysis)

Principal Component Models

	18 to 30 Years of Age			Over 30 Years of Age		
	"Carbs & fats" Prevalence Ratio e^{β}, (95% CI)	"Sugar-Sweetened beverages & sandwiches" Prevalence Ratio e^{β}, (95% CI)	"Milk & cereal" Prevalence Ratio e^{β}, (95% CI)	"Carbs & fats" Prevalence Ratio e^{β}, (95% CI)	"Sugar-Sweetened beverages & sandwiches" Prevalence Ratio e^{β}, (95% CI)	"Milk & cereal" Prevalence Ratio e^{β}, (95% CI)
Model 1[1]	0.99 (0.97, 1.01)	1.03* (1.003, 1.06)	0.99 (0.95, 1.04)	1.02* (1.003, 1.03)	1.00 (0.993, 1.01)	0.99* (0.98, 0.999)
Model 2[2]	0.98 (0.94, 1.01)	1.04* (1.002, 1.09)	1.00 (0.94, 1.06)	1.01 (0.99, 1.02)	1.01^{+} (0.998, 1.03)	0.99* (0.98, 0.998)
Model 3[3]	0.98 (0.94, 1.02)	1.04^{+} (0.99, 1.08)	0.99 (0.93, 1.05)	1.01 (0.99, 1.02)	1.01^{+} (0.996, 1.03)	0.99* (0.97, 0.999)
Model 4[4]	0.97 (0.93, 1.02)	1.04 (0.99, 1.09)	1.00 (0.93, 1.06)	1.01 (0.99, 1.03)	1.01+ (0.996, 1.03)	0.99* (0.97, 0.999)

Individual Food Models

	18 to 30 Years of Age			Over 30 Years of Age		
Highest Loading Food Group	Bread Prevalence Ratio	Sweetened Beverages Prevalence Ratio	Cereals Prevalence Ratio	Fats/Oils Prevalence Ratio	Sweetened Beverages Prevalence Ratio	Cereals Prevalence Ratio
Food Categories Model[5]	0.11 (0, 23.67)	1.21 (0.73, 2.01)	0.99 (0.89, 1.11)	2.12 (0.59, 7.62)	0.98 (0.81, 1.19)	0.99 (0.95, 1.03)
Second Highest Loading Food Group	Cheese Prevalence Ratio	Water (Negative Loading) Prevalence Ratio	Milk Prevalence Ratio	Bread Prevalence Ratio	Vegetables (Negative Loading) Prevalence Ratio	Milk Prevalence Ratio
Food Categories Model[5]	0.97 (0.85, 1.11)	1.02 (0.73, 1.44)	1.31 (0.61, 2.80)	0.88 (0.43, 1.83)	1.01 (0.5, 2.01)	0.87 (0.61, 1.25)

*p<0.05 $^{+}$p<0.10

[1] Model 1 included all three Principal Component (PC) variables as quartile ranking variables modeled ordinally. Each prevalence ratio corresponds to the change from one quartile to the next subsequent quartile.

[2] Model 2 contained the three PC variables described above and the following sociodemographic variables: gender, age (continuous), head of household education indicator variable for ≥ high school education and ratio of family income to poverty (continuous).

[3] Model 3 contained all the same variables as Model 2 and the following dietary variables: mean daily energy (dichotomous,< or = 2000 kilocalories vs >2000 kilocalories average), body mass index (continuous variable).

[4] Model 4 was contained the same variables as Model 2 with the addition of the average snacking occasions per day and average breakfast per day variables

[5] The highest and second highest loading food group (based on absolute value) from each principal component were included as predictors in models as percents' gram consumption / total gram consumption averaged over two days. Low gram percentage consumption of cereals and cheese resulted in very wide confidence intervals. These models included the sociodemographic and dietary variables listed in footnotes 2 and 3.

Appendix Table 9: Percent Change in Number of Decayed, Missing and Filled Teeth (DMFT) for each Quartile Increase in Identified Principal Components or Top Loading Food Groups from Principal Components Among those with DMFT>0, by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014 (Sensitivity Analysis)

Principal Component Models						
	18 to 30 Years of Age			Over 30 Years of Age		
	"Carbs & fats" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Sugar-Sweetened beverages & sandwiches" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Milk & cereal" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Carbs & fats" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Sugar-Sweetened beverages & sandwiches" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Milk & cereal" Percent Change $(e^{\beta}-1)*100$ (95% CI)
Model 1[1]	-0.64% (-3.02%, 1.8%)	2.67% (-0.82%, 6.28%)	-1.08% (-4.2%, 2.15%)	2.07%** (0.60%, 3.57%)	0.75% (-0.36%, 1.87%)	1.38%* (0.005%, 2.77%)
Model 2[2]	-0.64% (-3.3%, 2.11%)	2.4% (-1.42%, 6.37%)	-1.16% (-4.45%, 2.24%)	1.44%* (0.37%, 2.56%)	1.88%* (0.54%, 3.23%)	-0.06% (-1.32%, 1.22%)
Model 3[3]	-0.65% (-3.53%, 2.33%)	2.02% (-1.9%, 6.1%)	-1.13% (-4.7%, 2.58%)	1.53%* (0.47%, 2.60%)	1.84%* (0.41%, 3.28%)	-0.08% (-1.38%, 1.23%)
Model 4[4]	-0.95% (-4.21%, 2.41%)	2.76% (-1.8%, 7.53%)	-1.3% (-5.42%, 3%)	1.47%* (0.31%, 2.64%)	2.03%* (0.38%, 3.72%)	-0.24% (-1.68%, 1.22%)
Individual Food Models						
	18 to 30 Years of Age			Over 30 Years of Age		
First Highest Loading Food Group	Bread Percent Change	Sweetened Beverages Percent Change	Cereals Percent Change	Fats/Oils Percent Change	Sweetened Beverages Percent Change	Cereals Percent Change
Food Categories Model[5]	-81.73% (-99.45%, 502.06%)	17.6% (-25.17%, 84.82%)	5.78% (-2.01%, 14.18%)	4.23%$^{+}$ (-0.44%, 9.12%)	25.31%* (3.13%, 52.30%)	-2.19%$^{+}$ (-4.99%, 0.69%)
Second Highest Loading Food Group	Cheese Percent Change	Water (Negative Loading) Percent Change	Milk Percent Change	Bread Percent Change	Vegetables (Negative Loading) Percent Change	Milk Percent Change
Food Categories Model[5]	-0.6% (-8.37%, 7.83%)	5.61% (-21.77%, 42.57%)	-24.57% (-59.35%, 39.98%)	-42.54% (-79.12%, 58.15%)	2.69% (-36.12%, 65.10%)	13.75% (-18.69%, 59.12%)

**p<0.01 *p<0.05 $^{+}$p<0.10

[1] Model 1 included all three Principal Component (PC) variables as quartile ranking variables modeled ordinally. Each exponentiated coefficient corresponds to the change from one quartile to the next subsequent quartile.
[2] Model 2 contained the three PC variables described above and the following sociodemographic variables: gender, age (continuous), head of household education indicator variable for ≥ high school education and ratio of family income to poverty (continuous).
[3] Model 3 contained all the same variables as Model 2 and the following dietary variables: mean daily energy (dichotomous,< or = 2000 kilocalories vs >2000 kilocalories average), body mass index (continuous variable).
[4] Model 4 was contained the same variables as Model 2 with the addition of the average snacking occasions per day and average breakfast per day variables
[5] The highest and second highest loading food group (based on absolute value) from each principal component were included as predictors in models as percents' gram consumption / total gram consumption averaged over two days. Low gram percentage consumption of cereals and cheese resulted in very wide confidence intervals. These models included the sociodemographic and dietary variables listed in footnotes 2 and 3.

Alternative exposure variable parameterization for statistical models in main analysis

Appendix Tables 10 and 11 show the same models described in the main body of the text, with the exception that Principal Components are no longer modeled as ordinal quartiles but are instead modeled as continuously. We chose to use ordinal quartiles for the main models as they are more interpretable than the continuous principal component coordinates. The only significant difference between using continuous principal components coordinates and using ordinal quartile rankings is that in over-30-year-olds, Principal Component 1 is no longer a significant predictor of increasing severity of dental decay (Appendix Table 10). Concerned that this may be due to a non-linear relationship between this Principal Component and the outcome, we examined partial regression plots for the log linear models in over-30-year-olds (Appendix Figure 5). However, no non-linear relationship was apparent.

Appendix Table 10: Associations of Principal Components (PC) with Any Decayed, Missing and Filled Teeth (DMFT >0) for Continuous PC Scores by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014

Principal Component Models						
	18 to 30 Years of Age			Over 30 Years of Age		
	"Carbs & fats" Prevalence Ratio e^{β}, (95% CI)	"Sugar-Sweetened beverages & sandwiches" Prevalence Ratio e^{β}, (95% CI)	"Milk & cereal" Prevalence Ratio e^{β}, (95% CI)	"Carbs & fats" Prevalence Ratio e^{β}, (95% CI)	"Sugar-Sweetened beverages & sandwiches" Prevalence Ratio e^{β}, (95% CI)	"Milk & cereal" Prevalence Ratio e^{β}, (95% CI)
Model[1]	1.00 (0.98, 1.01)	1.01 (0.99, 1.03)	0.99 (0.96, 1.02)	1 (0.98, 1.01)	1.01 (0.99, 1.03)	0.99 (0.96, 1.02)
Model 2[2]	1.00 (0.98, 1.02)	1.02 (0.99, 1.06)	1.01 (0.98, 1.04)	1 (0.99, 1.01)	1.01* (1.004, 1.02)	0.99^{+} (0.98, 1.002)
Model 3[3]	1.00 (0.98, 1.02)	1.02 (0.98, 1.06)	1.01 (0.97, 1.04)	1 (0.99, 1.01)	1.01* (1.003, 1.02)	0.99^{+} (0.98, 1.002)
Model 4[4]	1.00 (0.97, 1.02)	1.03 (0.98, 1.07)	1.01 (0.97, 1.04)	1 (0.99, 1.01)	1.01* (1.004, 1.02)	0.99 (0.98, 1.003)

*p<0.05 $^{+}$p<0.10

[1] Model 1 included all three Principal Component (PC) variables as quartile ranking variables modeled ordinally. Each prevalence ratio corresponds to the change from one quartile to the next subsequent quartile.

[2] Model 2 contained the three PC variables described above and the following sociodemographic variables: gender, age (continuous), head of household education indicator variable for ≥ high school education and ratio of family income to poverty (continuous).

[3] Model 3 contained all the same variables as Model 2 and the following dietary variables: mean daily energy (dichotomous,< or = 2000 kilocalories vs >2000 kilocalories average), body mass index (continuous variable).

[4] Model 4 was contained the same variables as Model 2 with the addition of the average snacking occasions per day and average breakfast per day variables

[5] The highest and second highest loading food group (based on absolute value) from each principal component were included as predictors in models as percent gram consumption / total gram consumption averaged over two days. Low gram percentage consumption of cereals and cheese resulted in very wide confidence intervals. These models included the sociodemographic and dietary variables listed in footnotes 2 and 3.

Appendix Table 11: Percent Change in Number of Decayed, Missing and Filled Teeth (DMFT) for Continuous Principal Component Coordinates Among those with DMFT>0, by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014 (Main Analysis)

Principal Component Models

	18 to 30 Years of Age			Over 30 Years of Age		
	"Carbs & fats" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Sugar-Sweetened beverages & sandwiches" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Milk & cereal" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Carbs & fats" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Sugar-Sweetened beverages & sandwiches" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Milk & cereal" Percent Change $(e^{\beta}-1)*100$ (95% CI)
Model[1]	-0.39% (-2.21%, 1.46%)	2.73% (0.14%, 5.38%)	-0.2% (-2.78%, 2.44%)	1.03% (0.08%, 1.98%)	1.14% (-0.09%, 2.39%)	-0.36% (-2.63%, 1.97%)
Model 2[2]	-0.61% (-2.47%, 1.28%)	2.19% (-1.21%, 5.7%)	-0.89% (-4.22%, 2.55%)	1.02% (-0.03%, 2.07%)	1.97% (0.44%, 3.51%)	-0.84% (-2.47%, 0.81%)
Model 3[3]	-0.73% (-2.66%, 1.23%)	1.99% (-1.66%, 5.78%)	-0.86% (-4.32%, 2.72%)	1.07% (-0.04%, 2.2%)	1.88% (0.32%, 3.46%)	-0.79% (-2.47%, 0.92%)
Model 4[4]	-0.96% (-3.25%, 1.39%)	2.78% (-0.91%, 6.6%)	-0.94% (-4.69%, 2.96%)	1.02% (-0.26%, 2.31%)	2.05% (0.3%, 3.83%)	-1.03% (-2.93%, 0.91%)

**p<0.01 *p<0.05 $^{+}$p<0.10

[1] Model 1 included all three Principal Component (PC) variables as quartile ranking variables modeled ordinally. Each exponentiated coefficient corresponds to the change from one quartile to the next subsequent quartile.

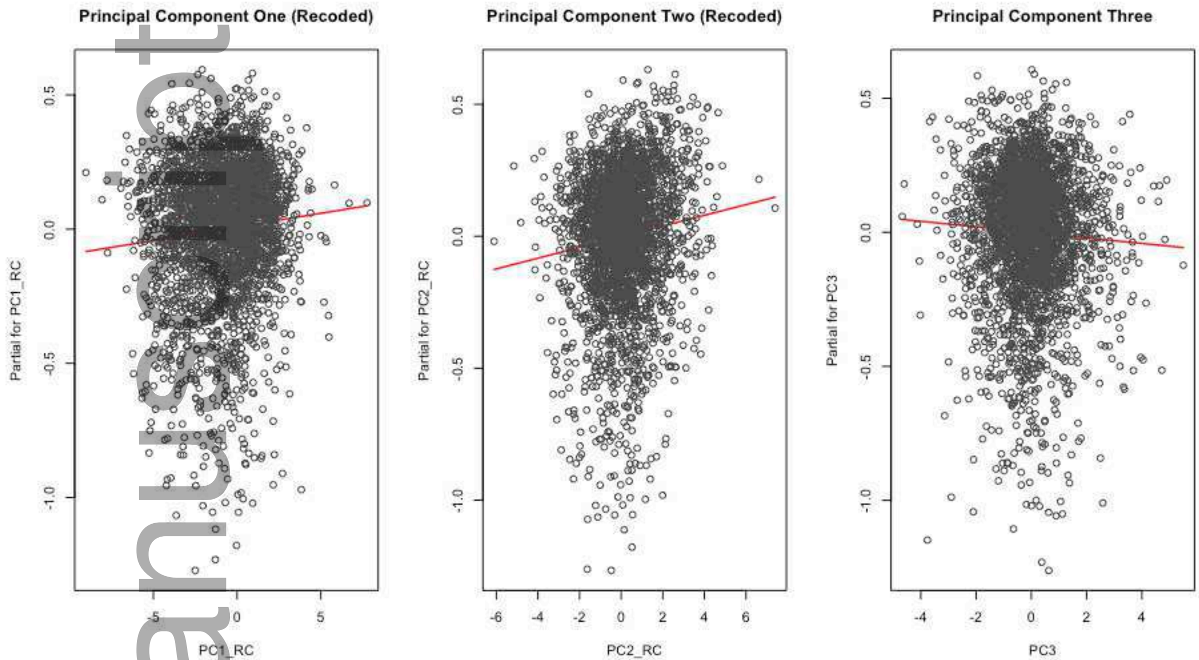
[2] Model 2 contained the three PC variables described and the following sociodemographic variables: gender, age (continuous), head of household education indicator variable for ≥ high school education and ratio of family income to poverty (continuous).

[3] Model 3 contained all the same variables as Model 2 and the following dietary variables: mean daily energy (dichotomous,< or = 2000 kilocalories vs >2000 kilocalories average), body mass index (continuous variable).

[4] Model 4 was contained the same variables as Model 2 with the addition of the average snacking occasions per day and average breakfast per day variables

[5] The highest and second highest loading food group (based on absolute value) from each principal component were included as predictors in models as percent gram consumption / total gram consumption averaged over two days. Low gram percentage consumption of cereals and cheese resulted in very wide confidence intervals. These models included the sociodemographic and dietary variables listed in footnotes 2 and 3.

Appendix Figure 5: Partial Regression Plots of Continuous Principal Component coordinates in Log Linear Models for Over 30 Year Olds from National Health and Nutrition Examination Survey 2013-14 (Main Analysis)



Sensitivity analysis: Separation of high- and low-sugar cereals:

For our main analysis, we used broader food categories which encompass several "What We Eat in America" (WWEIA) food categories. For example, the food group "Milk" which we use includes the WWEIA food categories "Milk, whole", "Milk, reduced fat", "Milk, low-fat" and "Milk, nonfat" (See Appendix Table 2). Although the known cariogenicity of WWEIA food categories are generally similar within these broader food groups, for the "Ready-to-eat-cereal" food group, this may not be the case, as this food group contains the WWEIA food categories "Ready-to-eat-cereal (higher sugar)" and "Ready-to-eat-cereal (lower sugar)", which may have different cariogenicity due to the differing sugar content. To address this, we carried out a sensitivity analysis in which the "Ready-to-eat-cereal" food group is split into its respective food categories of "Ready-to-eat-cereal (higher sugar)" and "Ready-to-eat-cereal (lower sugar)" and the PCA is rerun. All other food groups remain the same. Appendix tables 12 and 13 show the loadings of the first three PCs from this sensitivity analysis. The highly-loading food groups characterizing each PC remain essentially the same as in the main analysis. "Ready-to-eat-cereal (higher sugar)" and "Ready-to-eat-cereal (lower sugar)" both load most strongly in the third PC for 18-30-year-olds and over-30-year-olds. Appendix tables 14 and 15 show results of multivariate models using these new PC quartiles; again, results are similar to the main analysis. However, low sugar cereals now appear to be protective for severity of dental decay in the individual food log-linear model.

Appendix Table 12: Loadings from principal component analysis; 18-30 year olds in the National Health and Nutrition Examination Survey 2013-14 (Sensitivity Analysis - Cereals)

	PC1	PC2	PC3
Milk	0.18	-0.13	0.31
Flavored Milk	0.08	0.07	0.27
Dairy Drinks & Substitutes	0.06	-0.22	0.01
Cheese	0.29	0.09	0.08
Yogurt	0.08	-0.19	0.13
Meats	0.16	-0.01	-0.11
Poultry	0.14	-0.03	-0.19
Seafood	0.03	-0.05	-0.04
Eggs	0.18	-0.18	-0.22
Cured Meats & Poultry	0.25	0.08	0.16
Plant-based Protein	0.14	-0.22	-0.08
Mixed Meat, Poultry or Seafood Dishes	0.06	0.01	0.11
Mixed Grain-based Dishes	0.09	0.14	0.15
Asian Dishes	0.02	-0.05	0.09
Mixed Mexican Dishes	0.01	0.08	0.06
Pizza	0.09	0.2	0
Mixed Sandwiches	0.03	0.23	-0.1
Soups	0.05	0.04	0.07
Cooked Grains	0.05	-0.22	-0.1
Breads, Rolls, & Tortillas	0.34	0.01	0.14
Quick Breads and Products	0.2	0.16	-0.13
Ready-to-eat Cereals (lower sugar)	0.11	-0.03	0.35
Ready-to-eat Cereals (higher sugar)	0.09	-0.2	0.17
Cooked Cereals	0.07	-0.11	0.1
Savory Snacks	0.19	0.13	0.11
Crackers	0.2	0.02	-0.1
Snack/Meal Bars	0.07	-0.09	-0.01
Sweet Bakery Products	0.15	0.07	0.08
Candy	0.14	0.16	0.1
Other Desserts	0.11	0.03	0.19
Fruits	0.16	-0.3	0.09
Vegetables, excluding Potatoes	0.24	-0.21	-0.17
White Potatoes	0.16	0.18	-0.08
100% Juice	0.14	0.04	0.26
Diet Beverages	0.03	0.09	0.02
Sweetened Beverages	0.11	0.34	-0.01
Coffee & Tea	0.22	0.01	-0.29
Alcoholic Beverages	0.05	0.05	-0.07
Plain Water	0.13	-0.26	0.04
Flavored or Enhanced Water	0.04	0.04	-0.05
Fats & Oils	0.25	0.05	-0.24
Condiments & Sauces	0.21	0.01	-0.16
Sugars	0.21	0.15	-0.16
Baby Foods	-0.02	0	0
Baby Beverages	0.01	0.1	-0.05
Protein & Nutritional Powders	0.06	-0.26	-0.06
Not included in a food category	0.04	-0.18	-0.08

Appendix Table 13: Loadings from principal component analysis; over 30 year olds in the National Health and Nutrition Examination Survey 2013-14 (Sensitivity Analysis – Cereals)

	PC1	PC2	PC3
Milk	-0.17	0.11	-0.38
Flavored Milk	-0.03	0.04	-0.11
Dairy Drinks & Substitutes	-0.05	-0.11	-0.24
Cheese	-0.24	-0.06	0.06
Yogurt	-0.14	-0.23	-0.05
Meats	-0.17	0.03	0.21
Poultry	-0.11	-0.07	0.11
Seafood	-0.07	-0.14	0.07
Eggs	-0.14	-0.07	0.16
Cured Meats & Poultry	-0.2	0.14	0.05
Plant-based Protein	-0.11	-0.18	-0.01
Mixed Meat, Poultry or Seafood Dishes	-0.12	0.08	-0.04
Mixed Grain-based Dishes	-0.1	0.07	0.05
Asian Dishes	-0.04	-0.11	0.01
Mixed Mexican Dishes	-0.04	0.12	-0.07
Pizza	-0.09	0.19	-0.11
Mixed Sandwiches	-0.08	0.34	-0.07
Soups	-0.04	-0.13	-0.11
Cooked Grains	-0.01	-0.27	-0.01
Breads, Rolls, & Tortillas	-0.27	0.01	-0.02
Quick Breads and Products	-0.12	-0.01	0.15
Ready-to-eat Cereals (lower sugar)	-0.15	0.1	-0.35
Ready-to-eat Cereals (higher sugar)	-0.14	0	-0.28
Cooked Cereals	0	-0.1	0.07
Savory Snacks	-0.21	0.14	-0.03
Crackers	-0.2	0.04	-0.12
Snack/Meal Bars	-0.14	-0.02	-0.18
Sweet Bakery Products	-0.2	0.11	0.09
Candy	-0.17	0.15	-0.08
Other Desserts	-0.18	0.07	-0.11
Fruits	-0.17	-0.28	-0.25
Vegetables, excluding Potatoes	-0.24	-0.35	0.07
White Potatoes	-0.18	0.16	0.16
100% Juice	-0.09	-0.05	-0.08
Diet Beverages	-0.15	0.12	-0.01
Sweetened Beverages	-0.08	0.38	0
Coffee & Tea	-0.24	0	0.27
Alcoholic Beverages	-0.08	0.02	0.25
Plain Water	-0.15	-0.24	-0.13
Flavored or Enhanced Water	-0.1	-0.07	-0.04
Fats & Oils	-0.27	-0.07	0.16
Condiments & Sauces	-0.21	-0.05	0.14
Sugars	-0.18	0.08	0.2
Protein & Nutritional Powders	-0.03	-0.08	-0.13
Not included in a food category	0	-0.05	-0.02

Appendix Table 14: Associations of Principal Components (PC) with Any Decayed, Missing and Filled Teeth (DMFT >0) for Quartile Increases in Principal Component (PC) Scores by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014 (Sensitivity Analysis - Cereals)

Principal Component Models

	18 to 30 Years of Age			Over 30 Years of Age		
	"Carbs & fats" Prevalence Ratio e^{β}, (95% CI)	"Sugar-Sweetened beverages & sandwiches" Prevalence Ratio e^{β}, (95% CI)	"Milk & cereal" Prevalence Ratio e^{β}, (95% CI)	"Carbs & fats" Prevalence Ratio e^{β}, (95% CI)	"Sugar-Sweetened beverages & sandwiches" Prevalence Ratio e^{β}, (95% CI)	"Milk & cereal" Prevalence Ratio e^{β}, (95% CI)
Model 1[1]	0.99 (0.96, 1.02)	1.02 (0.98, 1.06)	0.98 (0.94, 1.02)	1.01 (0.997, 1.03)	1 (0.99, 1.02)	0.99 (0.98, 1.001)
Model 2[2]	0.99 (0.94, 1.03)	1.02 (0.98, 1.07)	0.99 (0.95, 1.04)	1 (0.99, 1.02)	1.01 (1.002, 1.03)*	0.99 (0.97, 1.003)
Model 3[3]	0.99 (0.94, 1.03)	1.02 (0.98, 1.08)	0.99 (0.94, 1.04)	1 (0.99, 1.02)	1.02 (1.001, 1.03)*	0.99 (0.97, 1.004)
Model 4[4]	0.98 (0.93, 1.04)	1.03 (0.97, 1.1)	0.99 (0.94, 1.04)	1 (0.98, 1.03)	1.02 (1.001, 1.03)*	0.99 (0.97, 1.005)

Individual Food Models

	18 to 30 Years of Age			Over 30 Years of Age		
Highest Loading Food Group	Bread Prevalence Ratio	Sweetened Beverages Prevalence Ratio	Cereals – lower sugar Prevalence Ratio	Bread Prevalence Ratio	Sweetened Beverages Prevalence Ratio	Milk Prevalence Ratio
Food Categories Model[5]	1.08 (0.01, 88.26)	1.12 (0.7, 1.79)	1.00 (0.91, 1.11)	0.93 (0.33, 2.65)	1.04 (0.83, 1.3)	0.82 (0.52, 1.29)
Second Highest Loading Food Group	Cheese Prevalence Ratio	Fruit (Negative Loading) Prevalence Ratio	Milk Prevalence Ratio	Fats/Oils Prevalence Ratio	Vegetables (Negative Loading) Prevalence Ratio	Cereals-lower sugar Prevalence Ratio
Food Categories Model[5]	0.97 (0.86, 1.09)	0.22 (0.01, 8.99)	1.88 (0.53, 6.61)	2.47 (0.64, 9.53)	1.02 (0.49, 2.12)	1.00 (0.96, 1.05)

$^{*}p<0.05$ $^{+}p<0.10$

[1] Model 1 included all three Principal Component (PC) variables as quartile ranking variables modeled ordinally. Each prevalence ratio corresponds to the change from one quartile to the next subsequent quartile.

[2] Model 2 contained the three PC variables described above as well as sociodemographic variables: gender, age (continuous), head of household education indicator variable for ≥ high school education and ratio of family income to poverty (continuous).

[3] Model 3 contained all the same variables as Model 2 and additionally dietary variables: mean daily energy (dichotomous,< or = 2000 kilocalories vs >2000 kilocalories average), body mass index (continuous variable).

[4] Model 4 was contained the same variables as Model 2 with the addition of the average snacking occasions per day and average breakfast per day variables

[5] The highest and second highest loading food group (based on absolute value) from each principal component were included as predictors in models as percents' gram consumption / total gram consumption averaged over two days. Low gram percentage consumption of cereals and cheese resulted in very wide confidence intervals. These models included the sociodemographic and dietary variables listed in footnotes 2 and 3.

Appendix Table 15: Percent Change in Number of Decayed, Missing and Filled Teeth (DMFT) for each Quartile Increase in Identified Principal Components or Top Loading Food Groups from Principal Components Among those with DMFT>0, by Age Group. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014 (Sensitivity Analysis - Cereals)

Principal Component Models						
	18 to 30 Years of Age			Over 30 Years of Age		
	"Carbs & fats" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Sugar-Sweetened beverages & sandwiches" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Milk & cereal" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Carbs & fats" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Sugar-Sweetened beverages & sandwiches" Percent Change $(e^{\beta}-1)*100$ (95% CI)	"Milk & cereal" Percent Change $(e^{\beta}-1)*100$ (95% CI)
Model 1[1]	-2.12% (-5.75%, 1.64%)	3.37% (-0.58%, 7.49%)	-0.37% (-3.79%, 3.18%)	2.43% (1.09%, 3.78%)**	1.24% (-0.17%, 2.68%)$^{+}$	0.26% (-1.66%, 2.22%)
Model 2[2]	-2.49% (-6.19%, 1.36%)	2.66% (-2.28%, 7.85%)	-1.03% (-4.65%, 2.72%)	2.2% (0.76%, 3.66%)**	1.74% (0.16%, 3.34%)*	-0.47% (-1.96%, 1.04%)
Model 3[3]	-2.71% (-6.47%, 1.19%)	2.41% (-2.76%, 7.85%)	-0.92% (-4.73%, 3.04%)	2.25% (0.77%, 3.75%)**	1.63% (0.01%, 3.27%)*	-0.45% (-1.95%, 1.07%)
Model 4[4]	-3.16% (-7.68%, 1.59%)	3.52% (-2.23%, 9.61%)	-1.05% (-5.28%, 3.36%)	2.2% (0.5%, 3.93%)*	1.8% (-0.04%, 3.67%)$^{+}$	-0.67% (-2.34%, 1.03%
Individual Food Models						
	18 to 30 Years of Age			Over 30 Years of Age		
First Highest Loading Food Group	Bread Percent Change	Sweetened Beverages Percent Change	Cereals – low sugar Percent Change	Breads Percent Change	Sweetened Beverages Percent Change	Cereals – low sugar Percent Change
Food Categories Model[5]	-92.39% (-99.87%, 355.1%)	9.01% (-29.99%, 69.73%)	5.13% (-6.13%, 17.74%)	5.16% (-2.52%, 13.45%)	24.78% (-1.02%, 57.32%)$^{+}$	-4.24% (-7.41%, -0.97%)*
Second Highest Loading Food Group	Cheese Percent Change	Fruit (Negative Loading) Percent Change	Milk Percent Change	Fats/oils Percent Change	Vegetables (Negative Loading) Percent Change	Milk Percent Change
Food Categories Model[5]	1.16% (-9.76%, 13.41%)	-17.31% (-83.78%, 321.67%)	-0.77% (-65.05%, 181.71%)	-53.75% (-84.09%, 34.42%)	-8.83% (-49.01%, 63.01%)	13.5% (-24.73%, 71.15%)

**p<0.01 *p<0.05 $^{+}$p<0.10

[1] Model 1 included all three Principal Component (PC) variables as quartile ranking variables modeled ordinally. Each exponentiated coefficient corresponds to the change from one quartile to the next subsequent quartile.
[2] Model 2 contained the three PC variables described above the following sociodemographic variables: gender, age (continuous), head of household education indicator variable for ≥ high school education and ratio of family income to poverty (continuous).
[3] Model 3 contained all the same variables as Model 2 and the following dietary variables: mean daily energy (dichotomous,< or = 2000 kilocalories vs >2000 kilocalories average), body mass index (continuous variable).
[4] Model 4 was contained the same variables as Model 2 with the addition of the average snacking occasions per day and average breakfast per day variables
[5] The highest and second highest loading food group (based on absolute value) from each principal component were included as predictors in models as percents' gram consumption / total gram consumption averaged over two days. Low gram percentage consumption of cereals and cheese resulted in very wide confidence intervals. These models included the sociodemographic and dietary variables listed in footnotes 2 and 3.

Sensitivity analysis: Utilization of a negative binomial model with dichotomous parameterization of average energy variable

Negative binomial model (Table 16): Since DMFT is over dispersed in both the younger and older analysis sets (18-30-year-olds mean: 5.2, variance: 24.4 and > 30 year olds mean: 12.1, variance:50.9), a negative binomial model was more appropriate than a Poisson model. Due to convergence issues when the average energy variable was parameterized continuously, this variable was instead parameterized dichotomously, as either less than or equal to 2000 or greater than 2000.

In 18-30-year-olds, membership in each subsequently higher quartile of a diet high in breads and oils was associated with on average, 0.91 (95% CI: 0.84, 0.98) times more DMFT. Additionally, membership in each subsequent quartile of a diet high in sugar-sweetened beverages and sandwiches were associated with, on average, 1.12 (95% CI: 1.03, 1.21) times more DMFT. Bread, rolls and tortillas was the only food group from the highest loading food groups which associated with DMFT (exponentiated coefficient=0.003, 95% CI: (0.0001, 0.06)).

In over-30-year-olds, membership in each subsequently higher quartile of a diet high in breads and oils was associated with on average, 1.04 (95% CI: 1.01, 1.07) times more DMFT. Membership in each subsequent quartile of a diet high in sugar-sweetened beverages and sandwiches were associated with, on average, 1.06 (95% CI: 1.04, 1.1) times more DMFT. Membership in each subsequent quartile of a diet high in milk and cereal was associated with, on average, 0.98 (95% CI: 0.96, 1) times more DMFT. When we examined the individual food groups which loaded strongly in the PCs, higher gram percentage of sugar-sweetened beverages consumption associated with count of DMFT (=1.7, 95% CI: 1.3, 2.2) as did reporting any fats/oil consumption (=1.1, 95% CI: 1, 1.3). Higher gram percentage of bread consumption associated inversely with count of DMFT (=0.2, 95% CI: 0.1, 0.6), as did reporting any cereal consumption (=0.9, 95% CI: 0.9, 1).

Table 16: Associations of Principal Components (PC) with Count of Decayed, Missing and Filled Teeth for Quartile Increases in Principal Component (PC) Scores by Age Group using a negative binomial model. Participants in the National Health and Nutrition Examination Survey (NHANES) 2013-2014

Principal Component Models

	18 to 30 Years of Age			Over 30 Years of Age		
	"Breads & fats" e^{β}, (95% CI)	"Sugar-Sweetened beverages & sandwiches" e^{β}, (95% CI)	"Milk & cereal" e^{β}, (95% CI)	"Breads & fats" e^{β}, (95% CI)	"Sugar-Sweetened beverages & sandwiches" e^{β}, (95% CI)	"Milk & cereal" e^{β}, (95% CI)
Model 1[1]	0.95 (0.88, 1.03)	1.09 (1.02, 1.17)*	0.99 (0.9, 1.08)	1.05 (1.03, 1.08)***	1.05 (1.02, 1.07)***	0.99 (0.96, 1.02)
Model 2[2]	0.93 (0.86, 0.997)*	1.08 (0.997, 1.16)$^{+}$	0.99 (0.92, 1.06)	1.04 (1.01, 1.07)*	1.06 (1.03, 1.1)***	0.98 (0.96, 1.001)$^{+}$
Model 3[3]	0.93 (0.87, 0.996)*	1.08 (1, 1.16)*	0.99 (0.92, 1.06)	1.04 (1.01, 1.07)*	1.06 (1.03, 1.09)***	0.98 (0.96, 1)*
Model 4[4]	0.91 (0.84, 0.98)*	1.12 (1.03, 1.21)**	0.99 (0.93, 1.05)	1.04 (1.01, 1.07)*	1.06 (1.03, 1.1)***	0.98 (0.96, 0.995)*

Individual Food Models

	18 to 30 Years of Age			Over 30 Years of Age		
Highest Loading Food Group	Bread Prevalence Ratio	Sweetened Beverages Prevalence Ratio	Cereals Prevalence Ratio	Fats/Oils Prevalence Ratio	Sweetened Beverages Prevalence Ratio	Cereals Prevalence Ratio
Food Categories Model[5]	0.003 (0.0001, 0.06)**	1.34 (0.76, 2.33)	1.03 (0.91, 1.18)	1.13 (1.02, 1.26)*	1.70 (1.31, 2.22)***	0.93 (0.88, 0.99)*
Second Highest Loading Food Group	Cheese Prevalence Ratio	Fruits (Negative Loading) Prevalence Ratio	Milk Prevalence Ratio	Bread Prevalence Ratio	Vegetables (Negative Loading) Prevalence Ratio	Milk Prevalence Ratio
Food Categories Model[5]	0.98 (0.85, 1.12)	0.16 (0.004, 5.92)	1.98 (0.59, 6.70)	0.21 (0.07, 0.63)**	0.69 (0.27, 1.80)	1.18 (0.64, 2.16)

***p<0.001**<0.01*p<0.05 $^{+}$p<0.10

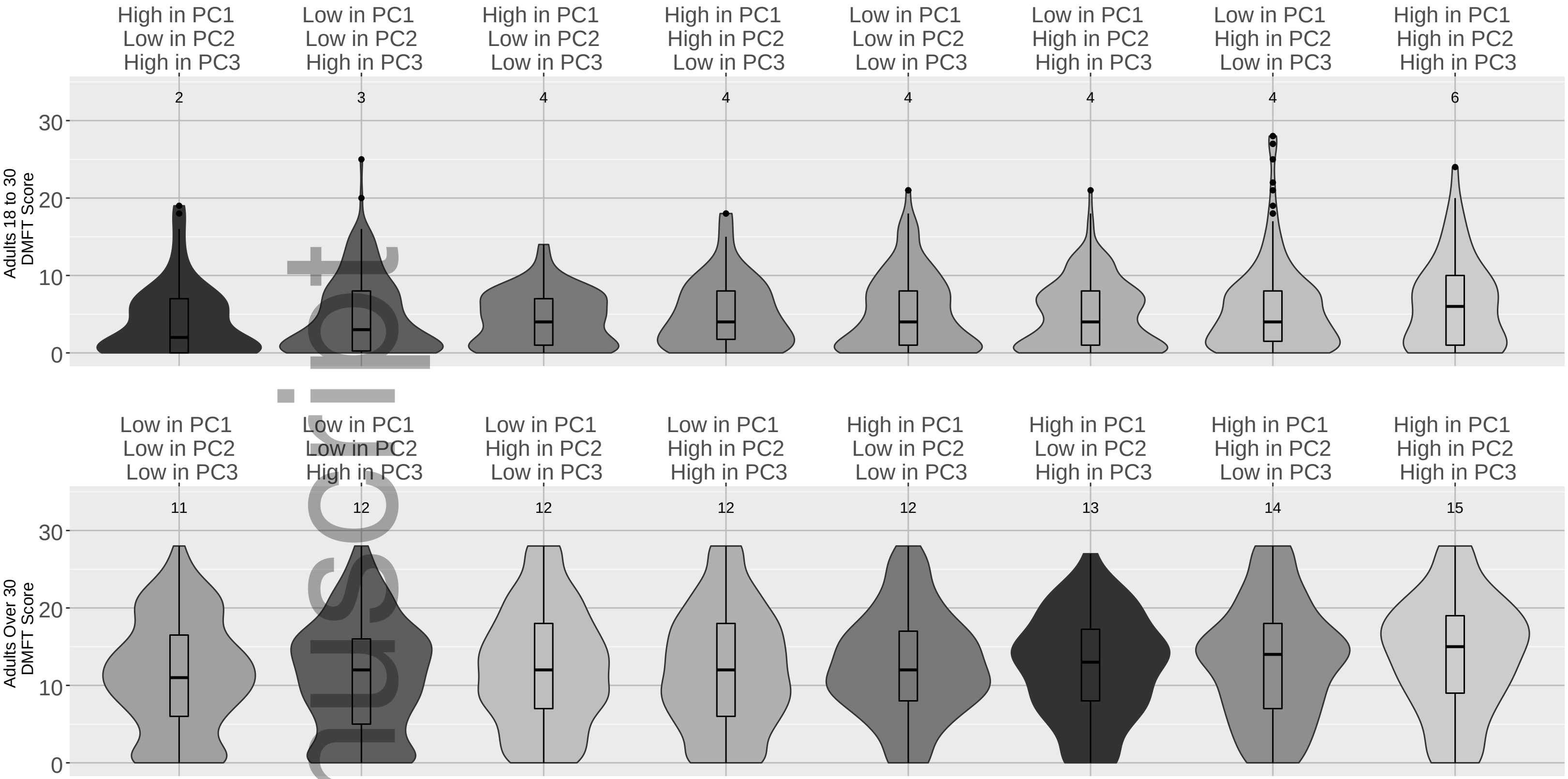
[1] Model 1 included all three Principal Component (PC) variables as quartile ranking variables modeled ordinally. Each prevalence ratio corresponds to the change from one quartile to the next subsequent quartile.

[2] Model 2 contained the three PC variables described above and the following sociodemographic variables: gender, age (continuous), head of household education indicator variable for ≥ high school education and ratio of family income to poverty (continuous).

[3] Model 3 contained all the same variables as Model 2 and the following dietary variables: mean daily energy (dichotomous,< or = 2000 kilocalories vs >2000 kilocalories average), body mass index (continuous variable).

[4] Model 4 was contained the same variables as Model 2 with the addition of the average snacking occasions per day and average breakfast per day variables

[5] The highest and second highest loading food group (based on absolute value) from each principal component were included as predictors in models as percents' gram consumption / total gram consumption averaged over two days. Low gram percentage consumption of cereals and cheese resulted in very wide confidence intervals. These models included the sociodemographic and dietary variables listed in footnotes 2 and 3.



cdoe_12509_f1.eps