Body mass index multiple regression formula testable by eight Bradford Hill causality criteria: Worldwide ecological cohort data analysis to inform dietary guidance

David K Cundiff (0000-0002-3206-9665), independent researcher, Chunyi Wu (0000-0002-2186-3433), Research Epidemiologist/Statistician

1 Long Beach, California, USA
2 Area Specialist Lead in Epidemiology and Statistics, Michigan Medicine, Ann Arbor, Michigan, USA
3 Volunteer collaborators with the Institute of Health Metrics and Evaluation, Seattle, Washington, USA

Correspondence to: David K Cundiff davidkcundiff@gmail.com Phone: 01-562-438-8805

Word count: abstract=399, text=3603
Abstract

Objectives (1) Derive a worldwide body mass index (BMI) multiple regression formula (BMI formula), (2) Quantify the percent weights attributable to dietary and other risk factors from this BMI formula, and (3) Test relevant dietary guidelines and diets using the BMI formula generated BMI estimates.

Design BMI and risk factor cohort data from 1990-2017 from the Institute of Health Metrics and Evaluation (IHME) were formatted and population-weighted. We empirically explored the univariate and multiple regression correlations of risk factors with BMI to maximize BMI formula functionality.

Setting Worldwide

Ecological cohort data Over 12,000 Global Burden of Disease (GBD) risk factor surveys of people 15-69 years old from 195 countries analysed and synthesised into representative cohort risk factor values by staff and volunteer collaborators with the IHME.

Main outcome measures Performance of a worldwide BMI formula when tested with eight Bradford Hill causality criteria, each scored with 0-5 scale: 0=negative to 5=very strong support (40 possible score).

Results In the BMI formula derived, all foods were expressed in kilocalories/day (kcal/day), and all risk factor coefficients were adjusted to equate with their percent weight impacts on BMI.

BMI formula = 0.29*processed meat + 11.87*red meat + 0.02*fish + 3.74*milk + 11.09*poultry + 5.46*eggs + 1.95*alcohol + 6.29*sugary beverages + 0.29*corn + 3.53*potatoes + 2.71 * saturated fatty acids + 0.64*polyunsaturated fatty acids + 0.06*trans fatty acids - 0.37*fruit - 0.52*vegetables - 0.03*nuts and seeds - 0.12 *whole grains - 0.47*legumes - 8.26*rice - 18*sweet potatoes - 9.93*physical activity (METs/week) + 5.78*total kcals/day available - 4.13*child
underweight + 0.92 * discontinued breast feeding. BMI increasing foods have + signs in the BMI formula and BMI decreasing foods have - signs. Bradford Hill causality criteria test scores (0-5):

1) strength = 5, 2) consistency = 5, 3) dose-response = 5, 4) temporality = 5, 5) analogy = 4, 6) plausibility = 5, 7) experimentation = 5 and 8) coherence = 5. Total = 39/40. Based on the USA BMI trend data, the mean adult BMI in 1990 was 25.45, is projected to be 28.13 in 2020, and 31.81 in 2050. BMI formula based mean BMI estimates included the following: Dietary Guidelines for Americans = 26.34, USA with 50% reduction of BMI increasing food intake = 23.34, USA with 25% reduction of BMI increasing food intake and increasing physical activity by running at 6 mph for 1 hour/day on average = 23.67, vegetarian diet = 22.54, low carbohydrate, high fat diet = 31.76.

**Conclusions** Eight relevant Bradford Hill causality criteria strongly supported that the BMI formula derived was causally related to mean BMIs of worldwide cohorts, subgroups, and individual risk factor patterns. The statistical analysis methodology introduced could inform individual, clinical, and public health strategies regarding overweight/obesity prevention/treatment and other health outcomes.
Introduction

The scientific validity of the Dietary Guidelines for Americans for 2015-2020 was challenged by Journalist Nina Teicholz in the *BMJ*.\(^1\) The Center for Science in the Public Interest called for the *BMJ* to retract the article. The peer reviewers selected to adjudicate the far reaching dispute concluded that, “Teicholz’s criticisms of the methods used by Dietary Guidelines for Americans Committee are within the realm of scientific debate.”\(^2\) Regarding scientific study of treatments of overweight/obesity, *JAMA* published a representative randomised trial comparing a low fat diet with a low carbohydrate diet. It reported no difference in the very modest weight loss achieved at one year.\(^3\) This suggests that current short term dietary intervention trials (< 5 years) are meaningless in determining the long-term relationship of diet (> 20 years) with overweight/obesity.

Influential Stanford University meta researcher, Dr. John Ioannidis, called for radical reform of all nutritional epidemiology methodologies used to influence food/agricultural policies and to produce dietary guidelines for clinicians and the public.\(^4\) Currently, no methodology for relating body mass index (BMI) long-term or BMI change/year to food intake and physical activity has been generally accepted as rigorous, replicable, and scientifically valid. This study analysing worldwide data for dietary and other risk factors for BMI attempts to answer Dr. Ioannidis’ call for a more rigorous and reliable nutritional epidemiology methodology to base public health policies and individual dietary guidance.

The first objective of this analysis was to derive a multiple regression risk factor formula using worldwide ecological data on BMIs of male and female cohorts (dependent variable) and dietary
and other risk factors (independent variables). Secondly, we used Bradford Hill causality criteria to test the BMI formula containing worldwide diet and other risk factors. Satisfying Bradford Hill causality criteria is considered validating in epidemiological research.\(^5\) Third we wanted to apply the BMI formula to assorted diet and other risk factor patterns to derive mean BMI estimates associated with long-term (>20 years) following of those diet and other risk factor patterns. We only hypothesised that the objectives were achievable.
Methods

As volunteer collaborators with the Institute of Health Metrics and Evaluation (IHME), we received raw Global Burden of Disease (GBD) ecological data (≈1.4 Gigabytes) on mean BMIs of male and female cohorts 15-49 years old and 50-69 years old from each year 1990-2017 from 195 countries and 365 subnational locations (n=1120 cohorts). We also utilised GBD data on exposures to 32 risk factors and covariates potentially related to BMI. IHME dietary covariate data originally came from Food and Agriculture Organization surveys of animal and plant food commodities available per capita—as opposed to consumed per capita—in countries worldwide. Food risk factors came as g/day consumed per capita. Other variables were utilised as part of deriving the BMI formula and testing the BMI formula according to the Bradford Hill causality criteria: physical activity (METs/day), kilocalories per day available per capita (kcal/day, a covariate), severe child underweight (2 SD below the mean weight for age), and discontinued breast feeding before 6 months. Other available variables tested included socio-demographic index, LDL-cholesterol, fasting plasma glucose and, systolic blood pressure. Supplementary Table 1 lists the relevant GBD risk factors, covariates, and other available variables with definitions of those risk factor exposures.

GBD worldwide citations of over 12,000 surveys constituting ecological data inputs for this analysis are in Appendix 1. The main characteristics of IHME GBD data sources for BMI and all risk factor values have been published by IHME GBD data researchers and discussed elsewhere. These included detailed descriptions of categories of input data, potentially important biases, and methodologies of analysis. We did not clean or pre-process any of the GBD data. GBD cohort risk factor and BMI data from the IHME had no missing records.
Because this was a post hoc analysis and the GBD data for the study came from IHME, no ethics committee approval or institutional review board review was needed for this statistical analysis. The raw data for this analysis may be obtained by volunteer researchers collaborating with IHME.\textsuperscript{12}

IHME made available GBD BMI, risk factor, and covariate exposure data for each year from 1990-2017 for male and female cohorts from ages 15-49 years old and from 50-69 years old. To maximally utilise the available data, we averaged the values for ages 15-49 years old together with 50-69 years old for BMI and for each risk factor exposure for each male and female cohort for each year. Finally, for each male and female cohort, data from all 28 years (1990-2017) on mean BMI and on each of the risk factor exposures were averaged using the computer software program R.

To weigh the country and subnational data according to population, internet searches (mostly Wikipedia) yielded the most recent population estimates for countries and subnational states, provinces, and regions. World population data from the World Bank and the Organisation for Economic Co-operation and Development could not be used because they did not include all 195 countries or any subnational data.

Using the above described formatted dataset of risk factors, covariates, and BMIs, a software program in R generated a population-weighted analysis dataset. Each male or female cohort in the population-weighted analysis dataset represented approximately 1 million people (range: < 100,000 to 1.5 million). The analysis dataset had n=7886 cohorts (rows of data), half male and
half female, representative of over seven billion people. For example, India with about 1.234 billion people had 617 rows of the same data for males and 617 rows for females. Maldives, with about 445,000 people, had a single row of data for males and another for females. Without population-weighting the data, cohorts in India and Maldives each would have had one row in the analysis dataset, invalidating the analysis results.

This report follows the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) guidelines for reporting global health estimates.\textsuperscript{13}

Supplementary Table 2 details how omega-3 fatty acid g/day was converted to fish g/day using data on the omega-3 fatty acid content of frequently eaten fish from the National Institutes of Health Office of Dietary Supplements (USA).\textsuperscript{14} As shown in Supplementary Table 3, we converted all of the animal and plant food data, including alcohol and sugary beverage consumption, from g/day to kcal/day. For the g to kcal conversions, we used the Nutritionix track app,\textsuperscript{15} which tracks types and quantities of foods consumed. Saturated fatty acids risk factor (0-1 portion of the entire diet) was not available with GBD data from 2017, so GBD saturated fatty acids risk factor data from 2016 was used. Polyunsaturated fatty acid and trans fatty acid GBD risk factor data from 2017 (0-1 portion of the entire diet) was also utilised. These fatty acid data were converted to kcal/day by multiplying by the kcal/day available for each cohort.

\textbf{Statistical methods}

To determine the strengths of the risk factor correlations with mean BMIs of population weighted worldwide cohorts (7886 cohorts) or subgroups of cohorts (e.g., continents, socio-
demographic quartiles, etc.), we utilised Pearson correlation coefficients: $r$, 95% confidence intervals (95% CIs), and $P$ values.

In this first of its kind, post hoc, exploratory analysis; the methodology was determined as we proceeded by experimenting with strategies to optimise the functioning of the BMI formula derived. We sought to derive a BMI multiple regression formula from worldwide data with the following characteristics:

1. including as many as possible of the available animal and plant food variables,
2. including physical activity and other plausibly informative available ecological variables,
   and
3. combining dietary variables if appropriate to utilise more variables and improve BMI formula strength (i.e., $R^2$ with BMI) and consistency at predicting subgroup mean BMIs.

Appendix 2 further details the statistical methodology and explains the use of Bradford Hill causality criteria to assess whether the worldwide BMI was causally related to the risk factors in the BMI formula. Briefly, we tested the BMI formula output with the Bradford Hill causality criteria (1) strength, (2) consistency, (3) dose response, (4) temporality, (5) analogy, (6) plausibility, (7) experimentation, and (8) coherence. For each criterion, we used a 0-5 scale to assess the magnitude of support of the BMI formula output being causally related to the BMIs worldwide (0=no support of causality to 5=very strong support of causality).
In determining the variables to include and exclude in worldwide BMI formula, we set the statistical threshold for a variable to enter and to remain in the formula at $P < 0.25$. We used SAS and SAS Studio statistical software 9.4 (SAS Institute, Cary, NC) for the data analysis.
Results

Table 1 shows the basic statistics and univariate correlations of mean BMI with dietary and other risk factors and with the BMI formula derived from worldwide cohorts. Table 2 shows the univariate correlations of each of the 24 BMI formula risk factors with each other. All six animal foods (processed meat, red meat, fish, milk, poultry, and eggs), alcohol, sugary beverages, corn availability, potato availability, saturated fatty acids, polyunsaturated fatty acids, and trans fatty acids all positively correlated with BMI. Table 2 shows that corn availability (kcal/day percapita, a covariate) correlated moderately strongly with sugary beverages ($r=0.419$, 95% CI 0.400 to 0.437, $P<0.0001$), suggesting that high fructose corn syrup may account for the positive correlation with BMI. Potato availability (kcal/day percapita, a covariate), which positively correlated with BMI, included $\geq 50\%$ highly processed potato products worldwide according to the International Potato Center. Four starchy plant foods (whole grains, legumes, rice, and sweet potatoes) all negatively correlated with BMI.

The strong positive correlations of fruits, vegetables, and nuts and seeds with BMI were unexpected since nutritional literature broadly supports these foods as non-obesogenic. It suggested possible multicollinearity (when an independent variable is highly correlated with one or more known or unknown other independent variables). Supplementary Table 4 demonstrates the multicollinearity of fruits, vegetables, and nuts and seeds with BMI increasing foods:

- BMI increasing foods (animal foods + alcohol + sugary beverages + corn + potatoes) as a group were very strongly correlated with BMI ($r=0.597$, 95% CI 0.583 to 0.611, $P<0.0001$).
Fruits, vegetables, nuts and seeds grouped together were strongly positively correlated with BMI (r= 0.655, 95% CI 0.642 to 0.677, \(P< 0.0001\)) but also with BMI increasing foods (r= 0.323, 95% CI 0.304 to 0.343, \(P< 0.0001\)).

Supplementary Table 4 also shows that the top sociodemographic index (SDI) cohort quartile’s mean BMI was about 3.6 BMI units higher than the three lowest SDI quartiles (24.54 versus 20.88), meaning that the more developed countries had much higher mean BMIs. BMI increasing foods in the top SDI quartile exceeded intake in the three lowest quartiles by 58% (905/kcal/day / 571 kcal/day=1.58). Fruits, vegetables, and nuts and seeds in the top SDI quartile exceeded intake in the three lower quartiles by 1.9 to 1 (199/kcal/day / 106 kcal/day=1.88). Because of the multicollineairities of fruits, vegetables and nuts and seeds with the 10 BMI increasing foods, we grouped fruits, vegetables and nuts and seeds with BMI decreasing foods (including whole grains, legumes, rice, and sweet potatoes) in a combination variable used to derive the BMI formula, gives the coefficients of risk factors equated to their percent weights (shown in Supplementary Table 5):

\[
\text{BMI formula} = 0.29 \times \text{processed meat} + 11.87 \times \text{red meat} + 0.02 \times \text{fish} + 3.74 \times \text{milk} + 11.09 \times \text{poultry} + 5.46 \times \text{eggs} + 1.95 \times \text{alcohol} + 6.29 \times \text{sugary beverages} + 0.29 \times \text{corn} + 3.53 \times \text{potatoes} + 2.71 \times \text{saturated fatty acids} + 0.64 \times \text{polyunsaturated fatty acids} + 0.06 \times \text{trans fatty acids} - 0.37 \times \text{fruit} - 0.52 \times \text{vegetables} - 0.03 \times \text{nuts and seeds} - 0.12 \times \text{whole grains} - 0.47 \times \text{legumes} - 8.26 \times \text{rice} - 0.18 \times \text{sweet potatoes} - 9.93 \times \text{physical activity (METs/week)} + 5.78 \times \text{total kcals/day available} - 4.13 \times \text{child underweight} + 0.92 \times \text{discontinued breast feeding}.\]

Percent weights of risk factors were equated to risk factor coefficients. Total percent weights=78.64. (BMI
formula $R^2=0.7864$, $r=0.887$, 95% CI 0.882 to 0.891, $P<0.0001$, + signs mean BMI increasing and - signs mean BMI decreasing).

As detailed in the methods and Appendix 2, the standardised BMI formula coefficients of the dietary and other risk factors were all adjusted to equate to the percent weights of risk factors in the BMI formula.

As shown in Supplementary Table 5, the 13 BMI increasing foods and seven BMI decreasing foods underwent three adjustments: (1) multiplied times their mean kcal/day values (Column C), (2) multiplied times their $R^2$ values in univariate correlation with BMI (Column E), and (3) multiplied times the ratio of the mean kcal/day of the second, third, and fourth sociodemographic index (SDI) quartiles divided by the mean kcal/day of the first SDI quartile (Column G). Table 3 demonstrates that these adjustments improved the consistency of the BMI formula as measured by the closeness of fit of the multiple regression derived BMI formula outputs with the mean BMIs of 37 subgroups of the worldwide data.

**BMI formula output analysed by Bradford Hill causality criteria**

Eight Bradford Hill causality criteria tested the functionality of the BMI formula. With the methodology detailed in Appendix 2, we assessed the relationship of the worldwide BMI formula output versus the mean BMI as follows:

1. **Strength=5**—The correlation coefficient of the BMI formula regressed with BMI: $r=0.887$ (95% CI: 0.882 to 0.891), $P<0.0001$. 

13
2. Consistency score=5— Table 3 shows the average absolute differences between subgroup mean BMIs and BMI formula estimates of the BMI formula compared with two other variations of the BMI formula. The BMI formula with fruits, vegetables, and nuts and seeds grouped with BMI decreasing foods as negatively correlated with BMI had a slightly closer fit than a comparison BMI formula with fruits, vegetables, and nuts and seeds grouped with BMI increasing foods positively correlated with BMI (average BMI - BMI formula estimate (absolute difference): 0.249 versus 0.255 BMI units). Without the three adjustments to each of the food variables, the average BMI - BMI formula estimate absolute difference was 0.368 BMI units.

3. Dose-response (Biological gradient) score=5—Table 3 shows that the mean of the BMI absolute differences between the BMI formula estimates and mean BMIs and in the four dose-response quartiles was 0.302 BMI units.

4. Temporality score=5—a multiple regression with the worldwide BMI trend 1990-2007 (dependent variable) versus the risk factor trends (independent variables) generated a BMI trend formula versus risk factor BMI trends: $r= 0.615$, (95% CI: 0.601 to 0.628), $P<0.0001$). Fifteen out of 24 risk factor trends contributed to the BMI trend formula. The derivation of standardised BMI trend versus risk factors trends formula with sign concordant coefficients adjusted to equal the percent weights of the significant risk factors shown in Supplementary Table 7 was as follows:

$$
\text{BMI trend formula} = 4.84 \times \text{Red meat} + 0.63 \times \text{Milk} + 4.61 \times \text{Poultry} + 2.62 \times \text{Eggs} + 3.91 \times \text{Alcohol} + 2.69 \times \text{Sugary beverages} + 0.40 \times \text{corn} + 3.60 \times \text{Potatoes} + 3.90 \times \text{PUFA} - 0.65 \times \text{Legumes} - 3.08 \times \text{Rice} - 0.80 \times \text{Sweet potatoes} + 6.06 \times \text{kcal available (total percent weight 37.79, } R^2=0.3779).$$
5. Analogy score=4—Systolic blood pressure had 15/24 risk factors concordant in sign with BMI. SBP correlated weakly with the BMI formula (r= 0.024, 95% CI 0.002 to 0.046, P =0.0361). The low density lipoprotein cholesterol (LDL-C) and the fasting plasma glucose (FPG) formulas had 23/24 and 21/24 risk factors concordant in sign with the BMI formula, respectively. The LDL-C and the FPG were both strongly correlated with BMI (r=0.757, 95% CI 0.747 to 0.766, P < 0.0001 and (r=0.565, 95% CI 0.550 to 0.580, P < 0.0001, respectively). They were also strongly correlated with the BMI formula (r=0.757, 95% CI 0.747 to 0.766, P < 0.0001 and r=0.581, 95% CI 0.566 to 0.595), P < 0.0001, respectively.

6. Plausibility: Score=5—Based on systematic medical literature reviews, physical activity inversely correlated with BMI and BMI directly correlated with intakes of sugar, alcohol, and animal foods. The relationship of adult BMI with early childhood severe underweight has not been reported worldwide. Since people in poor countries have less animal foods and alcohol and more physical exercise than in developed countries, it is plausible that childhood severe underweight positively correlates with lower BMI in adulthood.

7. Experiment: Score=5—Table 4 demonstrates cross validation of the BMI formula with a very high degree of reproducibility in the percent weight distributions from 20 randomly selected subgroups of the dataset (n=100 for each, which are shown in Supplementary Table 7).

8. Coherence: Score=5—As evidenced by the near perfect score 34/35 on the first seven criteria, the Bradford Hill causality criteria overall were strongly supportive that the...
worldwide BMI formula accurately modeled the 24 risk factors that led to increased and decreased BMI—total causality criteria score=39/40.

Relative to other countries, the USA had one of the world’s highest levels of kcal/day of BMI increasing foods (exclusive of fatty acids), 1015 kcal/day and below average consumption of BMI decreasing foods, 326 kcal/day, corresponding to an average adult BMI=26.66. Based on the USA BMI trend data, the mean adult BMI in 1990 was 25.45 and is projected to be 28.13 in 2020, and 30.81 in 2050. Table 5 shows BMI formula estimates for various relevant patterns of diet and/or other BMI formula risk factors. Following the Dietary Guidelines for Americans 2015 -2020 would result in a mean adult BMI of 26.34, which is almost the mean USA adult BMI from 1990 - 2017 (26.66). A major increase in exercise on a regular basis alone would decrease the BMI formula estimated mean long term BMI. In addition to the current mean level of physical activity (3853 METs/week), running two hours/day on average at 6 miles per hour (8400 METs/week)\(^\text{21}\) would reduce the estimated mean BMI by 2.68 BMI units to 23.98.
Discussion

The eight Bradford Hill criteria used in this analysis of long term worldwide data all support that the BMI formula is causally related to adult BMI.

Short of extreme and unrealistic for most people increases in exercise to control weight, dietary change from the USA current dietary patterns or even the USA dietary guidelines recommended diet would be required to reach the normal range for BMI (BMI ≥ 18.5 and BMI < 25). For example, shifting 50% of kcal/day from BMI increasing foods to BMI decreasing foods equates to a BMI formula estimate=23.26 BMI units over the long term. While following a low carbohydrate, high fat diet has been demonstrated to cause modest short term weight loss in obese people, the BMI formula projects that the long-term effect (> 20 years) would be obesity (projected mean BMI=31.76).

The perhaps counterintuitive findings that fruits, vegetables, and nuts and seeds were positively correlated with BMI (r=0.655, 95% CI 0.642 to 0.677, P < 0.0001) may be largely explained by the multicollinearity between fruits, vegetables, and nuts and seeds and BMI increasing foods (r=0.323, 95% CI 0.304 to 0.343, P < 0.0001, Table 2). The high cost of fruits, vegetables, and nuts and seeds may account for much of this multicollinearity both worldwide and within wealthy countries that have high economic inequality levels like the USA. A systematic review of the literature on food cost relative to nutrient quality found that the median costs of starches (€0.14/100 kcal) was quite low relative to fruits and vegetables (€0.82/100 kcal), meat/eggs/fish (€0.64/100 kcal), fresh dairy (€0.32/100 kcal), and nuts (€0.25/100 kcal).
The findings in this study, particularly those in Table 5, should influence government food policy decisions. The Supplemental Nutrition Assistance Program (SNAP—formerly Food Stamps) spends an estimated 22.6% of its $73 billion/year budget on payments to low-income Americans for “sweetened beverages, prepared desserts, salty snacks, candy, and sugar.”

Additionally, the US Department of Agriculture (USDA) subsidises crops that go primarily for animal feed or that are processed into sugars while not subsidizing fruits and vegetables. While the USDA recognises the relatively low intake of fruits and vegetables in the USA and sponsors a publicity campaign to increase fruits and vegetables, USDA expenditures should promote reduced prices of BMI decreasing foods and increased prices of BMI increasing foods.

The strong correlation of discontinuation of breastfeeding with BMI ($r=0.802$, 95% CI 0.794 to 0.810, $P<0.0001$) was in accord with a meta-analysis of breast feeding related to subsequent childhood and adult BMI by Jeanne Stolzer. However, the estimated benefit in reducing adult BMI of breastfeeding in this study was very modest (0.19 BMI unit, see Table 5). Breastfeeding for at least six months is recommended by the American Academy of Pediatrics and the National Institute of Clinical Excellence.

**Limitations**

The GBD data on animal and plant foods were not comprehensive and comprised only 1199 kcal/day on average. Subnational data were available on only four countries. Because the data formatting and statistical methodology were new, this was necessarily a post hoc analysis and no pre-analysis protocol was possible. As detailed in the Foresight Report on obesity, obesity is affected by a complex system of interacting factors besides diet, physical activity, and breast
feeding. So genes,\textsuperscript{31} gut microbiome,\textsuperscript{32} ultra processing of food,\textsuperscript{33,34} and other influences on BMI were outside of the purview of this analysis.

359 \textbf{Generalisability}

360 Given the strength and consistency of the relationship of the BMI formula in modeling BMI (SDI quartiles, continents, and subnationals, etc., see Tables 3 and 5), the findings should be generalisable to people all over the world.
Conclusion

Eight Bradford Hill causality criteria strongly supported that the worldwide obesity epidemic is causally related to the 24 risk factors in proportion to their coefficients in the BMI formula. The findings in this study should be considered by health policymakers drafting dietary guidelines for healthy weight management. While this study deals only with dietary and other risk factors for BMI (overweight/obesity), the methodology introduced could easily apply to estimating percent weights of multiple dietary and other risk factors that pertain to dozens of non communicable diseases, for which the IHME have GBD data.
Table 1. BMI risk factor basic statistics and Bradford Hill criterion strength

| BMI and risk factors: n=7886 | Mean | Std Dev | Minimum | Maximum | r     | 95% CI       | P     | R²   |
|-------------------------------|------|---------|---------|---------|-------|--------------|-------|------|
| BMI kg/M²                     | 21.79| 2.31    | 17.95   | 29.39   | 0.887 | 0.882 to 0.891 | <.0001| 0.791 |
| BMI formula                   | 21.79| 2.31    | 18.15   | 30.32   | 0.864 | 0.590 to 0.618  | <.0001| 0.365 |
| Processed kcal/day            | 5.49 | 9.94    | 0.20    | 68.8    | 0.662 | 0.649 to 0.674  | <.0001| 0.438 |
| Red meat kcal/day             | 50.70| 45.45   | 3.21    | 235.9   | 0.613 | 0.805 to 0.820  | <.0001| 0.661 |
| Fish kcal/day                 | 10.01| 36.43   | 0.40    | 370.4   | 0.103 | 0.081 to 0.124  | <.0001| 0.111 |
| Milk kcal/day                 | 25.37| 27.39   | 1.06    | 146.8   | 0.684 | 0.672 to 0.696  | <.0001| 0.468 |
| Poultry kcal/day avail        | 45.06| 51.03   | 1.06    | 411.9   | 0.813 | 0.805 to 0.820  | <.0001| 0.596 |
| Eggs kcal/day avail           | 19.47| 14.76   | 0.79    | 69.6    | 0.688 | 0.676 to 0.699  | <.0001| 0.473 |
| Alcohol kcal/day              | 81.71| 58.06   | 4.25    | 430     | 0.166 | 0.145 to 0.188  | <.0001| 0.028 |
| Sugary beverages kcal/day     | 297.65| 152.33  | 73      | 1472    | 0.119 | 0.097 to 0.141  | <.0001| 0.014 |
| Corn kcal/day avail           | 34.67| 48.16   | 0.16    | 305.2   | 0.073 | 0.051 to 0.095  | <.0001| 0.005 |
| Potatoes kcal/day avail       | 84.16| 74.43   | 3.07    | 533.9   | 0.211 | 0.189 to 0.232  | <.0001| 0.044 |
| Saturated FA % kcal available | 191.27| 64.49   | 70.79   | 481.1   | 0.710 | 0.699 to 0.721  | <.0001| 0.504 |
| Polyunsaturated fatty acids kcal% | 82.24| 74.98   | 2.93    | 381.31  | 0.737 | 0.727 to 0.747  | <.0001| 0.543 |
| Trans fatty acids kcal%       | 13.4 | 13.82   | 1.99    | 77.76   | 0.488 | 0.471 to 0.504  | <.0001| 0.238 |
| Fruits kcal/day               | 40.39| 22.59   | 3.58    | 161.4   | 0.622 | 0.609 to 0.636  | <.0001| 0.387 |
| Vegetables kcal/day           | 80.14| 43.36   | 9.48    | 304.2   | 0.520 | 0.503 to 0.536  | <.0001| 0.270 |
| Nuts and seeds kcal/day       | 8.51 | 8.44    | 0.05    | 103.0   | 0.486 | 0.469 to 0.503  | <.0001| 0.237 |
| Whole grains kcal/day         | 55.65| 30.85   | 1.14    | 235.1   | -0.201| -0.222 to -0.179| <.0001| 0.040 |
| Legumes kcal/day              | 51.66| 32.16   | 0.51    | 194.7   | -0.384| -0.403 to -0.365| <.0001| 0.148 |
| Rice kcal/day avail           | 141.23| 116.38  | 1.42    | 461.8   | -0.561| -0.576 to -0.545| <.0001| 0.314 |
| Sweet potatoes kcal/day avail | 22.67| 35.88   | 0.02    | 364.7   | -0.152| -0.173 to -0.130| <.0001| 0.023 |
| Physical activity METs        | 4709 | 1367    | 1609    | 7669    | -0.445| -0.462 to -0.427| <.0001| 0.198 |
| Kcal available/day            | 2579 | 421     | 1579    | 3898    | 0.843 | 0.837 to 0.850  | <.0001| 0.711 |
| Child/infant underweight      | 0.19 | 0.17    | 0.00    | 0.53    | -0.794| -0.802 to -0.786| <.0001| 0.631 |
| Discontinued breast feeding< 6 months | 0.12 | 0.06 | 0.02 | 0.24 | 0.802 | 0.794 to 0.810 | <.0001 | 0.644 |
| Sodium g/day                  | 4.45 | 2.33    | 1.33    | 9.21    | -0.019| -0.041 to -0.003| 0.099  | 0.000 |
| Calcium g/day                 | 0.3  | 0.18    | 0.08    | 1.04    | 0.771 | 0.762 to 0.779  | <.0001| 0.594 |
| Dietary fiber g/day           | 9.22 | 3.15    | 2.72    | 22.68   | 0.310 | 0.290 to 0.330  | <.0001| 0.096 |
| Sugar available kcal/day       | 447.1| 283.8   | 48.3    | 1249    | 0.423 | 0.405 to 0.441  | <.0001| 0.179 |
| Fasting plasma glucose mmol/dl | 4.31 | 0.35    | 3.32    | 5.58    | 0.565 | 0.550 to 0.580  | <.0001| 0.32  |
| Systolic BP mm/Hg              | 133.9| 4.33    | 123.4   | 147.9   | 0.087 | 0.065 to 0.109  | <.0001| 0.008 |
| LDL cholesterol mmol/dl       | 2.36 | 0.4     | 1.27    | 3.25    | 0.757 | 0.747 to 0.766  | <.0001| 0.573 |
| Socio-demographic index        | 0.54 | 0.17    | 0.11    | 0.9     | 0.738 | 0.728 to 0.748  | <.0001| 0.544 |
Table 2. Correlations between the risk factors that are in Table 1 (continued)

| Variables (foods in kcal/day) | Processed meat | Red meat | Fish | Milk | Poultry | Eggs | Alcohol |
|------------------------------|----------------|----------|------|------|---------|------|---------|
| Processed meat               | 1.000          |          |      |      |         |      |         |
| Red meat                     | 0.656          | 1.000    |      |      |         |      |         |
| Fish                         | 0.281          | 0.118    | 1.000|      |         |      |         |
| Milk                         | 0.763          | 0.682    | 0.109| 1.000|         |      |         |
| Poultry                      | 0.675          | 0.670    | 0.162| 0.684| 1.000   |      |         |
| Eggs                         | 0.574          | 0.786    | 0.392| 0.532| 0.633   | 1.000|         |
| Alcohol                      | 0.592          | 0.478    | 0.292| 0.402| 0.372   | 0.352| 1.000   |
| Sugary beverages             | -0.118         | 0.037    | -0.163| 0.059| 0.126   | 0.014| -0.098  |
| Corn                         | -0.052         | -0.144   | -0.018| -0.108| 0.008   | -0.100| -0.108  |
| Potatoes                     | 0.160          | 0.124    | -0.058| 0.202| 0.104   | 0.032| -0.007  |
| Saturated fatty acids        | 0.633          | 0.709    | 0.039| 0.772| 0.709   | 0.601| 0.304   |
| Poly unsaturated fatty acids | 0.711          | 0.699    | 0.162| 0.691| 0.844   | 0.635| 0.400   |
| Trans fatty acids            | 0.479          | 0.203    | -0.006| 0.445| 0.465   | 0.217| 0.174   |
| Fruits                       | 0.391          | 0.380    | 0.099| 0.499| 0.613   | 0.374| 0.075   |
| Vegetables                   | 0.308          | 0.392    | 0.271| 0.362| 0.360   | 0.519| 0.209   |
| Nuts and seeds               | 0.520          | 0.399    | 0.073| 0.590| 0.519   | 0.344| 0.253   |
| Whole grains                 | -0.022         | -0.111   | 0.091| -0.250| -0.023  | -0.076| 0.065   |
| Legumes                      | -0.199         | -0.404   | 0.061| -0.197| -0.222  | -0.418| 0.133   |
| Rice                         | -0.396         | -0.357   | -0.071| -0.568| -0.396  | -0.277| -0.139  |
| Sweet potatoes               | -0.197         | 0.000    | -0.075| -0.395| -0.242  | 0.045| -0.068  |
| Physical activity METs       | -0.215         | -0.072   | -0.163| -0.431| -0.397  | -0.119| 0.097   |
| Kcal /day available          | 0.602          | 0.649    | 0.098| 0.775| 0.754   | 0.644| 0.231   |
| Child/infant 2SD underweight | -0.416         | -0.737   | -0.189| -0.409| -0.618  | -0.804| -0.141  |
| Discontinued breast feeding <6 mo | 0.583 | 0.812    | 0.219| 0.639| 0.727   | 0.847| 0.307   |

Table 2. Correlations between the risk factors that are in Table 1 (continued)

| Variables (foods in kcal/day) | Sugary beverages | Corn | Potatoes | Fruits | Vegetables | Nuts and seeds | Whole grains |
|------------------------------|-------------------|------|----------|--------|------------|----------------|--------------|
| Corn                         | 0.419             | 1.000|          |        |            |                |              |
| Potatoes                     | -0.026            | 0.158| 1.000    |        |            |                |              |
| Fruits                       | 0.189             | 0.100| 0.243    | 1.000  |            |                |              |
| Vegetables                   | -0.173            | -0.203| -0.044   | 0.394  | 1.000      |                |              |
| Nuts and seeds               | -0.154            | -0.114| 0.188    | 0.586  | 0.556      | 1.000          |              |
| Whole grains                 | -0.204            | 0.428| -0.058   | 0.012  | -0.346     | -0.208         | 1.000        |
| Legumes                      | 0.199             | 0.269| 0.018    | -0.032 | -0.160     | -0.028         | 0.100        |
| Rice                         | -0.142            | -0.248| -0.378   | -0.325 | -0.237     | -0.359         | 0.547        |
| Sweet potatoes               | -0.215            | -0.027| 0.238    | -0.100 | -0.099     | -0.079         | -0.039       |
| Physical activity METs       | -0.099            | 0.028| 0.021    | -0.393 | -0.212     | -0.306         | 0.268        |
| Kcal /day available          | 0.114             | -0.044| 0.182    | 0.640  | 0.615      | 0.601          | -0.187       |
| Child/infant 2SD underweight | -0.060            | -0.081| -0.216   | -0.497 | -0.457     | -0.300         | 0.089        |
| Discontinued breast feeding <6 mo | 0.133 | -0.064| 0.132    | 0.564  | 0.480      | 0.373          | -0.125       |
Table 2. Correlations between the risk factors that are in Table 1 (continued)

| Variables (foods in kcal/day) | Legumes | Rice  | Sweet potatoes | Physical activity | Kcal    | Child/ infant underweight |
|-------------------------------|---------|-------|----------------|-------------------|---------|---------------------------|
| Rice                          | -0.047  | 1.000 |                |                   |         |                           |
| Sweet potatoes                | -0.053  | 0.072 | 1.000          |                   |         |                           |
| Physical activity METs        | -0.066  | 0.388 | 0.328          | 1.000             |         |                           |
| Kcal /day available           | -0.250  | -0.429| -0.257         | -0.500            | 1.000   |                           |
| Child/infant 2SD underweight  | 0.507   | 0.377 | -0.164         | 0.138             | -0.621  | 1.000                     |
| Discontinued breast feeding <6 mo | -0.406  | -0.419| -0.062         | -0.227            | 0.727   | -0.836                    |
Table 3. Bradford Hill criteria consistency and dose response: Mean BMI and BMI formula output differences in 37 subgroups

| Subsets of worldwide data on risk factors related to BMI | n  | Mean BMI | BMI formula output with mean and SD equated with mean BMI | BMI and BMI formula‡ output difference | BMI and BMI formula¶ output difference | BMI and BMI formula† output difference |
|----------------------------------------------------------|----|----------|----------------------------------------------------------|---------------------------------------|----------------------------------------|---------------------------------------|
| Socio-demographic index quartile 1                       | 1966 | 24.54    | 24.75                                                   | 0.210                                 | 0.243                                  | 0.289                                  |
| SDI quartile 2                                           | 2096 | 21.82    | 21.69                                                   | 0.130                                 | 0.210                                  | 0.402                                  |
| SDI quartile 3                                           | 2220 | 20.40    | 20.30                                                   | 0.099                                 | 0.104                                  | 0.107                                  |
| SDI quartile 4                                           | 1604 | 20.31    | 20.31                                                   | 0.007                                 | 0.079                                  | 0.012                                  |
| Africa                                                   | 1682 | 21.67    | 21.25                                                   | 0.411                                 | 0.009                                  | 0.519                                  |
| Asia                                                     | 4188 | 20.47    | 20.46                                                   | 0.013                                 | 0.065                                  | 0.058                                  |
| Europe                                                   | 880  | 24.31    | 24.66                                                   | 0.343                                 | 0.388                                  | 0.234                                  |
| North America                                            | 592  | 25.97    | 26.00                                                   | 0.032                                 | 0.079                                  | 0.312                                  |
| Oceana                                                   | 54   | 24.43    | 24.66                                                   | 0.229                                 | 0.277                                  | 0.206                                  |
| South America                                            | 474  | 23.48    | 24.24                                                   | 0.761                                 | 0.667                                  | 0.463                                  |
| Four countries with subnational data                    | 730  | 25.18    | 25.45                                                   | 0.267                                 | 0.300                                  | 0.964                                  |
| Quartile 1 BMI increasing foods                          | 2007 | 24.14    | 24.70                                                   | 0.565                                 | 0.572                                  | 0.654                                  |
| Q2                                                       | 1834 | 21.96    | 21.86                                                   | 0.096                                 | 0.103                                  | 0.274                                  |
| Q3                                                       | 2050 | 20.60    | 20.31                                                   | 0.293                                 | 0.281                                  | 0.205                                  |
| Q4                                                       | 1995 | 20.49    | 20.28                                                   | 0.212                                 | 0.226                                  | 0.224                                  |
| Quartile 1 BMI decreasing foods                          | 1475 | 20.61    | 20.16                                                   | 0.454                                 | 0.570                                  | 0.197                                  |
| Q2 BMI decreasing foods                                  | 2275 | 21.12    | 21.03                                                   | 0.092                                 | 0.114                                  | 0.114                                  |
| Q3 BMI decreasing foods                                  | 2211 | 22.48    | 22.74                                                   | 0.258                                 | 0.294                                  | 0.239                                  |
| Q4 BMI decreasing foods                                  | 1925 | 22.69    | 22.82                                                   | 0.125                                 | 0.199                                  | 0.112                                  |
| Quartile 1 physical activity                             | 2371 | 20.48    | 20.56                                                   | 0.081                                 | 0.089                                  | 0.019                                  |
| Q2 physical activity                                     | 1573 | 21.66    | 21.13                                                   | 0.533                                 | 0.598                                  | 0.554                                  |
| Q3 physical activity                                     | 2018 | 21.37    | 21.55                                                   | 0.175                                 | 0.198                                  | 0.308                                  |
| Q4 physical activity                                     | 1924 | 23.95    | 24.07                                                   | 0.118                                 | 0.136                                  | 0.125                                  |
| Male                                                     | 3943 | 21.51    | 21.91                                                   | 0.406                                 | 0.401                                  | 0.368                                  |
| Female                                                   | 3943 | 22.07    | 21.65                                                   | 0.423                                 | 0.419                                  | 0.382                                  |
| Random controlled trial #1                               | 100  | 21.92    | 21.96                                                   | 0.038                                 | 0.054                                  | 0.061                                  |
| RCT #2                                                   | 100  | 22.00    | 21.76                                                   | 0.236                                 | 0.226                                  | 0.198                                  |
| RCT #3                                                   | 100  | 21.67    | 21.78                                                   | 0.109                                 | 0.111                                  | 0.149                                  |
| RCT #4                                                   | 100  | 21.89    | 21.81                                                   | 0.074                                 | 0.069                                  | 0.064                                  |
| UK                                                       | 66   | 24.99    | 25.13                                                   | 0.132                                 | 0.132                                  | 0.393                                  |
| USA                                                      | 376  | 26.66    | 26.74                                                   | 0.082                                 | 0.132                                  | 0.380                                  |
| Japan                                                    | 158  | 21.89    | 22.87                                                   | 0.978                                 | 0.964                                  | 2.888                                  |
| Mexico                                                   | 130  | 24.99    | 24.99                                                   | 0.006                                 | 0.062                                  | 0.606                                  |
| Quartile 1 Dose-response: BMI formula versus BMI         | 1862 | 24.82    | 25.33                                                   | 0.509                                 | 0.538                                  | 0.663                                  |
| Q2 Dose-response                                         | 1802 | 22.44    | 22.06                                                   | 0.379                                 | 0.350                                  | 0.539                                  |
| Q3 Dose-response                                         | 1773 | 21.10    | 20.84                                                   | 0.266                                 | 0.034                                  | 0.201                                  |
| Q4 Dose-response                                         | 2449 | 19.51    | 19.56                                                   | 0.057                                 | 0.158                                  | 0.128                                  |
| Mean of BMI - BMI formula absolute values                |      | 0.249    | 0.255                                                   | 0.368                                  |
In combination variables grouped by BMI increasing foods and BMI decreasing foods, each food was multiplied times (1) its kcal/day (2) its coefficient of determination (R^2) with BMI, and (3) its ratio of kcal/day in SDI quartiles 2-4 /SDI quartile 1. Fruits, vegetables, and nuts/seeds grouped with plants that negatively correlate with BMI. 

† As with ‡ but with fruits, vegetables, and nuts/seeds grouped with plants that positively correlate with BMI 

‡ As with § but with food risk factor unadjusted but with food risk factors unadjusted
Table 4. Cross validation: summary percent weight data for 20 RCT derived BMI formulas

| Risk factors                        | Mean | Std Dev | Minimum | Maximum |
|------------------------------------|------|---------|---------|---------|
| Processed meat kcal/day            | 0.26 | 0.09    | 0.11    | 0.43    |
| Red meat kcal/day                  | 9.82 | 3.05    | 5.41    | 14.38   |
| Fish kcal/day                      | 0.18 | 0.51    | 0.01    | 1.97    |
| Milk kcal/day                      | 3.13 | 1.22    | 1.38    | 4.88    |
| Poultry kcal/day available         | 9.10 | 2.88    | 5.35    | 15.53   |
| Eggs kcal/day available            | 4.66 | 1.09    | 2.50    | 7.73    |
| Alcohol kcal/day                   | 1.78 | 2.08    | 0.01    | 7.92    |
| Sugary beverages kcal/day          | 7.24 | 1.51    | 0.01    | 14.62   |
| Corn kcal/day available            | 0.80 | 0.70    | 0.00    | 2.00    |
| Potatoes kcal/day available        | 2.79 | 1.94    | 0.01    | 7.81    |
| Saturated fatty acids kcal/day     | 3.09 | 2.09    | 0.90    | 8.18    |
| Poly unsaturated fatty acids kcal/day | 0.71 | 0.38    | 0.18    | 1.50    |
| Trans fatty acids kcal/day         | 0.09 | 0.05    | 0.03    | 0.18    |
| Fruits kcal/day                    | 0.46 | 0.21    | 0.12    | 0.88    |
| Vegetables kcal/day                | 0.63 | 0.30    | 0.23    | 1.31    |
| Nuts and seeds kcal/day            | 0.03 | 0.02    | 0.01    | 0.07    |
| Whole grains kcal/day              | 0.13 | 0.02    | 0.00    | 0.39    |
| Legumes kcal/day                   | 0.56 | 3.51    | 0.09    | 1.54    |
| Rice kcal/day available            | 8.89 | 2.65    | 3.40    | 13.19   |
| Sweet potatoes kcal/day available  | 0.32 | 0.32    | 0.02    | 1.13    |
| Physical activity METs             | 9.76 | 4.09    | 4.24    | 13.29   |
| Kcal available/day                 | 7.89 | 3.23    | 3.68    | 13.56   |
| Child/infant 2SD underweight       | 5.41 | 4.51    | 2.49    | 8.29    |
| Discontinued breast feeding <6 mo  | 1.19 | 0.84    | 0.20    | 2.96    |
| Total percent weights              | 78.90| 2.88    | 74.65   | 83.04   |

See Supplementary Table 7 for the data on the 20 RCTs each with n=100
Table 5. BMI formula estimates with common dietary and other risk factor scenarios

| Testing common dietary and other risk factor scenarios with standardised BMI formula estimates | BMI kg/M² | BMI formula estimate |
|---------------------------------------------------------------------------------------------|----------|---------------------|
| World (n=7886 cohorts)                                                                        | 21.79    | 21.71               |
| United Kingdom (n=66 cohorts)                                                                | 24.99    | 25.09               |
| USA (n=376 cohorts)                                                                            | 26.66    | 26.67               |
| Following USA Dietary Guidelines 2015-2020 recommendations                                        |          |                     |
| USA with 25% reduction of BMI increasing food intake†                                        |          | 25.04               |
| USA with 50% reduction of BMI increasing food intake†                                        |          | 23.34               |
| USA mean physical activity plus 2 hours/day running at 6 mph                                 |          | 23.98               |
| USA mean physical activity plus 1 hour/day running at 6 mph and 25% reduction of BMI increasing food intake† |          | 23.67               |
| USA with no red or processed meat†                                                           |          | 24.81               |
| USA with no sugary beverage intake†                                                          |          | 26.31               |
| USA vegetarian (no meat, poultry, fish)†                                                     |          | 22.54               |
| USA vegan (no meat, poultry, fish, dairy or eggs)†                                            |          | 21.11               |
| USA breast feeding continued at least 6 months in all                                          |          | 26.68               |
| USA breast feeding discontinued before 6 months in all                                        |          | 26.87               |
| USA no occurrence of child severe underweight                                                 |          | 26.73               |
| USA all children with severe underweight (theoretically)                                     |          | 25.54               |
| EAT-Lancet diet                                                                              |          | 22.47               |
| Low Carbohydrate Mediterranean Diet                                                          |          | 31.76               |

BMI formula estimates based on 28 years of following dietary and/or risk factor patterns
Kcal/day isocalorically shifted to the 7 BMI decreasing foods in the BMI formula, distributed equally
**Acknowledgments**

**Authors’ contributions:** DKC acts as guarantor; conceived and designed the study, acquired and analysed the data, interpreted the study findings, drafted the manuscript, critically reviewed and edited the manuscript and tables, and approved the final version for publication.

CW designed software programs in R to format and population weight the data, aided with the SAS statistical analysis, critically reviewed the manuscript, and approved the final version for publication.

The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

**Copyright:** The retain the copyright to the paper

The Patient and Public Involvement

- When and how were patients/public first involved in the research?

IHME acquired, catalogued and extracted information from over 12,000 surveys from government and non government agencies in order to compile the GBD database.

- How were the research question(s) developed and informed by their priorities, experience, and preferences?

The diverse surveys contributing to the GBD database had many different priorities and intentions unrelated to this post hoc analysis. The current analysis research questions were
developed to give researchers, policymakers, and the public the methodological tool to quantify the BMI impacts of dietary and other risk factor patterns.

● How were patients/public involved in

(a) the design and conduct of the study?

The way patients/public were involved in the collection of the surveys has not been systematically studied and reported on. Patients/public are the intended beneficiaries of this analysis of the GBD data.

(b) choice of outcome measures?

The outcome measures chosen all related to human health and to adult BMI.

(c) recruitment to the study?

The recruitment methods varied by the survey.

● How were (or will) patients/public be involved in choosing the methods and agreeing plans for dissemination of the study results to participants and linked communities?

IHME and all the health surveyors that contributed will make these decisions. This analysis of the data will be available to all by open access.

• Competing interests statement: Both authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Contributors: Martin Sebera, from the Department of Kinesiology, Faculty of Sports Studies
Masaryk University, Czech Republic, critiqued statistical aspects of the manuscript and provided useful input. Pavel Grasgruber, from Masaryk University, Czech Republic, provided suggestions after reviewing the manuscript.

Transparency declaration: The manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Ethics: Studies based solely on data from IHME GBD database do not need approval from any bioethics committee.

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. The Bill and Melinda Gates Foundation funded the acquisition of the data for this analysis by the IHME. The data were provided to the authors as volunteer collaborators with IHME.

Details of the role of the study sponsors: While IHME GBD faculty and staff by virtue of Gates Foundation grants provided the raw data for this analysis, they did not vet the analysis or sponsor the manuscript.

Statement of independence of researchers from funders: The researchers have received no funding. Gates Foundation funded IHME to collect and analyse the GBD data.
Dissemination declaration: Dissemination of this manuscript to the participants of the more than 12,000 surveys is not possible individually, but the information will become in the public domain.

Disclosures: We thank Scott Glenn and Brent Bell from IHME who supplied us with the GBD risk factor exposure data for the risk factors and for BMI data.

Ethics: Studies were based solely on data from the IHME GBD database and do not need approval from any bioethics committee.

Participant informed consents: Not applicable.

Author access to data: As volunteer collaborators with the Institute of Health Metrics and Evaluation, we received about 1.4 gigabytes of raw data on BMI and 32 relevant risk factors for BMI and other health outcomes.

Data sharing statement: The raw, unformatted data used in this analysis is now out of date. The 2019 GBD data on all the variables in this analysis may be obtained from the IHME by volunteer collaborating researchers. Our data formatting software code in R and SAS and our formatted database are available on request to researchers.

Protocol, submitted as a supplementary file: Not applicable.
STROBE checklist: Submitted.

Patient consent: Not applicable.

Clinical trial registration: Not applicable.
1. Teicholz N. The scientific report guiding the US dietary guidelines: is it scientific? BMJ. 2015; 351. http://www.bmj.com/bmj/351/bmj.h4962.full.pdf

2. Outcome of Post-publication review of article by Nina Teicholz. BMJ. 2015; 351: h4962. https://www.bmj.com/content/351/bmj.h4962/rr-48

3. Gardner CD, Trepanowski JF, Del Gobbo LC, et al. Effect of Low-Fat vs Low-Carbohydrate Diet on 12-Month Weight Loss in Overweight Adults and the Association With Genotype Pattern or Insulin Secretion: The DIETFITS Randomized Clinical Trial. JAMA. 2018; 319 (7): 667-679. https://jamanetwork.com/journals/jama/fullarticle/2673150

4. Ioannidis JA. The challenge of reforming nutritional epidemiologic research JAMA 2018; 320 (10): 969-70. http://dx.doi.org/10.1001/jama.2018.11025

5. Hill AB. THE ENVIRONMENT AND DISEASE: ASSOCIATION OR CAUSATION? Proceedings of the Royal Society of Medicine. 1965; 58 (5): 295-300. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1898525/

6. Global Burden of Disease Study 2017 (GBD 2017) Data Input Sources Tool. Seattle, Washington: Institute of Health Metrics and Evaluation Accessed August 10, 2019. http://ghdx.healthdata.org/gbd-2017/data-input-sources

7. GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017: Table of risk factor definitions. The Lancet. 2018; 392 (10159): 1923-94.
8. PROTOCOL FOR THE GLOBAL BURDEN OF DISEASES, INJURIES, AND RISK FACTORS STUDY (GBD) Version 3.0; Issue 26. Seattle, Washington: Institute for Health Metrics and Evaluation. Accessed August 2, 2019.

http://www.healthdata.org/sites/default/files/files/Projects/GBD/GBD_Protocol.pdf

9. Naghavi M, Abajobir AA, Abbafati C, et al. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980-2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet. 2017; 390(10100): 1151-1210. https://doi.org/10.1016/S0140-6736(17)32152-9

10. Flaxman AD, Lee YY, Vos T, et al. An Integrative Metaregression Framework for Descriptive Epidemiology. Seattle, WA: University of Washington Press, 2015.

11. GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2018; 392(10159): 1923-94. https://doi.org/10.1016/S0140-6736(18)32225-6

12. Call for Collaborators. Seattle, Washington: Institute for Health Metrics and Evaluation at the University of Washington 2019; Accessed April 25, 2019.

http://www.healthdata.org/gbd/call-for-collaborators

13. Stevens GA, Alkema L, Black PRE, et al. Guidelines for Accurate and Transparent Health Estimates Reporting: the GATHER statement. The Lancet. 2016; 388(10062): e19-e23. https://doi.org/10.1016/S0140-6736(16)30388-9
14. Omega 3 Fatty Acids: Fact Sheet for Health Professionals. The Office of Dietary Supplements 2018; Accessed September 1, 2018. https://ods.od.nih.gov/factsheets/Omega3FattyAcids-HealthProfessional/

15. Nutritionix Track App. Syndigo LLC Accessed April 25, 2019. https://www.nutritionix.com/

16. POTATO PROCESSING AND USES. Lima, Peru: International Potato Center Accessed May 28, 2020. https://cipotato.org/potato/potato-processing-uses/

17. Ruotsalainen H, Kyngäs H, Tammelin T, Kääriäinen M. Systematic review of physical activity and exercise interventions on body mass indices, subsequent physical activity and psychological symptoms in overweight and obese adolescents. *Journal of Advanced Nursing.* 2015; 71 (11): 2461-2477. https://onlinelibrary.wiley.com/doi/abs/10.1111/jan.12696

18. Te Morenga L, Mallard S, Mann J. Dietary sugars and body weight: systematic review and meta-analyses of randomised controlled trials and cohort studies. *BMJ* 2013; 346: e7492. https://www.bmj.com/content/bmj/346/bmj.e7492.full.pdf

19. Traversy G, Chaput J-P. Alcohol Consumption and Obesity: An Update. *Current obesity reports.* 2015; 4 (1): 122-130. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4338356/

20. Turner-McGrievy G, Mandes T, Crimaco A. A plant-based diet for overweight and obesity prevention and treatment. *Journal of geriatric cardiology : JGC.* 2017; 14 (5): 369-374. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5466943/

21. AINSWORTH BE, HASKELL WL, HERRMANN SD, et al. *Compendium of Physical Activities: A Second Update of Codes and MET Values.* *Medicine & Science in Sports*
& Exercise. 2011 43 (8): 1575-1581. https://journals.lww.com/acsm-msse/Fulltext/2011/08000/2011_Compendium_of_Physical_Activities___A_Second.25.aspx

22. Ebbeling CB, Feldman HA, Klein GL, et al. Effects of a low carbohydrate diet on energy expenditure during weight loss maintenance: randomized trial. BMJ. 2018; 363: k4583.

23. Darmon N, Drewnowski A. Contribution of food prices and diet cost to socioeconomic disparities in diet quality and health: a systematic review and analysis. Nutrition reviews. 2015; 73 (10): 643-660. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4586446/

24. Wakamo B. We Subsidize the Wrong Kind of Agriculture. Inequality.org Accessed. https://inequality.org/great-divide/subsidize-wrong-kind-agriculture/

25. Tagtow A. Fruits and Veggies Now Chock Full of Marketing Power. Washington, DC 2017; Accessed June 24, 2020. https://www.usda.gov/media/blog/2015/05/07/fruits-and-veggies-now-chock-full-marketing-power

26. Stolzer JM. Breastfeeding and obesity: a meta-analysis. Open Journal of Preventive Medicine. 2011; 1 (3): 88-93. http://www.scirp.org/journal/paperinformation.aspx?paperid=8509

27. AAP Reaffirms Breastfeeding Guidelines. The American Academy of Pediatrics 2012; Accessed August 10, 2019. https://www.aap.org/en-us/about-the-aap/aap-press-room/Pages/AAP-Reaffirms-Breastfeeding-Guidelines.aspx

28. Breastfeeding and the Use of Human Milk. Pediatrics. 2012; 129 (3): e827-e841. doi:10.1542/peds.2011-3552
29. Quality statement 5: Breastfeeding. National Institute of Clinical Excellence 2013; Accessed August 10, 2019. https://www.nice.org.uk/guidance/qs37/chapter/Quality-statement-5-Breastfeeding

30. FORESIGHT--Tackling Obesities: Future Choices -Follow Up Actions. London, England: Government Office for Science (UK) October 2007; Accessed August 5, 2019. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/295885/07-1526-obesity-action-plan.pdf

31. Mârginean CO, Mârginean C, Meliț LE. New Insights Regarding Genetic Aspects of Childhood Obesity: A Minireview. Frontiers in pediatrics. 2018; 6: 271-271. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6180186/

32. Castaner O, Goday A, Park Y-M, et al. The Gut Microbiome Profile in Obesity: A Systematic Review. International journal of endocrinology. 2018; 2018: 4095789-4095789. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5933040/

33. Monteiro CA, Moubarac J-C, Levy RB, Canella DS, Louzada MLdC, Cannon G. Household availability of ultra-processed foods and obesity in nineteen European countries. Public Health Nutrition. 2017; 21 (1): 18-26. https://www.cambridge.org/core/article/household-availability-of-ultraprocessed-foods-and-obesity-in-nineteen-european-countries/D63EF7095E8EFE72BD825AFC2F331149

34. Monteiro CA, Cannon G, Levy RB, et al. Ultra-processed foods: what they are and how to identify them. Public Health Nutrition. 2019; 22 (5): 936-941. https://www.cambridge.org/core/article/ultraprocessed-foods-what-they-are-and-how-to-identify-them/E6D744D714B1FF09D5B3A3E74D53A185
| Variables                  | Definition                                                                                                                                                                                                 |
|----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Alcohol                    | Any alcohol consumption (g/day)                                                                                                                                                                           |
| Body-mass index            | Body mass index (BMI) (kg/m²)—the dependent variable of interest                                                                                                                                           |
| Child underweight          | Proportion of children - 3 SD to - 2 SD of the WHO 2006 standard weight-for-age curve (0-1)                                                                                                                  |
| Corn                       | Corn availability per capita (g/day), a covariate                                                                                                                                                           |
| Discontinued breast feeding| Proportion of children aged 6-23 months who do not receive any breast milk                                                                                                                                   |
| Eggs                       | Eggs availability per capita (g/day) a covariate                                                                                                                                                            |
| Fasting plasma glucose     | Fasting plasma glucose (mmol/L)                                                                                                                                                                            |
| Fish                       | This variable expressed in g/day was derived by determining the weight of fish in g corresponding to 1 g of omega-3 fatty acids (eicosapentaenoic acid and docosahexaenoic acid) by averaging the fish g per 1 g of omega-3 fatty acids 20 species of fish= 117.04 g/day fish/1 g/day omega-3 fatty acids (Supplementary Table 2) |
| Fruits                     | Consumption of fruits (includes fresh, frozen, cooked, canned, or dried fruit but excludes fruit juices and salted or pickled fruits) (g/day)                                                             |
| Kilocalories available/day | The mean number of kilocalories per capita available per day to people in each location (kcal/day available), a covariate                                                                                     |
| LDL cholesterol            | Serum low-density lipoprotein cholesterol (mmol/L)                                                                                                                                                         |
| Legumes                    | Consumption of beans, lentils, pulses (g/day)                                                                                                                                                              |
| Milk                       | Consumption of milk including non-fat, low-fat, and full-fat milk but excluding soy milk and other plant derivatives (g/day)                                                                               |
| Nuts and seeds             | Consumption of nuts and seeds (g/day)                                                                                                                                                                       |
| Physical activity          | Average weekly physical activity at work, home, transport-related and recreational measured by MET min per week. Less than 3000 METs per week constitutes low physical activity. |
| Poultry                    | Poultry availability per capita (g/day), a covariate                                                                                                                                                        |
| Potatoes                   | Potatoes availability per capita (g/day), a covariate                                                                                                                                                      |
| Processed meat             | Consumption of any processed meat (includes meat preserved by smoking, curing, salting, or addition of chemical preservatives, including bacon, salami, sausages, or deli or luncheon meats like ham, turkey, and pastrami (g/day) |
| Red meat                   | Consumption of red meat (includes beef, pork, lamb, and goat but excludes poultry, fish, eggs, and all processed meats) (g/day)                                                                           |
| Rice                       | Rice availability per capita (g/day), a covariate                                                                                                                                                          |
| Seafood omega-3 fatty acids| Seafood omega-3 fatty acids (eicosapentaenoic acid and docosahexaenoic acid) in tablet or fish form (g/day)                                                                                              |
| Socio-demographic index    | SDI is a composite indicator of development status that was originally constructed for GBD 2015, and is derived from components that correlate strongly with health outcomes. It is the geometric mean for indices of the total fertility rate among women younger than 25 years, mean education for those aged 15 years or older, and lag-distributed income per capita. The resulting metric ranges from 0 to 1, with higher values corresponding to higher levels of development. |
| Sugar-sweetened beverages  | Consumption of any beverage with ≥50 calories of sugar per one-cup serving, including carbonated beverages, sodas, energy drinks, fruit drinks but excluding 100% fruit and vegetable juices (g/day) |
| Sweet potatoes             | Sweet potato availability per capita (g/day), a covariate                                                                                                                                                   |
| Systolic blood pressure    | Systolic blood pressure (mm Hg)                                                                                                                                                                            |
| Total sugar                | Total sugar availability per capita (g/day), a covariate                                                                                                                                                   |
| **Vegetables** | Consumption of frozen, cooked, canned, or dried vegetables (including legumes but excluding salted or pickled, juices, nuts and seeds, and starchy vegetables such as potatoes or corn) (g/day) |
| **Whole grains** | Consumption of whole grains (bran, germ, and endosperm in their natural proportions) from breakfast cereals, bread, rice, pasta, biscuits, muffins, tortillas, pancakes, and others (g/day) |
### Supplementary Table 2. Omega-3 Fatty Acid g to fish g calculation

| Fish                                | DHA g/3 ounce fish | EPA g/3 ounce fish | Omega-3 Fatty Acids (DHA - EPA) g/3 ounce fish mean | Fish 3 ounces = 85.02 g | Fish (g) per omega-3 Fatty Acids (g)=columns E/F |
|-------------------------------------|--------------------|--------------------|---------------------------------------------------|-------------------------|-----------------------------------------------|
| Salmon Atlantic farmed              | 1.24               | 0.59               |                                                   |                         |                                               |
| Salmon Atlantic wild                | 1.22               | 0.35               |                                                   |                         |                                               |
| Herring Atlantic                    | 0.94               | 0.77               |                                                   |                         |                                               |
| Sardines canned in tomato sauce drained | 0.74             | 0.45               |                                                   |                         |                                               |
| Mackerel Atlantic                   | 0.59               | 0.43               |                                                   |                         |                                               |
| Salmon pink canned drained          | 0.63               | 0.28               |                                                   |                         |                                               |
| Trout rainbow wild                  | 0.44               | 0.40               |                                                   |                         |                                               |
| Oysters eastern wild                | 0.23               | 0.30               |                                                   |                         |                                               |
| Sea bass                            | 0.47               | 0.18               |                                                   |                         |                                               |
| Shrimp                              | 0.12               | 0.12               |                                                   |                         |                                               |
| Lobster                             | 0.07               | 0.10               |                                                   |                         |                                               |
| Tuna light canned in water drained  | 0.17               | 0.02               |                                                   |                         |                                               |
| Tilapia                             | 0.11               |                    |                                                   |                         |                                               |
| Scallops                            | 0.09               | 0.06               |                                                   |                         |                                               |
| Cod Pacific                         | 0.1                | 0.04               |                                                   |                         |                                               |
| Tuna yellowfin                      | 0.09               | 0.01               |                                                   |                         |                                               |
| Mean DHA and EPA Omega-3 Fatty Acids g/3 ounce fish | 0.4531           | 0.2733             |                                                   |                         |                                               |
| Calculations total Omega-3 FA g to fish g |                |                    | 0.7264                                            | 85.02                   | 117.043                                       |

¶ Data on omega-3 fatty acid content of varieties of fish came from the National Institutes of Health Office of Dietary Supplements (USA)
Supplementary Table 3. Calculations of kcal/day from g/day of animal and plant foods

| Foods                  | Food sub-categories | kcal/serving | g/serving | kcal/g |
|------------------------|---------------------|--------------|-----------|--------|
| Milk (2% fat)          |                     | 122          | 244       | 0.5    |
| Fish                   |                     | 218          | 170       | 1.28   |
| Eggs                   |                     | 72           | 50        | 1.44   |
| Poultry                |                     | 187          | 85        | 2.91   |
| Red meat               |                     | 247          | 85        | 2.91   |
| Processed meat         |                     |              |           |        |
| Salami                 |                     | 222          | 59        | 3.76   |
| Pastrami               |                     | 104          | 71        | 1.46   |
| Ring baloney           |                     | 86           | 28        | 3.07   |
| Pepperoni              |                     | 94           | 100       | 0.94   |
| Average processed meat |                     | 126.5        | 64.5      | 1.96   |
| Fruits                 |                     | 97           | 162       | 0.60   |
| Vegetables             |                     | 59           | 91        | 0.65   |
| Legumes                |                     | 249          | 179       | 1.39   |
| Nuts                   |                     | 172          | 28        | 6.14   |
| Seeds                  |                     |              |           |        |
| Flax seeds             |                     | 55           | 10        | 5.5    |
| Chia seeds             |                     | 58           | 12        | 4.83   |
| Fennel seeds           |                     | 34.5         | 10        | 3.45   |
| Hemp seeds             |                     | 55.3         | 10        | 5.53   |
| Average of seeds       |                     | 50.7         | 10.5      | 4.83   |
| Average of nuts and seeds |               | 111.4        | 19.25     | 5.78   |
| Corn                   |                     | 99           | 103       | 0.96   |
| Potatoes               |                     | 161          | 173       | 0.93   |
| Sweet potatoes         |                     | 115          | 151       | 0.76   |
| Rice                   |                     | 205          | 158       | 1.3    |
| Whole grains           |                     | 120          | 52        | 2.31   |

¶ Source: NutritionIX app
# Supplementary Table 4. Multicollinearity of fruits, vegetables, and nuts and seeds

| High BMI risk factors (n=7886 cohorts) | Mean Kcal/day | BMI kg/M² | 95% CI | P | BMI increasing r † | 95% CI | P |
|---------------------------------------|--------------|-----------|--------|---|-------------------|--------|---|
| BMI kg/m²                              | 21.79        |           |        |   | 0.724             | 0.713 to 0.734 | <0.0001 |
| BMI increasing r                       | 654          | 0.597     | 0.583 to 0.611 | <0.0001 | 0.323 | 0.304 to 0.343 | <0.0001 |
| Fruits + vegetables + nuts and seeds  | 129          | 0.655     | 0.642 to 0.677 | <0.0001 | 0.498 | 0.482 to 0.515 | <0.0001 |
| Fruits                                | 40.4         | 0.622     | 0.609 to 0.636 | <0.0001 | 0.150 | 0.128 to 0.171 | <0.0001 |
| Vegetables                            | 80.1         | 0.520     | 0.503 to 0.536 | <0.0001 | 0.279 | 0.258 to 0.299 | <0.0001 |
| Nuts and seeds                        | 8.51         | 0.486     | 0.469 to 0.503 | <0.0001 |        |                   |    |
| Top socio-demographic index (SDI) quartile n=1996 | | | | |BMI kg/m² |           |        |   |-------------------|--------|---|
| BMI kg/m²                              | 24.54        |           |        |   | 0.364             | 0.325 to 0.402 | <0.0001 |
| BMI increasing r                       | 905          | 0.364     | 0.325 to 0.402 | <0.0001 | 0.346 | 0.306 to 0.364 | <0.0001 |
| Fruits + vegetables + nuts and seeds  | 199          | 0.104     | 0.060 to 0.148 | <0.0001 | -0.022 | -0.007 to -0.095 | 0.32 |
| Fruits                                | 61.1         | 0.296     | 0.255 to 0.336 | <0.0001 | 0.0043 | -0.425 to -0.350 | <0.0001 |
| Vegetables                            | 120          | -0.064    | -0.108 to -0.020 | 0.0043 | -0.388 | -0.161 to -0.074 | <0.0001 |
| Nuts and seeds                        | 17.5         | 0.314     | 0.273 to 0.353 | <0.0001 |        |                   |    |
| Bottom three SDI quartiles (n=5920 cohorts) | | | | |BMI kg/m² |           |        |   |-------------------|--------|---|
| BMI kg/m²                              | 20.88        |           |        |   | 0.364             | 0.342 to 0.386 | <0.0001 |
| BMI increasing r                       | 571          | 0.364     | 0.342 to 0.386 | <0.0001 | 0.083 | 0.108 to 0.133 | <0.0001 |
| Fruits + vegetables + nuts and seeds  | 106          | 0.536     | 0.517 to 0.554 | <0.0001 | 0.429 | 0.408 to 0.450 | <0.0001 |
| Fruits                                | 33.5         | 0.466     | 0.446 to 0.486 | <0.0001 | 0.103 | -0.128 to -0.077 | 0.0186 |
| Vegetables                            | 66.8         | 0.395     | 0.374 to 0.417 | <0.0001 |        |                   |    |
| Nuts and seeds                        | 5.52         | -0.040    | -0.066 to -0.015 | 0.0018 | -0.072 | -0.098 to -0.047 | <0.0001 |

† BMI increasing foods include all six animal products, alcohol, sugary beverages, corn (high fructose corn syrup included), and potatoes (highly processed potato products included).

Saturated fatty acids, PUFAs, and TFAs also correlate positively with BMI. However, for the purposes of this table, they were not included with BMI increasing foods, because animal foods contain most of the fatty acids.
## Supplementary Table 5. Derivation of the BMI formula (variables all standardized)

| Variables in BMI formula | A | B | C | D | E | F | G | H | I | J | K |
|--------------------------|---|---|---|---|---|---|---|---|---|---|---|
| **Processed meat kcal/day** | Processed meat kcal/day | * 5.486 | * 0.365 | * 0.099 | * 0.02972 | = 0.00586 |
| **Red meat kcal/day** | Red meat kcal/day | * 50.70 | * 0.438 | * 0.364 | * 0.02972 | = 0.24003 |
| **Fish kcal/day** | Fish kcal/day | * 10.02 | * 0.011 | * 0.106 | * 0.02972 | = 0.00033 |
| **Milk kcal/day** | Milk kcal/day | * 25.37 | * 0.468 | * 0.214 | * 0.02972 | = 0.07553 |
| **Poultry kcal/day availability** | Poultry kcal/day availability | * 45.06 | * 0.661 | * 0.253 | * 0.02972 | = 0.22416 |
| **Eggs kcal/day availability** | Eggs kcal/day availability | * 19.47 | * 0.473 | * 0.402 | * 0.02972 | = 0.11011 |
| **Alcohol kcal/day** | Alcohol kcal/day | * 81.71 | * 0.028 | * 0.587 | * 0.02972 | = 0.03948 |
| **Sugary beverages kcal/day** | Sugary beverages kcal/day | * 297.7 | * 0.014 | * 1.013 | * 0.02972 | = 0.12722 |
| **Corn kcal/day availability** | Corn kcal/day availability | * 34.67 | * 0.005 | * 1.081 | * 0.02972 | = 0.00590 |
| **Potatoes kcal/day availability** | Potatoes kcal/day availability | * 84.16 | * 0.044 | * 0.644 | * 0.02972 | = 0.07137 |
| **Saturated fatty acids** | Saturated fatty acids | * 191.3 | * 0.504 | * 0.653 | * 0.0008705 | = 0.05481 |
| **Polyunsaturated fatty acids** | Polyunsaturated fatty acids | * 82.24 | * 0.543 | * 0.334 | * 0.0008705 | = 0.01297 |
| **Trans fatty acids** | Trans fatty acids | * 13.39 | * 0.238 | * 0.471 | * 0.0008705 | = 0.00131 |
| **Fruits kcal/day** | Fruits kcal/day | * 40.39 | * 0.387 | * 0.549 | * 0.0008705 | = 0.00747 |
| **Vegetables kcal/day** | Vegetables kcal/day | * 80.14 | * 0.270 | * 0.555 | * 0.0008705 | = 0.01046 |
| **Nuts and seeds kcal/day** | Nuts and seeds kcal/day | * 8.511 | * 0.237 | * 0.315 | * 0.0008705 | = 0.00055 |
| **Whole grains kcal/day** | Whole grains kcal/day | * 55.65 | * 0.040 | * 1.248 | * 0.0008705 | = 0.00243 |
| **Legumes kcal/day** | Legumes kcal/day | * 51.65 | * 0.148 | * 1.432 | * 0.0008705 | = 0.00951 |
| **Rice kcal/day availability** | Rice kcal/day availability | * 141.2 | * 0.314 | * 4.320 | * 0.0008705 | = 0.16697 |
| **Sweet potatoes kcal/day availability** | Sweet potatoes kcal/day availability | * 22.67 | * 0.023 | * 7.977 | * 0.0008705 | = 0.00362 |
| **Physical activity METs** | Physical activity METs | 0.20087 | = 0.20087 |
| **Total Kcal/day availability** | Total Kcal/day availability |
| **Child/infant underweight** | Child/infant underweight |
| **Discontinued breast feed <6 months** | Discontinued breast feed <6 months |
| **Sums** | Sums | 1.37098 |
| **BMI formula R2** | BMI formula R2 | 0.73530 |
| **BMI formula R2 / sum of parameter estimates** | BMI formula R2 / sum of parameter estimates | 0.53633 |
| Variables in BMI formula | A | L | M | N | O | P | Q | R | S |
|--------------------------|---|---|---|---|---|---|---|---|---|
| Processed meat kcal/day  |   | 0.00315 |   | 0.00315 | + | Processed meat | * | 0.29 |
| Red meat kcal/day        |   | 0.12874 |   | 0.12874 | + | Red meat | * | 11.87 |
| Fish kcal/day            |   | 0.00018 |   | 0.00018 | + | Fish | * | 0.02 |
| Milk kcal/day            |   | 0.04051 |   | 0.04051 | + | Milk | * | 3.74 |
| Poultry kcal/day         |   | 0.12022 |   | 0.12022 | + | Poultry | * | 11.09 |
| Eggs kcal/day            |   | 0.05906 |   | 0.05906 | + | Eggs | * | 5.45 |
| Alcohol kcal/day         |   | 0.02117 |   | 0.02117 | + | Alcohol | * | 1.95 |
| Sugary beverages kcal/day|   | 0.06823 |   | 0.06823 | + | Sugary beverages | * | 6.29 |
| Corn kcal/day            |   | 0.00317 |   | 0.00317 | + | Corn | * | 0.29 |
| Potatoes kcal/day        |   | 0.03828 |   | 0.03828 | + | Potatoes | * | 3.53 |
| Saturated fatty acids    |   | 0.02940 |   | 0.02940 | + | SFA | * | 2.71 |
| Polyunsaturated fatty acids |   | 0.00696 |   | 0.00696 | + | PUFA | * | 0.64 |
| Trans fatty acids        |   | 0.00070 |   | 0.00070 | + | TFA | * | 0.06 |
| Fruits kcal/day          |   | 0.00401 |   | 0.00401 | - | Fruits | * | 0.37 |
| Vegetables kcal/day      |   | 0.00561 |   | 0.00561 | - | Vegetables | * | 0.52 |
| Nuts and seeds kcal/day  |   | 0.00030 |   | 0.00030 | - | Nuts and seeds | * | 0.03 |
| Whole grains kcal/day    |   | 0.00130 |   | 0.00130 | - | Whole grains | * | 0.12 |
| Legumes kcal/day         |   | 0.00510 |   | 0.00510 | - | Legumes | * | 0.47 |
| Rice kcal/day            |   | 0.08955 |   | 0.08955 | - | Rice | * | 8.26 |
| Sweet potatoes kcal/day  |   | 0.00194 |   | 0.00194 | - | Sweet Potatoes | * | 0.18 |
| Physical activity        |   | 0.10773 |   | 0.10773 | - | Physical activity | * | 9.93 |
| Total Kcal/day           |   | 0.5408 |   | 0.063 | 0.06270 | + | Total kcal | * | 5.78 |
| Child/infant underweight |   | 0.3865 |   | 0.045 | 0.04481 | - | Child underweight | * | 4.13 |
| Discontinued breast feed <=6 months |   | 0.0861 |   | 0.010 | 0.00999 | + | Discontinued breast feeding | * | 0.92 |
| Sums                     |   | 0.7353 |   | 1.0133 | 0.118 | 0.85280 | 78.64 |
| Sums                     |   | 1.0133 |   | 0.85280 |    |            |
| BMI formula R2           |   | 0.1175 |   | 0.78643 |    |            |
| BMI formula R2 / sum of parameter estimates |   | 0.1160 |   | 0.92218 |    |            |

\[ \mu \text{ Adjust column } K \text{ by } K28 = 0.5314 \text{ (foods and physical activity BMI formula } R^2 (K27)/\text{parameter estimate sums (K26)} \]
Supplementary Table 6. Spreadsheet derivation of BMI trend formula

| BMI trend risk factor trends | Mean | Risk factor with BMI17fts R2 | Parameter estimates from multiple regression | Column C * Column E * Column G= preliminary parameter estimates |
|-----------------------------|------|-------------------------------|---------------------------------------------|--------------------------------------------------|
| Processed meat kcal/day trend | 5.486 | 0.0001 | = 0.000 | |
| Red meat kcal/day trend | 50.695 | 0.1650 | 0.01645 = 0.138 | |
| Fish kcal/day trend | 10.015 | 0.0251 | = 0.000 | |
| Milk kcal/day trend | 25.372 | 0.0425 | 0.01645 = 0.018 | |
| Poultry kcal/day trend | 45.059 | 0.1768 | 0.01645 = 0.131 | |
| Eggs kcal/day trend | 19.466 | 0.2322 | 0.01645 = 0.074 | |
| Alcohol kcal/day trend | 81.713 | 0.0826 | 0.01645 = 0.111 | |
| Sugary beverages kcal/day trend | 297.648 | 0.0156 | 0.01645 = 0.077 | |
| Corn kcal/day available trend | 34.669 | 0.0198 | 0.01645 = 0.011 | |
| Potatoes available kcal/day available trend | 84.160 | 0.0739 | 0.01645 = 0.102 | |
| Saturated fatty acids trend | 191.27 | | | |
| Polyunsaturated fatty acids trend | 82.24 | 0.0820 | 0.01645 = 0.111 | |
| Trans fatty acids trend | 13.4 | | | |
| Fruits kcal/day trend | 40.388 | 0.2301 | = 0.000 | |
| Vegetables kcal/day trend | 80.141 | 0.1559 | = 0.000 | |
| Nuts and seeds kcal/day trend | 8.510 | 0.0152 | = 0.000 | |
| Whole grains kcal/day trend | 55.647 | 0.0107 | = 0.000 | |
| Legumes kcal/day trend | 51.657 | 0.0887 | 0.00404 = 0.019 | |
| Rice kcal/day available trend | 141.230 | 0.1536 | 0.00404 = 0.088 | |
| Sweet potatoes available kcal/day trend | 22.667 | 0.2487 | 0.00404 = 0.023 | |
| Physical activity mets trend | | | | |
| Kcal available/day trend | | | 0.17225 = 0.17225 | |
| Child/infant 2SD underweight trend | | | | |
| Discontinued breast feeding <6 mo trend | | | | |
| Sums | | | 1.07414 | |
| Sum | | | 1.07414 | |
| $R^2$ | | | 0.37790 | |
| Trend $R^2$/trend BMI formula parameter estimates sum | | | 0.35182 | |
Spreadsheet derivation of BMI trend formula continued

| A | I | J | K | L | M | N |
|---|---|---|---|---|---|---|
| BMI trend risk factor trends | Column C * Column E * Column G= preliminary parameter estimates | Column I adjusted to equate to parameter estimates sum to R2 of the BMI trend formula | Final BMI trend formula to check percent weights with SAS Studio | Column J * 100 to derive percent weights of risk factors |
| Processed meat kcal/day trend | 0.000 | | | pmeat17KCs | |
| Red meat kcal/day trend | 0.138 | 0.048 | + | rmeat17KCs | * 4.841 |
| Fish kcal/day trend | 0.000 | 0.000 | | fish17KCs | |
| Milk kcal/day trend | 0.018 | 0.006 | + | milk17KCs | * 0.624 |
| Poultry kcal/day trend | 0.131 | 0.046 | + | poultry16KCs | * 4.610 |
| Eggs kcal/day trend | 0.074 | 0.026 | + | eggs16KCs | * 2.616 |
| Alcohol kcal/day trend | 0.111 | 0.039 | + | Alcohol17KCs | * 3.908 |
| Sugary beverages kcal/day trend | 0.077 | 0.027 | + | Sugarb17KCs | * 2.692 |
| Corn kcal/day available trend | 0.011 | 0.004 | + | corn16KCs | * 0.397 |
| Potatoes available kcal/day available trend | 0.102 | 0.036 | + | potatoes16KCs | * 3.600 |
| Saturated fatty acids trend | | | 0.000 | |
| Polyunsaturated fatty acids trend | 0.111 | 0.039 | + | PUFA17KCs | * 3.904 |
| Trans fatty acids trend | | | 0.000 | |
| Fruits kcal/day trend | 0.000 | 0.000 | | fruits17KCs | |
| Vegetables kcal/day trend | 0.000 | 0.000 | | Vegetables17KCs | |
| Nuts and seeds kcal/day trend | 0.000 | 0.000 | | nutsseeds17KCs | |
| Whole grains kcal/day trend | 0.000 | 0.000 | | wgrains17KCs | |
| Legumes kcal/day trend | 0.019 | 0.007 | - | legumes17KCs | * 0.651 |
| Rice kcal/day available trend | 0.088 | 0.031 | - | rice16KCs | * 3.084 |
| Sweet potatoes available kcal/day available trend | 0.023 | 0.008 | - | swtpot16KCs | * 0.801 |
| Physical activity mets trend | | | 0.000 | | PAMETs17ts |
| Kcal available/day trend | 0.172 | 0.061 | + | KCal2016ts | * 6.060 |
| Child/infant 2SD underweight trend | | | | Childunwt17ts |
| Discontinued breast feeding <6 mo trend | | | | discbreastF17ts |
| Sums | 1.074 | 0.378 | | | 37.79 |
| Sum | 1.074 | | | | 0.37790 |
| R² | 0.378 | | | | 100.00 |
| Trend R² / trend BMI formula parameter estimates sum | 0.352 | | | | |
### Supplementary Table 7. Cross validation experiment with 20 randomised controlled trials of n=100

| Risk factors                        | RCT1 Percent Weights | RCT2 Percent Weights | RCT3 Percent Weights | RCT4 Percent Weights | RCT5 Percent Weights |
|-------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Processed meat kcal/day             | 0.39                 | 0.22                 | 0.14                 | 0.22                 | 0.11                 |
| Red meat kcal/day                   | 13.27                | 8.57                 | 6.41                 | 7.71                 | 5.41                 |
| Fish kcal/day                       | 0.00                 | 0.01                 | 0.01                 | 0.02                 | 0.01                 |
| Milk kcal/day                       | 5.96                 | 3.58                 | 1.38                 | 2.72                 | 1.42                 |
| Poultry kcal/day available          | 13.66                | 8.22                 | 5.35                 | 8.90                 | 5.53                 |
| Eggs kcal/day available             | 5.43                 | 3.72                 | 2.50                 | 4.35                 | 3.72                 |
| Alcohol kcal/day                    | 0.56                 | 1.58                 | 0.30                 | 1.76                 | 0.93                 |
| Sugary beverages kcal/day           | 0.01                 | 8.49                 | 13.37                | 9.67                 | 11.67                |
| Corn kcal/day available             | 0.01                 | 0.16                 | 0.19                 | 1.96                 | 1.04                 |
| potatoes kcal/day available         | 2.17                 | 0.83                 | 4.55                 | 2.36                 | 2.60                 |
| Saturated fatty acids kcal/day      | 1.79                 | 7.70                 | 3.12                 | 1.44                 | 3.71                 |
| Polyunsaturated fatty acids kcal/day| 0.51                 | 1.58                 | 0.88                 | 0.48                 | 0.86                 |
| Trans fatty acids kcal/day          | 0.05                 | 0.20                 | 0.08                 | 0.07                 | 0.07                 |
| Fruits kcal/day                     | 0.29                 | 0.73                 | 0.56                 | 0.28                 | 0.79                 |
| Vegetables kcal/day                 | 0.45                 | 1.05                 | 0.59                 | 0.63                 | 0.89                 |
| Nuts and seeds kcal/day             | 0.03                 | 0.06                 | 0.03                 | 0.02                 | 0.05                 |
| Whole grains kcal/day               | 0.18                 | 0.26                 | 0.04                 | 0.06                 | 0.01                 |
| Legumes kcal/day                    | 0.46                 | 1.10                 | 0.65                 | 0.24                 | 0.47                 |
| Rice kcal/day available             | 3.40                 | 12.76                | 9.29                 | 7.81                 | 8.14                 |
| Sweet potatoes kcal/day available   | 0.05                 | 0.74                 | 0.18                 | 0.53                 | 0.06                 |
| Physical activity METs             | 13.08                | 13.10                | 7.90                 | 9.28                 | 7.11                 |
| Kcal available/day                 | 6.52                 | 3.72                 | 10.22                | 8.31                 | 14.90                |
| Child/infant 2SD underweight       | 4.86                 | 2.49                 | 7.62                 | 6.03                 | 10.51                |
| Discontinued breast feeding <6 mo    | 1.52                 | 0.52                 | 0.42                 | 0.45                 | 2.34                 |
| Total percent weights in BMI formula| 74.65                | 81.39                | 75.77                | 75.28                | 82.35                |

### Supplementary Table 7. continued

| Risk factors                        | RCT6 Percent weights | RCT7 Percent weights | RCT8 Percent weights | RCT9 Percent weights | RCT10 Percent weights |
|-------------------------------------|----------------------|----------------------|----------------------|----------------------|-----------------------|
| Processed meat kcal/day             | 0.27                 | 0.34                 | 0.26                 | 0.39                 | 0.33                  |
| Red meat kcal/day                   | 8.14                 | 11.88                | 8.69                 | 13.49                | 14.38                 |
| Fish kcal/day                       | 0.01                 | 0.04                 | 0.02                 | 0.04                 | 0.03                  |
| Milk kcal/day                       | 2.89                 | 3.98                 | 1.89                 | 3.47                 | 4.88                  |
| Poultry kcal/day available          | 6.69                 | 11.11                | 5.89                 | 7.89                 | 15.53                 |
| Eggs kcal/day available             | 4.41                 | 5.84                 | 4.14                 | 6.21                 | 7.18                  |
| Alcohol kcal/day                    | 0.11                 | 4.39                 | 0.62                 | 7.92                 | 1.26                  |
| Sugary beverages kcal/day           | 10.44                | 4.11                 | 14.62                | 0.34                 | 1.79                  |
| Corn kcal/day available             | 0.01                 | 0.00                 | 0.82                 | 1.49                 | 0.01                  |
| potatoes kcal/day available         | 1.82                 | 1.97                 | 1.80                 | 2.15                 | 5.32                  |
| Saturated fatty acids kcal/day      | 8.18                 | 1.87                 | 1.97                 | 3.01                 | 2.42                  |
| Risk factors                              | RCT11 Percent weights | RCT12 Percent weights | RCT13 Percent weights | RCT14 Percent weights | RCT15 Percent weights |
|------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Processed meat kcal/day                  | 0.13                  | 0.18                  | 0.28                  | 0.35                  | 0.16                  |
| Red meat kcal/day                        | 6.31                  | 6.46                  | 12.56                 | 11.78                 | 6.90                  |
| Fish kcal/day                            | 0.01                  | 0.02                  | 0.02                  | 0.01                  | 0.01                  |
| Milk kcal/day                            | 1.70                  | 2.12                  | 3.83                  | 2.99                  | 2.35                  |
| Poultry kcal/day available               | 7.81                  | 6.50                  | 11.94                 | 11.25                 | 7.17                  |
| Eggs kcal/day available                  | 2.82                  | 3.48                  | 5.55                  | 5.44                  | 2.83                  |
| Alcohol kcal/day                         | 0.21                  | 0.06                  | 0.60                  | 0.13                  | 1.35                  |
| Sugary beverages kcal/day                | 10.33                 | 8.84                  | 4.75                  | 4.17                  | 10.87                 |
| Corn kcal/day available                  | 1.32                  | 0.33                  | 1.21                  | 0.81                  | 1.29                  |
| Potatoes kcal/day available              | 1.73                  | 2.12                  | 7.81                  | 2.91                  | 0.01                  |
| Saturated fatty acids kcal/day           | 2.90                  | 2.49                  | 2.40                  | 2.90                  | 3.65                  |
| Polyunsaturated fatty acids kcal/day     | 0.58                  | 0.70                  | 0.56                  | 0.44                  | 0.88                  |
| Trans fatty acids kcal/day               | 0.05                  | 0.15                  | 0.03                  | 0.08                  | 0.17                  |
| Fruits kcal/day                          | 0.28                  | 0.43                  | 0.34                  | 0.50                  | 0.75                  |
| Vegetables kcal/day                      | 0.64                  | 0.68                  | 0.40                  | 0.61                  | 0.84                  |
| Nuts and seeds kcal/day                  | 0.01                  | 0.02                  | 0.02                  | 0.03                  | 0.04                  |
| Whole grains kcal/day                     | 0.16                  | 0.39                  | 0.22                  | 0.21                  | 0.06                  |
| Legumes kcal/day                         | 0.38                  | 0.89                  | 0.77                  | 0.29                  | 0.65                  |
| Rice kcal/day available                  | 9.30                  | 10.10                 | 8.34                  | 9.08                  | 7.09                  |
| Sweet potatoes kcal/day available        | 0.12                  | 0.18                  | 0.04                  | 0.25                  | 0.67                  |
| Physical activity METs                  | 9.01                  | 10.97                 | 6.78                  | 9.80                  | 10.34                 |
| Kcal available/day                       | 13.56                 | 11.05                 | 4.34                  | 7.80                  | 12.42                 |
| Child/infant 2SD underweight             | 8.29                  | 8.08                  | 4.01                  | 4.05                  | 7.08                  |
| Discontinued breast feeding <6 mo        | 0.32                  | 0.20                  | 0.58                  | 0.47                  | 1.52                  |
| Total percent weights in BMI formula     | 77.97                 | 76.44                 | 77.38                 | 76.37                 | 79.09                 |
| Risk factors                        | RCT16 Percent weights | RCT17 Percent weights | RCT18 Percent weights | RCT19 Percent weights | RCT20 Percent weights |
|------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Processed meat kcal/day            | 0.33                  | 0.21                  | 0.29                  | 0.43                  | 0.26                  |
| Red meat kcal/day                  | 13.14                 | 7.99                  | 7.48                  | 14.07                 | 11.69                 |
| Fish kcal/day                      | 0.02                  | 0.02                  | 0.01                  | 1.97                  | 1.28                  |
| Milk kcal/day                      | 3.65                  | 2.02                  | 4.49                  | 4.10                  | 3.12                  |
| Poultry kcal/day available         | 11.06                 | 5.98                  | 9.24                  | 11.22                 | 11.02                 |
| Eggs kcal/day available            | 4.90                  | 3.25                  | 4.17                  | 7.73                  | 5.61                  |
| Alcohol kcal/day                   | 4.34                  | 1.52                  | 0.01                  | 4.10                  | 3.80                  |
| Sugary beverages kcal/day          | 1.58                  | 12.85                 | 7.01                  | 3.37                  | 6.49                  |
| Corn kcal/day available            | 2.00                  | 0.42                  | 1.95                  | 0.17                  | 0.77                  |
| potatoes kcal/day available        | 2.98                  | 0.18                  | 3.88                  | 6.20                  | 2.33                  |
| Saturated fatty acids kcal/day     | 1.20                  | 5.43                  | 1.81                  | 0.90                  | 2.87                  |
| Polyunsaturated fatty acids kcal/day| 0.29                  | 1.50                  | 0.49                  | 0.18                  | 1.07                  |
| Trans fatty acids kcal/day         | 0.05                  | 0.18                  | 0.06                  | 0.03                  | 0.07                  |
| Fruits kcal/day                    | 0.23                  | 0.88                  | 0.35                  | 0.12                  | 0.47                  |
| Vegetables kcal/day                | 0.32                  | 1.17                  | 0.54                  | 0.23                  | 0.38                  |
| Nuts and seeds kcal/day            | 0.02                  | 0.05                  | 0.03                  | 0.01                  | 0.04                  |
| Whole grains kcal/day              | 0.03                  | 0.00                  | 0.04                  | 0.12                  | 0.02                  |
| Legumes kcal/day                   | 0.17                  | 0.68                  | 0.40                  | 0.09                  | 0.56                  |
| Rice kcal/day available            | 7.18                  | 12.12                 | 9.14                  | 6.79                  | 13.19                 |
| Sweet potatoes kcal/day available  | 0.02                  | 1.13                  | 0.04                  | 0.22                  | 0.21                  |
| Physical activity METs             | 13.29                 | 4.32                  | 12.97                 | 9.16                  | 9.16                  |
| Kcal available/day                 | 8.12                  | 9.24                  | 7.45                  | 6.10                  | 4.78                  |
| Child/infant 2SD underweight       | 4.37                  | 5.50                  | 6.14                  | 3.27                  | 2.89                  |
| Discontinued breast feeding <6 mo  | 1.29                  | 2.87                  | 0.49                  | 1.22                  | 0.95                  |
| Total percent weights in BMI formula| 80.57                 | 79.47                 | 78.50                 | 81.80                 | 83.04                 |
Appendix 1. Worldwide surveys contributing to the IHME GBD risk factor data.

Online only.
Appendix 2. Bradford Hill causality criteria based assessment methodology detailed

A literature search revealed no published methodological precedents for statistically modeling the correlation between mean BMIs of worldwide countries and subnational regions/provinces/states and their corresponding dietary and other risk factors. Since the Bradford Hill causality criteria (enumerated by the English occupational physician and epidemiologist Sir Austin Bradford Hill) are the gold standard assessment tools to test causality of risk factors related to health outcomes, we explored the data looking for mean BMI to risk factor correlations that could be tested with the Bradford Hill criteria.

The relevant causality criteria included #1 strength, #2 consistency, #3 dose-response (biological gradient), #4 temporality, #5 analogy, #6 plausibility, #7 experimentation, and #8 coherence. The non applicable criterion is specificity. Since diet, physical activity, and other risk factors may affect many non-communicable disease health outcomes, specificity is not relevant. As Dr. Bradford Hill said, “In short, if specificity exists we may be able to draw conclusions without hesitation; if it is not apparent, we are not thereby necessarily left sitting irresolutely on the fence.”

In considering many candidate methodologies involving univariate and multiple regression analysis, we sought a technique that produced the worldwide BMI versus risk factors multiple regression formula which most accurately predicted the mean BMI of subgroups (e.g., continents, socio-economic index quartiles, etc.). Such a BMI formula with quantifiable high accuracy in predicting the mean BMI of subgroups would necessarily score high with the
Bradford Hill criteria strength, consistency, and dose-response (biological gradient). Of the candidate statistical modeling strategies, the methodology that created the worldwide BMI formula that most accurately estimated subgroup mean BMIs had the following steps:

1. Transform all the food group risk factors (i.e., alcohol, total sugar, plant foods, animal foods, and fatty acids) from g/day to kilocalories/day (kcal/day).
2. Multiply each food group risk factor by its mean kcal/day consumption.
3. Multiply each risk factor by the $R^2$ (coefficient of determination) of its univariate correlation with BMI.
4. Adjust for multicollinearities (any animal or plant risk factors in the combination variables that have significant multicollinearity (when an independent variable is highly correlated with one or more of the other independent variables in a multiple regression equation).
5. After performing #2, #3, and #4 above, form combination dietary variables with individual foods each expressed in kcal/day to be added together as the data suggests are appropriate:

   a. For example: processed meat * processed meat kcal/day * processed meat $R^2$ * processed meat kcal/day in bottom three SDI quartiles/top quartile + red meat * red meat $R^2$ * red meat kcal/day * red meat kcal/day in bottom three SDI quartiles/ kcal/day in the top quartile + …+ eggs * eggs kcal/day, etc.
b. Explore methods to adjust for multicollinearities based in part on maximizing the closeness of fit of the multiple regression derived BMI formula outputs with the mean BMIs of 37 subgroups of the worldwide data.

6. Add physical activity before performing multiple regression.

7. Perform multiple regression analysis with the food combination variables in #5 and physical activity to generate a BMI formula. (e.g., BMI formula = combination variable_a * parameter estimate_a + combination variables_b * parameter estimate_b … + physical activity * parameter estimate of physical activity.

8. Explore integrating the variables kcal available, child underweight, and discontinued breast feeding into the BMI formula.

9. Discard any of these variables that's sign in regression with BMI switches from univariate to the multivariate analysis (e.g., + in univariate and - in multivariate formula).

10. To simplify the calculations, use the Excel spreadsheet and transfer back and forth to SAS Studio.

11. To test the BMI formula with subgroups, match the BMI formula mean and standard deviation (SD) with that of mean worldwide BMI:

   a. Adjust the SD of the BMI formula output to equal the SD of the worldwide BMI by multiplying the BMI formula output times the BMI SD / BMI formula SD.

   b. Adjust the BMI formula constant to equal the worldwide population’s mean BMI by adding the difference between the mean BMI and the BMI formula constant.
12. Explore the BMI formula functionality with different Bradford Hill criteria with standardising the BMI and BMI risk factors or with all variables non standardised.

We estimated percent weights attributable to each risk factor with the standardised BMI formula from the above methodology with the following additions:

1. Equate the sum of the individual food and other risk factor parameter estimates to the total $R^2$ of the BMI formula:
   a. Total the absolute values of the parameter estimates for each food and any non food risk factors.
   b. Divide the worldwide BMI formula $R^2$ by the sum of the risk factor parameter estimates.
   c. Multiply the result of b above times each of the individual risk factor parameter estimates.

2. Multiply the results in #1 above times 100 to generate percent weights for each risk factor. The sum of the risk factor percent weights will equal the BMI formula $R^2$ times 100.

To cross validate the generated worldwide BMI formula experimentally, create 20 random subgroups of 100 cohorts with the macro function in SAS. Determine each of the 20 subgroup BMI formulas, and, from those 20 BMI formulas, create a table with: mean, SD, minimum, and maximum for each of the risk factors.
Bradford Hill causal criteria testing methodology applied to current study

The relevant eight of the nine original Bradford Hill criteria were each scored as:

“5” very strongly supporting causality,
“4” strongly supporting causality,
“3” moderately strongly supporting causality,
“2” supporting causality,
“1” weakly supporting causality, and
“0” not supporting causality

The scoring for each Bradford Hill causal criterion was as follows:

1. Strength: The correlation coefficient, r, of the worldwide multiple regression derived BMI formula with BMI (dependent variable) and BMI risk factors (independent variables) assessed strength.

Scoring of strength:

5=BMI formula regressed with BMI $r \geq 0.50$ and $P<0.0001$
4=BMI formula regressed with BMI $0.50>r \geq 0.40$ and $P<0.0001$
3=BMI formula regressed with BMI $0.40>r \geq 0.30$ and $P<0.0001$
2=BMI formula regressed with BMI $0.30>r \geq 0.20$ and $P<0.0001$
1=BMI formula regressed with BMI $0.20>r \geq 0.10$ and $P<0.0001$
0=BMI formula regressed with BMI $r<0.10$ or $P \geq 0.0001$
2. Consistency: For the purposes of this study, consistency between BMI and BMI formula output was determined by comparing the mean BMI and the mean BMI formula output in each of the following 37 subgroups:

1) A variable socio-demographic index (SDI)—see Supplementary Table 1 for definition of SDI—divided the world’s population by quartiles of SDI.

2) A variable “continents” allowed for analyses of countries from each of the six inhabited continents.

3) The four countries (UK, USA, Mexico, and Japan) with subnational data on BMI and the risk factors were grouped and assessed to compare the BMI formula output with the overall mean BMI in those countries.

4) Based on the total kcal/day of all foods that increased BMI, a combination variable was constructed and the world’s population divided into quartiles from the highest to lowest total kcal/day.

5) Similarly, based on the total kcal/day of all foods that decreased BMI, we divided the world’s population into quartiles from the highest to lowest by total kcal/day.

6) Based on physical activity (METs/week), the world’s population was divided into quartiles.

7) We evaluated dose response by dividing the BMI formula output into quartiles from the highest to lowest.

8) The four countries with subnational data were individually evaluated.

9) The first four of the 20 random number generated database subgroups were included in the consistency analysis.

10) Male and female were separately assessed.
11) Assessing dose-response was done in quartiles by the BMI formula outputs after the BMI formula was harmonized with the mean and SD of worldwide mean BMI. See the dose-response criteria (#4).

For each of the 37 subgroups, the absolute differences between the means of BMI and the BMI formula output in were totaled (e.g., continent Africa BMI formula output - mean BMI for Africa, etc.).

Scoring of consistency for BMI compared with BMI formula output for each of the 37 subgroups:

5=The mean of the absolute differences between mean BMI and BMI formula output was \( \leq 0.40 \) units.

4=The mean of the absolute differences between mean BMI and BMI formula output was \( \leq 0.50 \) units.

3=The mean of the absolute differences between mean BMI and BMI formula output was \( \leq 0.60 \) units.

2=The mean of the absolute differences between mean BMI and BMI formula output was \( \leq 0.70 \) units.

1=The mean of the absolute differences between mean BMI and BMI formula output was \( \leq 1.0 \) units.

0=The mean of the absolute differences between mean BMI and BMI formula output was > 1.0 units.
3. Dose-response (biological gradient): Dr. Bradford Hill thought that a clear dose-response effect on the incidence of disease with exposure to a single risk factor was the clearest evidence of a causal relationship. More recently, it has been realized that cause and effect relationships are often more complex. In this analysis of dose-response, instead of using single risk factor levels related to BMI, levels of a multivariable regression derived BMI formula outputs in quartiles were related to mean BMIs in those quartiles. Dose-response of BMI formula estimates versus mean BMI were included in the testing of consistency above.

Scoring of dose-response (biologic gradient) was based on this mean absolute difference when the BMI formula output was divided into quartiles:

5= The BMI formula output versus mean BMI absolute differences from each of the four quartiles averages \( \leq 0.40 \) BMI units.

4= The BMI formula output versus mean BMI absolute differences from each of the four quartiles averages \( \leq 0.50 \) units.

3= The BMI formula output versus mean BMI absolute differences from each of the four quartiles averages \( \leq 0.60 \) BMI units.

2= The BMI formula output versus mean BMI absolute differences from each of the four quartiles averages \( \leq 0.70 \) BMI units.

1= The BMI formula output versus mean BMI absolute differences from each of the four quartiles averages \( \leq 1.0 \) BMI units.

0= The BMI formula output versus mean BMI absolute differences from each of the four quartiles averages > 1.0 BMI units.
4. Temporality: The dictionary defines the noun temporality as: “The state of existing within or having some relationship with time.” Bradford Hill said, "Temporality refers to the necessity that the cause to precede the effect in time. This criterion is unarguable, insofar as any claimed observation of causation must involve the putative cause C preceding the putative effect D. It does not, however, follow that a reverse time order is evidence against the hypothesis that C can cause D. Rather, observations in which C followed D merely shows that C could not have caused D in these instances; they provide no evidence for or against the hypothesis that C can cause D in those instances in which it precedes D."

Dr. Bradford Hill was an occupational physician before the current availability in nutritional epidemiology of data on trends over 28 years of 18 components of worldwide diets along with the global BMI trend. Consequently, now it is fair to test temporality by deriving a standardised multiple regression formula with BMI trend, measured by the slope of the least squared regression line (LSRL) over 1990-2017, as (dependent variable). The independent variables would consist of the LSRL trends over 1990-2017 of the same dietary components, physical activity, and other variables as in the original BMI formula. We considered the strength (r) of the BMI trend formula versus the BMI trend an appropriate measure of temporality. Any risk factor was excluded from the trend BMI formula if its sign (+ or -) did not match the sign in the original BMI formula.

With a standardised multiple regression methodology similar to that of deriving risk factor percent weights above, a standardised BMI trend (dependent variable) versus
standardised risk factor trends (independent variables) formula was derived with the risk factor parameter estimates adjusted to equate to trend percent weights.

Scoring of temporality: The Pearson correlation coefficient, $r$, of the resulting multiple regression derived BMI trend versus risk factors trends formula:

5 = $r \geq 0.50$ and $P < 0.0001$

4 = $0.50 > r \geq 0.40$ and $P < 0.0001$

3 = $0.40 > r \geq 0.30$ and $P < 0.0001$

2 = $0.30 > r \geq 0.20$ and $P < 0.0001$

1 = $0.20 > r \geq 0.10$ and $P < 0.0001$

0 = $r < 0.10$ or $P \geq 0.0001$

Analogy: High BMI is among the four metabolic risk factors that are strongly associated with cardiovascular diseases, cancers, and other non-communicable diseases. The other major metabolic risk factors for non-communicable diseases are high systolic blood pressure (SBP), high low density lipoprotein cholesterol (LDL-C), and high fasting plasma glucose (FPG).

We tested analogy by the number of the diet, physical activity, and other variables in the percent weight of risk factor BMI formula that were also in the percent weight of risk factor formulas of the other metabolic factors and including only risk factors concordant in direction (+ or -) with the coefficients of percent weights of risk factors in the BMI formula.

Scoring of analogy:
5= At least three-quarters of BMI formula risk factors are also in the risk factor formulas of all three of the other metabolic factors (SBP, LDL-C, and FPG) and have concordant signs (+ or -) for each risk factor.

4= At least three-quarters of BMI formula risk factors are also in the risk factor formulas of two of the three other metabolic factors (SBP, LDL-C, and FPG) and have concordant signs (+ or -) for each risk factor.

3= At least three-quarters of BMI formula risk factors are also in the risk factor formulas of one of the three other metabolic factors (SBP, LDL-C, and FPG) and have concordant signs (+ or -) for each risk factor.

2= At least two-thirds of BMI formula risk factors are also in the risk factor formulas of one of the three other metabolic factors (SBP, LDL-C, and FPG) and have concordant signs (+ or -) for each risk factor.

1= At least half of BMI formula risk factors are also in the risk factor formulas of one of the three other metabolic factors (SBP, LDL-C, and FPG) and has concordant signs (+ or -) for each risk factor.

0= None of the above.

6. Plausibility: To test plausibility, we looked to find if any of our findings were at variance with the preponderance of studies published. We searched the medical literature particularly for systematic reviews of the relationships of foods and other variables with BMI.

Scoring of plausibility:
5= None of the current findings were at variance with the preponderance of the medical literature.

4= One of the current findings was at variance with the preponderance of the medical literature.

3= Two of the current findings were at variance with the preponderance of the medical literature.

2= Three of the current findings were at variance with the preponderance of the medical literature.

1= Four of the current findings were at variance with the preponderance of the medical literature.

0= Five or more of the current findings were at variance with the preponderance of the medical literature.

7. Experiment: Dr. Bradford Hill thought that evidence drawn from experimentation, including in epidemiologic studies, may lead to the strongest support for causal inference.¹

We used a cross validation method to assess Bradford Hill’s “experiment” criterion.

Random number generation of 20 subgroups each with 100 cohorts derived 20 standardised BMI formulas to compare with the standardised worldwide BMI formula.

Scoring of experiment:

5= At least 15 subgroups BMI formulas included all same risk factors with the same signs of coefficients as the worldwide BMI formula.
4=At least 10 subgroups of BMI formulas included all same risk factors with the same
signs of coefficients as the worldwide BMI formula.

3=All 20 subgroups of BMI formulas included at least three-quarters of same risk factors
with the same signs of coefficients as the worldwide BMI formula.

2=At least 10 subgroups of BMI formulas included at least three-quarters of same risk
with the same signs of coefficients factors as the worldwide BMI formula.

1= At least 10 subgroups of BMI formulas included at least half of same risk factors with
the same signs of coefficients as the worldwide BMI formula.

0=None of the above.

8. Coherence: According to Dr. Bradford Hill, “…cause and effect interpretation of our data
should not seriously conflict with the generally known facts of the natural history and
biology of the disease.”^1 In this analysis of BMI associated with BMI formula estimates,
coherence was the numerical total score of the above seven relevant causality criteria
each on a 0-5 scale. The maximum score was 40.

Scoring of coherence:

5=Score on the first seven Bradford Hill causation criteria=35-40.

4=Score on the first seven Bradford Hill causation criteria=30-34.

3=Score on the first seven Bradford Hill causation criteria=25-29.

2=Score on the first seven Bradford Hill causation criteria=20-24.

1=Score on the first seven Bradford Hill causation criteria=15-24.

0=Score on the first seven Bradford Hill causation criteria<15.
Reference

1. Hill AB. THE ENVIRONMENT AND DISEASE: ASSOCIATION OR CAUSATION?
Proceedings of the Royal Society of Medicine. 1965; 58(5):295-300.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1898525/
