

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- | | | |
|-------------------------------------|-------------------------------------|--|
| ☐ | ☒ | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☒ | ☐ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☒ | ☐ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☐ | ☒ | Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection This study analyzed secondary data; no software was used for data collection

Data analysis Stata MP version 17. Code available at: <https://github.com/clee321/NATFOOD-dietarysubs>

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

This study used data from the National Health and Nutrition Examination Surveys (NHANES) and from the Database of Food Recall Impacts on the Environment for Nutrition and Dietary Studies (dataFRIENDS). NHANES data are publicly available on the Centers for Disease Control website (<https://wwwn.cdc.gov/nchs/nhanes/>). The dataFRIENDS data are currently available from the authors upon request while analyses of these data are ongoing. The database on which dataFRIENDS was derived is publicly available at <https://css.umich.edu/page/datafield>.

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research.](#)

Reporting on sex and gender

The NHANES data include a variable on participants' gender. Interviewers ask participants about their "gender" "if not obvious" with response options of: male, female, don't know, and refused to answer.

The sample included 3,793 participants who identified as male and 3,960 participants who identified as female.

We examined and reported outcomes overall and disaggregated by gender (males vs. females).

Population characteristics

See research sample description below.

Recruitment

The NHANES sample represents the non-institutionalized civilian population residing in the 50 states and the District of Columbia. Since 1999, the sample design has consisted of multi-year, stratified, clustered four-stage samples, with data released in 2-year cycles. The NHANES sample is drawn in four stages: (a) PSUs (counties, groups of tracts within counties, or combinations of adjacent counties), (b) segments within PSUs (census blocks or combinations of blocks), (c) dwelling units (DUs) (households) within segments, and (d) individuals within households. PSUs are sampled from all U.S. counties. Screening is conducted at the DU level to identify sampled persons (SPs), based on oversampling criteria. Sampled participants are recruited in person by trained interviewers.

Ethics oversight

The Harvard Longwood Campus Institutional Review Board determined this study to be exempt (IRB #21-1624) and the Stanford University Institutional Review Board approved this study (IRB #67669).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences

☒ Behavioural & social sciences

☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

This is a simulation study using cross-sectional, quantitative survey data.

Research sample

The study used a population-based sample of US children and adults aged 2 and older. Participants were respondents in the National Health and Nutrition Examination Surveys (NHANES), a continuous, multistage probability survey that is nationally representative of the civilian, non-institutionalized US population. The study used data from the 2015-2016 NHANES cycle, the most recent cycle for which greenhouse gas emissions data were available.

The weighted sample was 51.5% female and 48.5% male. About 4% of participants were aged 2-5 years, 7% were 6-11 years, 9% were 12-17 years, 10% were 18-25 years, 19% were 26-39 years, 14% were 40-49 years, 15% were 50-59 years, and 22% were 60+ years. About 63% identified as non-Hispanic white, 12% as non-Hispanic Black, 17% as Mexican American or another Hispanic ethnicity, 5% as non-Hispanic Asian, and 4% as another race/ethnicity.

We chose this sample to enable calculation of nationally representative estimates of dietary quality and dietary carbon footprints.

Sampling strategy

NHANES uses a complex, multistage, probability sampling design to select participants representative of the civilian, non-institutionalized US population. In each cycle, oversampling of certain population subgroups is also done to increase the reliability and precision of health status indicator estimates for these particular subgroups. The NHANES sample represents the non-institutionalized civilian population residing in the 50 states and the District of Columbia. Since 1999, the sample design has consisted of multi-year, stratified, clustered four-stage samples, with data released in 2-year cycles. The NHANES sample is drawn in four stages: (a) PSUs (counties, groups of tracts within counties, or combinations of adjacent counties), (b) segments within PSUs (census blocks or combinations of blocks), (c) dwelling units (DUs) (households) within segments, and (d) individuals within households. PSUs are sampled from all U.S. counties. Screening is conducted at the DU level to identify sampled persons (SPs), based on oversampling criteria. All analyses adjusted for the complex survey design.

No power analyses were conducted for this study given our use of secondary data. Instead, we determined that the sample of approximately 7,700 participants from the most recent cycle of NHANES for which we had greenhouse gas emission data would be sufficient for our analyses by examining whether samples of this size yielded sufficient precision in similar analyses. The NHANES 2015-2016 cycle includes 7,753 participants with valid dietary recall data. Our primary outcomes -- Healthy Eating Index scores and dietary carbon footprints -- can be measured with high precision in samples of this size, providing estimates of changes in Healthy

Eating Index scores and dietary carbon footprints with narrow confidence intervals. For example, a recent analysis of NHANES found that if a subsample of 1,026 adults interested in dietary change replaced their beef intake with poultry, average Healthy Eating Index scores would increase by 0.88 points (95% CI: 0.76, 1.00) and participants' dietary carbon footprint would decline by 1.38 kg of CO₂ equivalents/day (95% CI: -1.58, -1.19) (see Willits-Smith 2020, doi: 10.1016/S2542-5196(20)30055-3).

Data collection

We used demographic and dietary intake data from NHANES.

Demographic data are collected by a trained interviewer in the participant's home. Interviewers collect and encrypt data on laptops, using printed materials to prompt and verify responses.

Dietary intake data are from a 24-hour dietary recall interview that is conducted for participants of all ages. Dietary recall interviews are conducted in person by trained dietary interviewers fluent in Spanish and English. Dietary recalls are collected using the Automated Multiple Pass Method. A proxy adult assists children aged 6–11 years and completes the recall for children aged ≤5 years. The setting of the interview is a private room in the Mobile Examination Center (MEC). Interviewers enter data on a laptop. Details are available here: https://www.cdc.gov/nchs/nhanes/measuring_guides_dri/measuringguides.htm.

Timing

Data collection occurred in 2015–2016.

Data exclusions

The initial sample included 8,505 participants. We excluded participants younger than 2 years of age (n=583, ~7% of the total sample). We additionally excluded participants who did not complete a valid 24-hour dietary recall (n=169, ~2% of the total sample). Our analytic sample included 7,753 participants aged 2+ with valid dietary recall data.

Non-participation

For NHANES 2015–2016, a total of 15,327 people were eligible, of which 9,971 were interviewed and 9,544 completed the health examination component. The unweighted response rates were 61% for the interviewed sample and 59% for the examined sample. Approximately 89% of participants who completed the health examination component completed the dietary recall.

Randomization

This is a modeling study using secondary data on dietary intake, so no randomization was used.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging