2023.05.19 Waist2 method, after talking with Mark and David via email for the 2nd time.

ANCOVA Model

Response = μ + DivGroup + Age + Gender + Ethnicity + Income + Education + kcal\*  + ε,

1 2 3 4 5 6 7

where DivGroup is the treatment effect, and Age, Gender, Ethnicity, Income, Education, and kcal intake were covariates, and ε is the error term.

QC-ed Total ()

|

No missing data in waist and BMI, and variables 1-7

|

Remove the 361 missing data in Income.

|

Make an OTU table

|

Remove outliers in nuts/seeds/legumes consumption amount in the OTU table.

|

Calc divGroup

|

Plot a boxplot of amount x DivGroup.

|

Make table of Demographic (var. 2-7) x DivGroup table. Percentages +- SD.

|

Make table of foods: PF\_LEG, PF\_TOTAL, V\_TOTAL, KCAL etc. x DivGroup.

There may be outliers that are only visible when I made divGroup table, and averages may be affected. If there is an outlier in those food variables, just remove them individually, I guess.

|

Kcal is different among DivGroups…

|

Make a table of nutrients: CARB, PUFA, etc./ 2000 kcal x DivGroup.

|

Test Waist vs. DivGroup and BMI vs. DivGroup using the ancova model shown above.

Done!!

Variable list <https://wwwn.cdc.gov/Nchs/Nhanes/2015-2016/DEMO_I.htm#RIAGENDR>

|  |  |
| --- | --- |
| BMI | BMXBMI |
| WAIST | BMXWAIST |
| Age | RIDAGEYR |
| Gender | RIAGENDR |
| Ethnicity | RIDRETH3 |
| Family IPR | INDFMPIR |
| Edu (<20 yo) | DMDEDUC3 |
| Edu (>=20) | DMDEDUC2 |
| kcal | KCAL |

Paragraphs that may be useful:

<Summary> We defined four dietary intake groups based on legume alpha-diversity (Shannon’s diversity, intake of multiple different nuts, seeds, and legumes consumed over 2 days) for n = 3,676. The highest legume alpha-diversity group, Div2, had a lower waist circumference (mean difference −3.6 cm, p <0.001, **Table**) and a lower BMI (mean difference −1.3, p <0.01) , after adjustment for age, gender, ethnicity, and education variables.

Legume diversity associated with higher energy intake

For KCAL, the ancova model with all covariates (except KCAL) indicated that all the covariates were significant, and KCAL was different among the DivGroups (p<0.0001, **Table 4**). The emmeans increased as the nuts/seeds/legumes consumption and their diversity increased (**Table 4**). The group that did not consume nuts/seeds/legumes had 175 kcal less calorie intake than those who consumed the most diverse nuts/seeds/legumes foods (p<0.0001, **Table 5**). In addition, the Div0 group that consumed only one nuts/seeds/legumes food had 131 kcal less calorie intake than Div2 (p<0.0001, **Table 5**), and DivNA group consumed 111 kcal less calories than the Div1 group (p<0.01, **Table 5**).