

## 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  |
|-------------------------------------|--|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of all covariates tested   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> 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) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> 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 All data was downloaded directly from sources specified in data availability statement. No software was used for data collection.

Data analysis Custom code was developed using R (4.3.1) and Stata SE (Version 18.0) for data cleaning, adjustments, algorithm scoring for FCS 2.0 for individual food and beverage items, cross-sectional analyses comparing FCS 2.0 to HSR, Nutri-Score, and NOVA, i.FCS 2.0 scoring for individuals, cross-sectional and survival analyses of i.FCS 2.0 with health outcomes, and summary statistics, tables, and figures. Statistical significance was defined as two-tailed  $\alpha=0.05$ . Tufts University is considering licensing of expertise, knowledge, and background work for potential commercial viability of Food Compass for the private sector and non-profit applications, and thus the code is not publicly available. There are no intellectual property or patent protections associated with Food Compass. The detailed Food Compass algorithm and its scoring are specified in our earlier Nature Food publication, with all updates and modifications described in the current supplementary materials (Table S9 and S10), and can be freely reproduced by any individual or organization.

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](#)

The attribute and domain scoring algorithm used to generate the Food Compass 2.0 is available in Supplementary Table S9 and S10. All data used in this analysis is publicly available from the following USDA and CDC sources: (1) Nutrient composition data for foods reported in NHANES dietary recalls [USDA Food and Nutrient Database for Dietary Studies 2001-2018, FNDDS: <https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fndds-download-databases/>]. (2) Food ingredients data for foods reported in NHANES dietary recalls [USDA Food Pattern Equivalents Database, 2001-2018: <https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fped-overview/>]. (3) Flavonoid data for select foods reported in NHANES dietary recalls [USDA Flavonoid Database, 2007-2010: [ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fndds-flavonoid-database/](https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fndds-flavonoid-database/)]. (4) National dietary recall, sociodemographic, physical activity, smoking, cardiometabolic biomarker, and prevalent condition data for US adults [National Health and Nutrition Examination Survey 1999-2018, NHANES: [www.cdc.gov/nchs/nhanes/Default.aspx](http://www.cdc.gov/nchs/nhanes/Default.aspx)] (5) All-cause and cause-specific mortality data for US adults [National Death Index 1999-2018, NDI: <https://www.cdc.gov/nchs/data-linkage/mortality-public.htm>]. The Nutri-Score categorizations for each food and beverage were calculated using the 2023 updated algorithm (Merz et al. Nat Food 2024). The Health Star Rating values for each food and beverage were calculated using the publicly available online calculator and guidance: [www.healthstarrating.gov.au/internet/healthstarrating/publishing.nsf/Content/excel-calculator](http://www.healthstarrating.gov.au/internet/healthstarrating/publishing.nsf/Content/excel-calculator), [www.healthstarrating.gov.au/internet/healthstarrating/publishing.nsf/Content/guide-for-industry](http://www.healthstarrating.gov.au/internet/healthstarrating/publishing.nsf/Content/guide-for-industry). The generated Food Compass 2.0, HSR, Nutri-Score and NOVA food processing classification scores for each of the 9,273 food items in the dataset are available in Supplementary Table S5.

## Human research participants

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

Reporting on sex and gender	<input type="text" value="NA"/>
Population characteristics	<input type="text" value="NA"/>
Recruitment	<input type="text" value="NA"/>
Ethics oversight	<input type="text" value="NA"/>

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](https://www.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	<input type="text" value="Quantitative, cross-sectional and prospective analysis using secondary data from publicly available USDA and CDC databases"/>
Research sample	<input type="text" value="For individual food and beverage score FCS analysis, we applied the original and updated (2.0) Food Compass algorithm to 9273 unique foods and beverages reported in the National Health and Nutrition Examination Survey (NHANES) from the years 2001-02 to 2017-18, contained within FNDDS 2001-02 to 2017-18. We excluded infant formula, baby foods, specialized dietary foods, alcohol, and products providing &lt;5 kcal per 100g as these products are not intended to be scored using Food Compass. For i.FCS analysis, we used a nationally representative sample of US adults 1999-2018 (n=47,999) from the National Health and Nutrition Examination Survey (NHANES) from 99/00 to 17/18. Study sample size was determined by NHIS staff to allow for a nationally representative sample. We included all NHANES respondents in the dietary portion of the survey, with a few exceptions as outlined in the exclusion section. This sample was chosen to ensure we are validating Food Compass 2.0 for use with foods and beverages consumed by the US population over the past two decades."/>
Sampling strategy	<input type="text" value="Sampling procedures including sample size calculation conducted by NHIS staff for NHANES, detailed here: https://wwwn.cdc.gov/nchs/nhanes/analyticguidelines.aspx#sample-design&lt;br/&gt;Since 1999, the sample design has consisted of multi-year, stratified, clustered four-stage samples, with data release 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. NHANES 2015-2018 oversampled some subgroups to increase precision for subgroup estimates.
Data collection	Data collection was conducted by NHIS staff for NHANES. Secondary data for the present analysis come from the Household interview, MEC Medical Examination including two 24-hour dietary recalls. NHANES gathers broad health and nutrition data for various public health purposes. The data is intended for multiple uses by different researchers with their own hypotheses. Thus, data collectors are not blinded because there are no specific hypotheses to be blinded against. More details provided here: <a href="http://www.cdc.gov/nchs/data/series/sr_01/sr01_056.pdf">www.cdc.gov/nchs/data/series/sr_01/sr01_056.pdf</a>
Timing	Data is collected continuously throughout the year but released in 2-year intervals (each release being nationally representative).
Data exclusions	We excluded NHANES respondents without valid dietary recalls (n=7434), extreme total energy intake (<500 or >5000kcal/d n=704), who only reported intake of alcoholic beverages (n=3), or with missing data on smoking status (n=34). For prospective mortality analyses we also excluded participants without valid linked mortality data (n=81).
Non-participation	Data on response rates for NHANES surveys for each 2-year survey cycle can be found here: <a href="http://wwwn.cdc.gov/nchs/nhanes/ResponseRates.aspx#response-rates">wwwn.cdc.gov/nchs/nhanes/ResponseRates.aspx#response-rates</a> .
Randomization	Participants were not allocated into experimental groups.

## 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
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging