June 11th, 2024

Files for the FRESH Tanzania 24-hour dietary recall among women of reproductive age

Authors: Fusta Azupogo; Charles D. Arnold (UC Davis)

Please note: "C." always denotes the SAS library location on Fusta's PC

Data cleaning and intake estimation steps

1. Cleaning took place in two phases
    a. Phase I - Data collected up to 11/30/2023
    b. Phase II - Data collected 12/1/2023 through the end of the baseline.
2. The file "C.FRESH_diet_5" is the processed file from phase I of the data cleaning and processing for intake estimation.
3. The "C.FRESHDIET_PART2_5" is the processed file from phase II of the data cleaning and processing for intake estimation.
4. The "C.FRESHdiet_24hr" combined both the Phase I and II files
    a. The "c.FRESHdiet_24hr" was exported to Excel, and the ingredient list of recipes such as ugali and pilau was inspected, and these recipes were re-classified, e.g. pilau without meat, pilau with beef, pilau with goat, pilau with chicken etc.
5. The file was re-imported into SAS as "C.FRESHdiet_24hr_1" and went through data steps to create new variables for date and day as C.FRESHdiet_24hr_2
6. The file "C.FRESHdiet_24hr_3" removes all duplicates in the dataset. The duplicates resulted from duplicating the same subject and day recalls in the Phase I and II data processing files.

Recipes disaggregation steps

7. The file from step 6 was merged with a file called "Food_recipe_list_1" to obtain c.FRESHdiet_24hr_4. The aim was to include a binary variable called "Recipe_Item" coded as 1(yes) and 0(No), denoting whether the food item from pass 1 of the 24hR was a recipe or not. This was subsequently used to disaggregate the recipes. Some food items were also re-coded as recipes in this file based on the ingredient description.
8. The files "c.Non_recipes" outputted all non-recipes that did not need to go through recipe disaggregation. This is important as all food codes not disaggregated will disappear from the file with disaggregated recipes. The file is later added to the file containing the disaggregated recipes.
9. The file "c.Standard_recipes_fresh" contains all recipe data for disaggregating the recipes in the intake file. The file was obtained by merging four different files described below. A column in this file "Recipe_source" denotes the source of the recipe data
    a. C.STANDARD_RECIPES_12: Contains recipes formulated from the recipe formulation work in the field. It includes 54 recipes (40.0%)
    b. C.formulated_RECIPES: These are recipes formulated based on the primary recipe file. For example, chicken pilau was formulated by replacing beef with chicken in "beef pilau" etc. It contains 24 recipes (17.8%)
    c. C.Caroline_RECIPES: Recipe data from a previous project conducted by UC Davis (led by Christine and Caroline) in Tanzania. It includes 17 recipes (12.7%)
    d. C.Tanzania_FCT_RECIPES: Recipe data taken from the Tanzania Food Table including 40 recipes (29.6%)

10. After going through some preliminary data steps, the file with disaggregated recipes is "c.fresh_24hR_recipes". The preliminary data files include "c.Freshdiet_24hr_5" and "c.Standard_recipes_fresh_1," which were obtained from "c.Freshdiet_24hr_4" and "c.Standard_recipes_fresh", respectively.
11. The dataset "c.FRESH_24hR" is obtained by adding up the file with non-recipes with the file containing the disaggregating recipes (c.fresh_24hR_recipes + c.Non_recipes)
12. The file "c.FRESH_24hR_1" is a preliminary data step, re-arranging the variables in a sequence, computing some new variables, etc.

Merging with the Food composition Table and Retention Factors file for Energy and Nutrient intake estimation

13. The cleaned and final Food Composition Table (FCT) file is "c.fct_fresh_final." The cleaned retention factors file, adopted from USDA, is "c.retention_factors_1"
14. The FCT file is merged by food ingredient code with the intake file containing all disaggregated recipes (merge c. FRESH_24hR_1 and c.fct_fresh_final) to obtain the file c.FRESH_24hR_2
15. The retention factors file is merged with the output file from step 14 to obtain c.FRESH_24hR_3 (merge c.retention_factors_1 and c.FRESH_24hR_2 by retention code).

Energy and nutrient intake estimation

16. The file "c.FRESH_24hR_4" outputs the estimated energy and nutrient intake at the ingredient level. This file is used to obtain all summary files of intake, e.g. intake by Subject ID and day, Subject ID, day, food grouping, etc. This file must always be the starting point for obtaining any desired summary file of the intake.

Summary files of Energy and nutrient intake

Note: These summary files are required for data analysis involving the 24hR. The interest of the analysis determines which summary file should be used.

17. The first summary file, "c.FRESH_24hR_6," outputs the estimated energy and nutrient intake for each subject (HHID) and recall day (recall 1 or 2). Preliminary data step files from this process include c.FRESH_24hR_5 and c.FRESH_24hR_7, which was used to obtain "c.FRESH_24hR_8", including the treatment arm, district, region, etc., based on comments from the FRESH team. Hence, the final cleaned file for energy and nutrient intake by subject ID (HHID) and recall number is **"**c.FRESH_24hR_8",. The file is in a wide format and does not require any transposing. The final file **was renamed as "mdr_final_hhid_recall"**
18. The summary file "FRESH_24hR_mealcode_1" outputs estimated energy and nutrient intake by subject ID, recall day and meal code (breakfast, lunch, dinner, snack, other). Updated to obtain **c**.FRESH_24hR_mealcode_3, which includes the treatment arm, district, region, etc. This file is in the long format. **The final file was renamed as mdr_final_mealcode**
19. The summary file "c.FRESH_24hR_TZ_foodgroup_1" contains estimated energy and nutrient intake by subject ID, recall day and Food grouping used in the TZ-FCT. This file was updated to obtain "c.FRESH_24hR_TZ_foodgroup_3" which includes the treatment arm, district, region, etc. This file is also in the long format**. The final file was renamed as "mdr_final_TZ_food_groups"**
20. The final diet metric summary files include the following: These files were transposed and put in a wide format for use; they do not require transposing.

a. c.FRESH_MDD_W_3: This outputs data on the food intake (gram/day) groups from the minimum dietary diversity for women (MDD-W) of reproductive age. This includes only the 10 food groups from the MDD-W. Files from preliminary data steps include c.FRESH_MDD_W, c.FRESH_MDD_W_1 and transposed files (c.FRESH_MDD_W_2_1 and c.FRESH_MDD_W_2_2) used to obtain c.FRESH_MDD_W_2. **The final file was renamed as mdr_final_MDD_W**

b. The final summary file "**c. FRESH_MDD_W_extended_3**". This outputs data on the food intake (gram/day) groups from the minimum dietary diversity for women (MDD-W) of reproductive age and includes the optional food groups. Files from preliminary data steps include c.FRESH_MDD_W_extended c.FRESH_MDD_W_extended_1, and two transposed files (c.FRESH_MDD_W_extended_2_1 and c.FRESH_MDD_W_extended_2_2) used to obtain c.FRESH_MDD_W_extended_2. **The final file was renamed as mdr_final_MDD_W_Extended**

c. The final summary file "c.FRESH_NOVA_3" outputs data on food intake (gram/day) following the NOVA classification. Files from preliminary data steps include c.FRESH_NOVA, c.FRESH_NOVA_1, c.FRESH_NOVA_2 and two transposed files (c.FRESH_NOVA_2_1, c.FRESH_NOVA_2_2) used to obtain c.FRESH_NOVA_2. **The final file was renamed as "mdr_final_NOVA"**

d. The final summary file **"**c.FRESH_GDQS_3**"** outputs data on food intake (gram/day) following the Global Diet Quality Score (GDQS). Preliminary data steps files include c.FRESH_GDQS, c.FRESH_GDQS_1 and two transposed files (c.FRESH_GDQS_2_1 and c.FRESH_GDQS_2_2) used to obtain c.FRESH_GDQS_2. **The final file was renamed as "mdr_final_GDQS"**

Dietary supplement intake data

21. The file **"**C.supplements_3**"** contains data on dietary supplement intake among women. The preliminary files from the data steps include C.supplements, C.supplements_1, and C.supplements_2. Interestingly, only 41 women reported any supplement intake, with 30 of this sample reporting iron and folate (IFAS) supplement intake. **The final file was renamed as "mdr_final_supplements."**

Please Note: All finale files for the 24hR begin with mdr_final…. and can be found in the Dropbox at

6. Data management/Data/Household/Baseline/3.Final/WRA