Supplemental Table 1: Components of Günther’s DASH Index

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **<1800 kcal/day** | **1800x<2300 kcal/day** | **2300x<2850 kcal/day** | **2850 kcal/day** |
| **Score component** | Max Score | Standard for Max Score | Standard for Min Score | Standard for Max Score | Standard for Min Score | Standard for Max Score | Standard for Min Score | Standard for Max Score | Standard for Min Score |
| **Grains** |  |  |  |  |  |  |  |  |  |
|  **Total** | 5 | ≥6 servings/day | 0 servings/day | ≥6 servings/day | 0 servings/day | ≥10 servings/day | 0 servings/day | ≥12 servings/day | 0 servings/day |
|  **High fiber** | 5 | ≥50% of daily servings | 0% of daily servings | ≥50% of daily servings | 0% of daily servings | ≥50% of daily servings | 0% of daily servings | ≥50% of daily servings | 0% of daily servings |
| **Vegetables**  | 10 | ≥3 servings/day | 0 servings/day | ≥4 servings/day | 0 servings/day | ≥5 servings/day | 0 servings/day | ≥6 servings/day | 0 servings/day |
| **Fruit**  | 10 | ≥4 servings/day | 0 servings/day | ≥4 servings/day | 0 servings/day | ≥5 servings/day | 0 servings/day | ≥6 servings/day | 0 servings/day |
| **Dairy** |  |  |  |  |  |  |  |  |  |
|  **Total**  | 5 | ≥2 servings/day | 0 servings/day | ≥2 servings/day | 0 servings/day | ≥3 servings/day | 0 servings/day | ≥3 servings/day | 0 servings/day |
|  **Low-fat** | 5 | ≥75% of daily servings | 0% of daily servings | ≥75% of daily servings | 0% of daily servings | ≥75% of daily servings | 0% of daily servings | ≥75% of daily servings | 0% of daily servings |
| **Meat, poultry, fish, eggs** | 10 | ≤2 servings/day | ≥4 servings/day | ≤2 servings/day | ≥4 servings/day | ≤2 servings/day | ≥4 servings/day | ≤2 servings/day | ≥4 servings/day |
| **Nuts, seeds, legumes**  | 10 | ≥3 servings/week | 0 servings/week | ≥4 servings/week | 0 servings/week | ≥1 serving/day | 0 servings/day | ≥1 serving/day | 0 servings/day |
| **Fats, oils**  | 10 | ≤2 servings/day | ≥4 servings/day | ≤3 servings/day | ≥6 servings/day | ≤3 servings/day | ≥6 servings/day | ≤4 servings/day | ≥8 servings/day |
| **Sweets** | 10 | ≤5 servings/week | ≥10 servings/week | ≤5 servings/week | ≥10 servings/week | ≤2 servings/day | ≥4 servings/day | ≤2 servings/day | ≥4 servings/week |

Intakes between minimum and maximum servings/day were scored proportionally according to the intake of the given food group. Scoring is from Günther ALB, *et al.* Hypertension ﻿2009;53:6-12 [1]. As an example, if an individual reported an average caloric intake of 2000 kcal/d and consumed 2 servings per day of total fruits, they would be assigned a score halfway between the maximum and minimum score (that is, 5), for Total Fruit because they were half way between the maximum threshold (4 or more servings per day), and the minimum, (0 servings per day). This score, together with scores from other food groups, were summed to achieve an overall score ranging from 0 to 80.

Supplemental Table 2: Food frequency questionnaire items utilized to form food groups according to Günther DASH diet scores.

|  |  |
| --- | --- |
|  | Günther’s Index |
| Total grains | Dark bread (including whole wheat, rye, pumpernickel, other high-fiber bread); high fiber, bran or granola cereals, shredded wheat; 0.5 \* cooked cereals (including oatmeal, cream of wheat, grits).Pizza; burritos, including breakfast burritos, soft taco with flour tortillas; enchiladas, tamales, tacos, tostadas, chalupas, other Mexican dishes with corn tortillas, including nachos with chili; 0.5\*other broth-based soups (including caldo, tortilla soup); white bread (including sandwiches, hamburger or hotdog buns, bagels, baguettes, pita bread, English muffin); Biscuits, scones, croissants, muffins, fry bread (popover), hush puppies; flour tortilla (by itself, not in burritos, etc.); corn tortilla (by itself, not in enchiladas, etc.); corn bread, corn muffins; highly fortified cereals such as Product 19, Total, or Most; cold cereals that are low in sugar such as Corn Flakes, Rice Krispies, KIX; Sweetened cold cereals, such as Frosted Flakes, Fruit Loops; cooked cereals (including oatmeal, cream of wheat, grits).Salty snacks such as crackers, potato chips, corn chips, tortilla chips, pretzels, popcorn.Rice (plain); pasta, noodles, fideo, couscous (without cheese or tomato sauce); spaghetti, lasagna, other pasta or mixed dishes with tomatoes or tomato sauce, including spanish rice; mixed dishes with cheese (including macaroni and cheese, chile rellenos, cheese quesadillas, quiche). |
| High-fiber grains | Dark bread (including whole wheat, rye, pumpernickel, other high-fiber bread); high fiber, bran or granola cereals, shredded wheat; 0.5 \* cooked cereals (including oatmeal, cream of wheat, grits). |
| Whole grains | - |
| Vegetables | Tomatoes, tomato juice, (including chopped tomatoes and onion as a condiment); 0.5\*salsa picante, taco sauce; spaghetti, lasagna, other pasta or mixed dishes with tomatoes or tomato sauce, including spanish rice; 0.5\*Pizza; 0.5\*enchiladas, tamales, tacos, tostadas, chalupas, other Mexican dishes with corn tortillas, including nachos with chili and cheese; 0.5\*vegetable and tomato soup (including vegetable beef, minestrone).Broccoli; cauliflower or brussel sprouts; cole slaw, cabbage, sauerkrautstring beans, green beans; Peas; corn, posole, chicos; other green salad; any other vegetable, including cooked onions, summer squash, asparagus, sweet peppers, bok choy, okra, etc); red chili con carne INT: when used as condiment count 1/2 frequency; green chili con carne INT: when used as a condiment count 1/2 frequency; Asian food (including Chinese, Thai, Vietnamese); 0.5\*vegetable and tomato soup (including vegetable beef, minestrone); winter squash, baked squash; mustard greens, turnip greens, collards; carrots, or mixed vegetables containing carrots; spinach (cooked); spinach (raw).Sweet potatoes, yams; other potatoes such as boiled, baked, mashed, potato salad.French fries, fried potatoes. |
| Fruit | Apples, applesauce, pears; Bananas; 0.85\*peaches, apricots, nectarines (canned, frozen, or dried, whole year); 0.15\*peaches apricots, nectarines (fresh, when in season); 0.15\*cantaloupe (when in season); 0.15\*strawberries (fresh, when in season); Oranges; Grapefruit; dried fruits, including raisins, prunes, figs; any other fruit (other melon, grapes, berries, fruit cocktails persimmon, etc. ); avocado, guacamole.Orange juice or grapefruit juice; other fruit juice with vitamin c, including fortified fruit drinks, hi-c, kool-aid, cranberry juice, tang; any other fruit juice (apple juice, grape juice). |
| Total dairy | Cottage cheese, ricotta cheese.Pizza; mixed dishes with cheese (including macaroni and cheese, chile rellenos, cheese quesadillas, quiche); 0.5\*enchiladas, tamales, tacos, tostadas, chalupas, other Mexican dishes with corn tortillas, including nachos with chili and cheese; cream soups; cheese (cheddar, American, cream cheese, parmesan, Velveeta, other cheeses or cheese spreads; including on sandwiches or as snacks).Low-fat flavored yogurt (2% or non-fat); 2% milk and beverages with 2% milk (not including on cereal); skim milk, 1% or buttermilk and beverages made with these (not including on cereal).Flavored yogurt (regular, from whole milk); whole milk and beverages with whole milk (not including on cereal); 0.5\*milk in coffee or tea; 0.5\*cream (real) or half-and-half in coffee or tea. |
| Low-fat dairy | Low-fat flavored yogurt (2% or non-fat); 2% milk and beverages with 2% milk (not including on cereal); skim milk, 1% or buttermilk and beverages made with these (not including on cereal). |
| Meat, poultry, eggs, fish | Fried fish or fish sandwich; tuna fish, salmon, sardines (including tuna salad, tuna casserole); shell fish (shrimp, lobster, crab, oysters, mussels, etc); other broiled or baked fish, including trout, sloe, halibut, etc.Eggs (include omelettes, fritatta).Hamburgers, cheeseburgers, meat loaf, picadillo, carne guisada (asada); beef (steaks, roasts, etc. including on sandwiches); beef stew or pot pie with carrots or other vegetables; pork, including chops, roasts or ribs; ham, ham hocks (including ham on sandwiches); game, including venison, rabbit; liver, including chicken livers; burritos, including breakfast burritos, soft taco with flour tortillas; green chili con carne (when used as a condiment count 1/2 frequency); Asian food (including Chinese, Thai, Vietnamese); liverwurst; hot dogs (include pork, beef, turkey); bologna, salami, spam, other lunch meats (excluding ham); bacon; sausage, chorizo.Chicken, turkey or wild fowl (roasted, broiled, or ground including on sandwiches); chicken or turkey stew or pot pie with carrots or other vegetables; fried chicken. |
| Red & processed meat | - |
| Nuts, seeds, dried beans | Nuts and seeds, including peanuts, peanut butter, pine nuts, sunflower seeds.Tofu/Tempeh.Refried beans (as side dish, not including those in burritos, etc); other beans such as pintos, black beans, garbanzos, baked beans, or lentils; 0.5\*burritos, including breakfast burritos, soft taco with flour tortillas. |
| Fats and oils | 0.5\*diet salad dressing, diet mayonnaises (including on sandwiches); salad dressing, mayonnaise, tartar sauce (including on sandwiches); butter, margarine, or other fat on vegetables, potatoes, rice, etc; 0.25\*tuna fish, salmon, sardines (including tuna salad, tuna casserole); margarine on bread or roll; butter on bread or rolls; gravies made with meat drippings, or white sauce; 0.5\*non-dairy creamer in coffee or tea. |
| Sweets | Regular soft drinks (including colas, 7-up, etc); lemonade, sweetened mineral water.Doughnuts, cookies, cakes, pastry, brownies, sopapillas, pan dulce; pumpkin pie, sweet potato pie, empanadas with pumpkin; other piesSugar, molasses, or honey added to cereal; chocolate including Hershey's kisses, M&M's, chocolate candy bars; other candy, jelly, honey, brown sugar, jams, or molasses, including on bread or other foods; sugar in coffee or tea, or honey in tea (not including artificial sweeteners). |
| Sugar-sweetened beverages | - |
| Sodium | - |

Supplemental Table 3: Baseline characteristics of IRAS participants included by incident type 2 diabetes mellitus at 5-year follow-up.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **No Incident T2DM** | **Incident T2DM** | **p-value** |
| **N** | 471 | 99 |  |
| **Age (years)**  | 54.67 (8.57) | 56.59 (7.93) | 0.040 |
| **Sex (Male (%))** | 214 (45.4%) | 38 (38.4%) | 0.241 |
| **Ethnicity** |  |  | 0.975 |
|  **African American** | 117 (24.8%) | 25 (25.3%) |  |
|  **Hispanic** | 153 (32.5%) | 31 (31.3%) |  |
|  **Caucasian** | 201 (42.7%) | 43 (43.4%) |  |
| **Smoking status** |  |  | 0.031 |
|  **Never Smoked** | 227 (48.2%) | 39 (39.4%) |  |
|  **Past Smoker** | 182 (38.6%)  | 37 (37.4%) |  |
|  **Current Smoker** | 62 (13.2%)  | 23 (23.2%) |  |
| **Energy Expenditure (kcal/day)** | 2117 (389) | 2098 (371) | 0.661 |
| **Energy Intake (kcal/day)** | 1850 (738)  | 1822 (777)  | 0.732 |
| **BMI (kg/m2 )** | 27.8 (5.08)  | 31.4 (6.54) | <0.001 |
| **Waist Circumference (cm)** | 89.1 (12.35)  | 95.6 (13.91)  | <0.001 |
| **Mean Günther Score** | 51.6 (9.48)  | 51.8 (8.79)  | 0.842 |

Values presented as Mean (SD) unless specified otherwise.

*Supplemental Figure 1: Consort diagram of included IRAS cohort sample.*

IRAS Sample (n=1625)

* Normal Glucose Tolerance (n=719)
* Impaired Glucose Tolerance (n=369)
* Type 2 Diabetes (n=537)

Exclusion for those with Type 2 Diabetes at baseline (n=537)

Exclusion for those who did not return for 5-year follow-up (n=181)

Exclusion for those with missing data (n=337)

* Missing food group data (n=20)
* Missing metabolite data (n=298)
	+ Subjects with all metabolite data missing (n=151)
	+ Subjects with missing data for excluded metabolites (n=147)
* Missing nutrient and/or sodium data (n=15)
* Missing caloric intake data (n=4)

Assessed for eligibility (n=3416)

Excluded (n=1791)

* Not meeting inclusion criteria (n=404)
* Declined to participate (n=1094)
* Eligible but not enrolled (n=289)
* Enrolled but excluded due to incomplete OGTT results (n=4)

Sample with Normal Glucose Tolerance and Impaired Glucose Tolerance (n=1088)

Sample with both Baseline and 5-year follow-up measurements (n=907)

**IRAS Sample Used: 570**

Normal Glucose Tolerance: 380

Impaired Glucose Tolerance: 190



Supplemental Figure 2: Percentage of missing measurements for each metabolite.



Supplemental Figure 3: Root mean squared error of prediction for number of components included in the partial least squares model conditioned on Günther’s DASH adherence scores. CV: Cross-validated, adjCV: Adjusted cross-validated.



*Supplemental Figure 4: Root mean squared error of prediction for number of components included in the partial least squares model conditioned on Gunther’s DASH adherence scores using imputed metabolite data. CV: Cross-validated, adjCV: Adjusted cross-validated.*

Supplemental Table 4: Table of metabolites with strongest loadings (0.2) on Components 1 and 2 of the PLS model conditioned on Günther’s DASH scores using imputed metabolite data.

|  |  |  |
| --- | --- | --- |
|  | Component 1 | Component 2 |
| L-carnitine (l-carn) | -0.30 |  |
| Stearoylcarnitine AC18:0 | -0.29 | 0.21 |
| Oleoylcarnitine AC18:1 (n-9) | -0.29 | 0.25 |
| Valerylcarnitine AC5:0 | -0.28 |  |
| Hexanoylcarnitine AC6:0 | -0.27 |  |
| Palmitoylcarnitine AC16:0 | -0.26 | 0.34 |
| Propionylcarnitine AC3:0 | -0.24 |  |
| Myristoylcarnitine AC14:0 | -0.24 | 0.34 |
| Eicosapentaenoic acid 20:5 (n-3) | 0.24 |  |
| Acetylcarnitine AC2:0 | -0.23 | 0.24 |
| Dodecanoylcarnitine AC12:0 | -0.23 | 0.31 |
| Fatty acid 18:4 (n-3)  | 0.23 |  |
| Linoleoylcarnitine AC18:2 (n-6) | -0.21 | 0.25 |
| Butyrylcarnitine AC4:0 | -0.21 |  |
| Butyrobetaine | -0.21 |  |
| Hexanoylcarnitine AC6:0 |  | 0.30 |
| Docosapentaenoic acid (DPA) 22:5 (n-3) |  | 0.28 |
| Eicosapentaenoic acid (EPA) 20:5 (n-3) |  | 0.27 |
| Octanoylcarnitine AC8:0 |  | 0.26 |
| Decanoylcarnitine AC10:0 |  | 0.26 |
| Docosahexaenoic acid (DHA) |  | 0.22 |



*Supplemental Figure 5: Circle plot from PLS model showing the clustering of metabolites conditioned on Gunther’s DASH score using imputed metabolite data. The percent explained variance of the complete metabolite dataset by each component is shown in brackets on each axis. The solid circular line represents an explained variance of 100% for a metabolite by a component, while the dashed line represents an explained variance of 50%. Metabolites between these lines represent variables that strongly explain the underlying structure of the data.*



*Supplemental Table 6: Association of PLS model generated component 1 (A) & 2 (B) with Gunther’s PLS scores using imputed metabolite data. Line represents a fitted linear regression model with its equation and coefficient of determination presented on the top left of the graph area. Shaded area indicates 95% confidence interval. A: y=51.7 + 1.63 x; R2=0.09; Correlation=0.30. B: y=51.7 + 0.754 x; R2=0.032; Correlation 0.17.*

*Supplemental Table 5: Results from multiple linear regression assessing the association between Gunther’s PLS component scores generated using imputed metabolite data and Insulin Sensitivity (SI).*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | β-Coefficients | 95% Confidence Interval | P-Value |  |
| Component 1 |  |  |  |  |
| Model 1 | 0.01 | -0.01; 0.04 | 0.28 |  |
| Model 2 | 0.00 | -0.02; 0.03 | 0.69 |  |
| Model 3 | 0.00 |  -0.03; 0.02 | 0.68 |  |
| COmponent 2 |  |  |  |  |
| Model 1 | 0.06 | 0.04; 0.08 | <0.001 |  |
| Model 2 | 0.04 | 0.03; 0.06 | <0.001 |  |
| Model 3 | 0.03 | 0.01; 0.04 | <0.001 |  |

Model 1: Incidence of Type 2 Diabetes Mellitus at 5-year follow-up ~ Component Scores. Model 2: Model 1 + Age + Sex + Ethnicity + Glucose tolerance status + Family history of diabetes + Education Level + Smoking status + Caloric intake + Total Energy Expenditure.

Model 3: Model 2 + BMI at baseline.

*Supplemental Table 6: Results from multiple linear regression assessing the association between Gunther’s PLS component scores generated using imputed metabolite data and Disposition Index (DI).*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | b-Coefficients | 95% Confidence Interval | P-Value |  |
| Component 1 |  |  |  |  |
| Model 1 | -0.06 | -0.11; -0.00 | 0.049 |  |
| Model 2 | -0.03 | -0.08; 0.03 | 0.363 |  |
| Model 3 | -0.04 | -0.91; 0.01 | 0.156 |  |
| COmponent 2 |  |  |  |  |
| Model 1 | 0.05 | 0.01; 0.10 | 0.028 |  |
| Model 2 | 0.03 | -0.01; 0.07 | 0.106 |  |
| Model 3 | 0.01 | -0.03; 0.05 | 0.519 |  |

Model 1: Incidence of Type 2 Diabetes Mellitus at 5-year follow-up ~ Component Scores. Model 2: Model 1 + Age + Sex + Ethnicity + Glucose tolerance status + Family history of diabetes + Education Level + Smoking status + Caloric intake + Total Energy Expenditure.

Model 3: Model 2 + BMI at baseline.

Supplemental Table 7: Results from multiple logistic regression assessing the association between Gunther’s PLS component scores generated using imputed metabolite data and incident type 2 diabetes at 5-year follow-up within the current IRAS sample.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Odds Ratio | 95% Confidence Interval | P-Value |  |
| Component 1 |  |  |  |  |
| Model 1 | 0.96 | 0.86 -1.08 | 0.54 |  |
| Model 2 | 0.98 | 0.85 - 1.13 | 0.81 |  |
| Model 3 | 1.02 | 0.88 -1.18 | 0.78 |  |
| COmponent 2 |  |  |  |  |
| Model 1 | 0.85 | 0.78 - 0.93 | <0.001 |  |
| Model 2 | 0.87 | 0.78 - 0.97 | 0.009 |  |
| Model 3 | 0.88 | 0.79 - 0.98 | 0.016 |  |

Model 1: Incidence of Type 2 Diabetes Mellitus at 5-year follow-up ~ Component Scores. Model 2: Model 1 + Age + Sex + Ethnicity + Glucose tolerance status + Family history of diabetes + Education Level + Smoking status + Caloric intake + Total Energy Expenditure.

Model 3: Model 2 + BMI at baseline.

Supplemental Table 8: Results from multiple logistic regression assessing the association between Gunther’s PLS component scores generated using imputed metabolite data and incident type 2 diabetes at 5-year follow-up within the current IRAS sample. Top and bottom 1% of caloric intakes were excluded.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Odds ratio** | **95% confidence interval** | **p-value** |
| **Component 1** |  |  |  |
|  **Model 1** | 0.97 | 0.86; 1.10 | 0.66 |
|  **Model 2** | 1.01 | 0.87; 1.16 | 0.94 |
|  **Model 3** | 1.02 | 0.88; 1.19 | 0.76 |
| **Component 2** |  |  |  |
|  **Model 1** | 0.86 | 0.78; 0.95 | 0.002 |
|  **Model 2** | 0.89 | 0.80; 0.99 | 0.027 |
|  **Model 3**Model 1: Incidence of Type 2 Diabetes Mellitus at 5-year follow-up ~ Component Scores. Model 2: Model 1 + Age + Sex + Ethnicity + Glucose tolerance status + Family history of diabetes + Education Level + Smoking status + Caloric intake + Total Energy Expenditure. Model 3: Model 2 + BMI at baseline. | 0.89 | 0.80; 0.99 | 0.040 |

Supplemental Table 9: Results from multiple logistic regression assessing the association between Gunther’s PLS component scores generated using imputed metabolite data and incident type 2 diabetes at 5-year follow-up within the current IRAS sample.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Odds ratio** | **95% confidence interval** | **p-value** |
| **Component 1** |  |  |  |
|  **Model 1** | 0.97 | 0.87; 1.10 | 0.68 |
|  **Model 2** | 1.02 | 0.88; 1.18 | 0.78 |
|  **Model 3** | 1.04 | 0.90; 1.21 | 0.61 |
| **Component 2** |  |  |  |
|  **Model 1** | 0.86 | 0.79; 0.95 | 0.002 |
|  **Model 2** | 0.89 | 0.80; 0.99 | 0.037 |
|  **Model 3** | 0.90 | 0.80; 1.00 | 0.052 |

Model 1: Incidence of Type 2 Diabetes Mellitus at 5-year follow-up ~ Component Scores.

Model 2: Model 1 + Age + Sex + Ethnicity + Glucose tolerance status + Family history of diabetes + Education Level + Smoking status + Caloric intake + Total Energy Expenditure + Blood Pressure Medications + Lipid Medications.

Model 3: Model 2 + BMI at baseline.

*Statistical Methods Appendix:*

PLS models were generated using a training dataset of 500 of the 570 included IRAS participants’ metabolite measurements. This training dataset provides both predictor and response variables with which the PLS algorithm can create models [2]. The first local minimum root mean squared error of prediction was used to identify the optimal number of components to retain in the PLS models. This test indicates the number of components to retain in the PLS model for optimal prediction while avoiding overfitting [3]. These PLS models restricted to the number of components previously identified to be retained were then validated using the “Leave One Out” method and were cross validated on a testing dataset of the remaining 70 IRAS participants. This testing dataset uses solely the predictor variables of metabolite measurements to estimate response variables of DASH index scores [2]. Based on generated prediction models, an assessment could be made of the performance of the PLS models.

*References*

[1] A. L. B. Günther, A. D. Liese, R. A. Bell, D. Dabelea, J. M. Lawrence, B. L. Rodriguez, D. A. Standiford, E. J. Mayer-Davis, *Hypertension* **2009**, *53*, 6–12.

[2] P. S. Gromski, H. Muhamadali, D. I. Ellis, Y. Xu, E. Correa, M. L. Turner, R. Goodacre, *Anal. Chim. Acta* **2015**, *879*, 10–23.

[3] B. Mevik, R. Wehrens, **2016**.