Hispanic Community Children's Health/ Study of Latino Youth (SOL Youth)

Dietary Data Overview

February 2016
Version INV2.0

Prepared by
HCHS/SOL Coordinating Center
Collaborative Studies Coordinating Center
UNC Department of Biostatistics

Daniela Sotres-Alvarez
Anna Maria Siega-Riz

Please send questions, suggestions and comments to dsotres@unc.edu

The Hispanic Community Children's Health
/ Study of Latino Youth (SOL Youth)

Dietary Data Overview, Methods and Guidelines
February 2016 INV 2.0

Table of Contents

Updates to SOL Youth Diet Data Release or Documentation

Version	Date	Description	Datasets	Documentation
1	10/03/2014	- 24-hr dietary recalls: seven NDSR files (2, 4, 7, 8, 9, 12 and 13) version 2012 - Dietary Behaviour Questionnaires: 8 forms (AFE,SFE, FNE, TVE, EDE, FSE, PCE and PPE)	_INV1	V1.0 (Oct 2014)
2	18/06/2015	- 24-hr dietary recalls: fifteen NDSR files (1,2,3, 4, 7, 8, 9, 12, 13, 14, 15, 16, 17, 18 and 19) version 2013	_INV1	V1.2 (Jun 2015)
3	02/05/2016	- NO changes to NDSR files released before; - NO changes to FOOD_GROUP_DERV; - NEW derived variable HEI2010 score and its intermediate variables are now included in SYOUTH_CHILD_PART_DERV	_INV2	V2.0 (Feb 2016)

FORWARD

Note to User of these SOL Youth Dietary Data Overview, Methods and Guidelines:

This document is not intended for direct citation or distribution outside of the immediate HCHS/SOL Study. It should be considered confidential and proprietary to HCHS/SOL investigators.

Diet Data Overview Section 1 presents a summary of dietary intake assessment conducted during SOL Youth baseline examination. **Section 2** describes SAS datasets from 24-hr dietary and supplement recalls and 30-day supplement recall. **Section 3** describes NCI method to predict usual nutrient intake. **Sections 4 and 5** have resources and references.

1 Dietary Intake Assessment in Sol Youth

As part of the SOL Youth clinic exam, dietary intake was assessed with two 24-hour dietary recalls. For details on 24-hour dietary recall, 24-hour and 30 day supplement use assessment procedures, see SOL Youth Manual of Operations 3 “Diet and Supplement”. Briefly,

1.1 Two 24-hour dietary and supplements recalls (24-hr and 30-day)

- **First dietary recall** was administered in-person at the field center.
- **Second dietary recall** was mostly telephone administered, and conducted at least five days and ideally within 45 days, following the initial examination interview.
- **Both 24hr dietary recalls were conducted using the NDSR software** developed by the Nutrition Coordinating Center (NCC) at the University of Minnesota which employs the multiple pass method. For the first dietary recall food models were used, and an amounts booklet was provided for estimating portion sizes in the 2nd recall. NDSR version 2012 (http://www.ncc.umn.edu/index.html) contains over 18,000 foods, 8,000 brand name products, many ethnic foods supplements and vitamins. The program provides values for 158 nutrients, nutrient ratios, and other food components. NDSR calculates nutrient intake and presents the data in several formats, including daily nutrient totals, nutrient amounts per individual food, daily totals compared to the Dietary Recommended Intakes (DRIs) and a new food group serving count system.
- **The NDSR Dietary Supplement Assessment Module (DSAM)** was used for the 24-hour dietary recall. This assessment utilizes the most currently available NHANES Supplement Database with enhancements from NCC and allows for the collection of 24-hour and/or 30-day intake of all dietary supplements and antacids. The goal of the dietary supplement recall is to assess use of all types of dietary supplements and over-the-counter antacids. Over-the-counter antacids are included in this assessment because many of these products contain calcium. **For the in-person interview**, the period covered for dietary supplement intake is the same time period covered by the 24-hour dietary recall and the past 30 days. **For the telephone (or second interview)** the DSAM was limited to the 24-hour dietary recall, and not past 30 day supplement use.

1.2 Dietary Behavior Questionnaires

Eight Sol Youth questionnaires collected information about dietary behavior. The Away from Home Foods form (**AFE**) solicits information regarding the frequency and location of foods consumed away from home. The School Food Environment form (**SFE**) assesses the youth’s perceptions of the school’s food environment and the frequency of foods consumed at school. The Food and Neighborhood Environment form (**FNE**) captures information about the availability of food in the participant’s surrounding area.

Other questionnaires that collect information on eating habits or feelings about eating include: Food Practices with TV/Video Viewing (**TVE**), Eating Disorders (**EDE**), Dietary and Physical Activity Support (**FSE**), Parenting for Eating and Physical Activity (**PCE**), and Parenting for Eating and Physical Activity (**PPE**) questionnaires.

2 SAS Datasets from 24-hr Dietary Recalls and Supplement Recalls (24-hr and 30-day)

Data release INV2 includes raw dietary data from 24-hr dietary recalls, 24 hr and 30 day supplement recalls. Among the 1,466 children in SOL Youth, 96% (n=1,428) have the 1st dietary recall (in-person at clinic visit) and 93% (n=1384) have the 2nd recall (most conducted by telephone).

2.1 24-hr Dietary and Supplement Recalls and 30-day Supplement Recalls (NDSR files)

NDSR files 1-4, 7-9, and 12-19 are raw data files for the 24 hour dietary and supplement recalls and the 30-day supplement recall. These are multiple-record-per-participant datasets (Table 2.1). NDSR files 5, 6, 10, 11, 20 and 21 were not used in SOL Youth. The included files have been reformatted from their original NDSR format to be consistent with HCHS/SOL and SOL Youth datasets and variable naming conventions (see SOL Youth Investigator Use Database Overview). Each dataset was given a four letter abbreviation (Y##A) where ## corresponds to NDSR file number. For example, dataset Y09A corresponds to NDSR file 9 "Food groups at the daily total level". The variables were named using the dataset name and a variable index, where the variable index corresponds to the column number in the NDSR manual. The SAS files contain variable labels with English descriptions of each variable.

The NDSR version used to collect the data for a participant can be found in Y04A17 (ranges from 2010-2012). All raw files were processed using version 13 (2013) of the NDSR software, which uses 2013 USDA database of nutrient contents of foods. **For information on NDSR Food Group Serving Count System and serving sizes see Appendix 10 of the NDSR 2013 Manual.** All NDSR files released (2-4, 7-9, and 12-19) excluded dietary recalls unreliable according to the interviewer (i.e. Y04A16 ≠ 0).

Table 2.1. SAS datasets from dietary and supplement recalls (NDSR raw data)

NDSR File	SOL Youth dataset	Dataset description	Key field 1	Key field 2[1]	Key field 3[1]
1	Y01A	Nutrients at the Component/Ingredient Level	ID	RECALLNUM	COMPID
2	Y02A	Nutrients at the whole food level	ID	RECALLNUM	FOODID
3	Y03A	Nutrients at the meal level	ID	RECALLNUM	MEALID
4	Y04A	Nutrients at the daily total level	ID	RECALLNUM	
7	Y07A	Food groups at the whole food level	ID	RECALLNUM	FOODID
8	Y08A	Food groups at the meal level	ID	RECALLNUM	MEALID
9	Y09A	Food groups at the daily total level	ID	RECALLNUM	
12	Y12A	Total 24 hour supplement intake	ID	RECALLNUM	
13[2]	Y13A	Averaged 30-day supplement intake	ID		
14	Y14A	Product file for 24 hour supplement intake	ID	RECALLNUM	PRDID
15[2]	Y15A	Product file for 30-day supplement intake	ID	SUPPLID	
16	Y16A	Product ingredients for 24 hour supplement intake	ID	RECALLNUM	INGID
17[2]	Y17A	Product ingredients for 30 day supplement intake	ID	INGID	
18	Y18A	Blend ingredients for 24 hour supplement intake	ID	RECALLNUM	BLDID
19[2]	Y19A	Blend ingredients for 30 day supplement intake	ID	BLDID	

NDSR files 5, 6, 10, 11, 20 and 21 were not used in SOL Youth nor in HCHS/SOL
[1]Descriptions of key field variables are as follows:
RECALLNUM – first or second 24 hour dietary recall
FOODID – Food file ID. MEALID plus a 3 digit index (ABBB or AABBB)
MEALID – 1-2 digit index for meal or eating occasion
PRDID or SUPPLID – 1-3 digit index for DSAM product file ID.
INGID – Ingredient ID. PRDID plus a 3 digit index (ABBB, AABBB, or AAABBB)
BLDID – Blend ingredient ID. INGID plus a 3 digit index (AAAABBB, AAAAABBB, or AAAAAABBB)
[2]30 day supplement intake was assessed only at the clinic visit. RECALLNUM was not needed as a key field.

2.2 Food Groups (SOLYOUTH_FOOD_GROUP_DERV)

The SAS multiple-record-per-participant dataset FOOD_GROUP_DERV contains derived variables for 50 broad food groups at the daily level created from Y09A (165 NDSR food codes at the daily level). Description of these 50 derived variables can be found in appendix A. Serving counts for individual NDSR food codes were added to create serving counts per day for each food group. **The dataset contains up to two observations per participant, depending on how many 24hr dietary recalls were collected, reliable according to interviewer (Y04A16) and cleaned at the 24hr recall level based on daily energy intake (Y04A20). See appendix B for details on data cleaning processes.**

2.3 Healthy Eating Index – 2010 (SOLYOUTH_CHILD_PART_DERV)

The Healthy Eating Index-2010 (HEI-2010; Guenther et al., 2013) is a measure of overall diet quality, independent of quantity, which can be used to assess compliance with the *2010 Dietary Guidelines for Americans* and to monitor changes in dietary patterns. It includes twelve dietary components (nine adequacy and three moderation components) that reflect key aspects of diet quality, including fruit, vegetables, grains, dairy, protein foods, fatty acids, sodium, and empty calories. Components scores can range from 0-5, 0-10, or 0-20, and 2010-HEI score ranges from 0 to 100 with a higher score indicating greater consistency of the diet with the 2010 Dietary Guidelines for Americans. **See the SOL Youth Child Derived Variable Dictionary for details on how it was scored.**

References

- Guenther PM, Kirkpatrick SI, Reedy J, Krebs-Smith SM, Buckman DW, Dodd KW, Casavale KO, Carroll RJ. The Healthy Eating Index-2010 is a valid and reliable measure of diet quality according to the 2010 Dietary Guidelines for Americans. J Nutr. 2014 Mar;144(3):399-407.
- Guenther PM, Casavale KO, Reedy J, Kirkpatrick SI, Hiza HA, Kuczynski KJ, Kahle LL, Krebs-Smith SM. Update of the Healthy Eating Index: HEI-2010. J Acad Nutr Diet. 2013 Apr;113(4):569-80.
- NDSR Guide to Creating Variables Needed to Calculate Scores for Each Component of the Healthy Eating Index-2010 (HEI-2010) developed by the Nutrition Coordinating Center (NCC), University of Minnesota, Minneapolis, MN

3 NCI Method to Estimate or Predict Usual Dietary Intake

NOTE: In SOL Youth, predicted usual were not calculated.

National nutritional surveys estimate usual nutrient or food intake to assess dietary deficiencies and excesses and adherence to dietary recommendations. **For assessment of dietary intake with multiple 24-hr dietary recalls, statistical models provide a better estimator of usual intake than simply averaging 24-hr dietary recalls. Dodd et al (Am Diet Assoc, 2006) provide an excellent review on several statistical methods for estimating usual intake:** NRC (National Research Council, ISU (Iowa State University) (Nusser et al, 1996a), Best Power Method (Nusser et al, 1996b), and the NCI Method (Tooze et al, 2006). More recently, Souverin et al (Aur J Clin Nutr (2011) compared these methods with MSM (Multiple Source Method) (Haubrock et al, 2011) and SPADE (Statistical Program for age-adjusted dietary assessment) (Waijers et al, 2006; Deckers).

The National Cancer Institute (NCI) method (http://riskfactor.cancer.gov/diet/usualintakes/method.html) can be used to predict usual nutrient intake. This method can be used to:

- estimate the distribution of usual intake for a population or subpopulation;
- assess the effects of individual covariates on consumption; and
- predict individual usual intake for use in a model to assess the relationship between diet and disease or other variable.

The premise of the NCI method is that usual intake is equal to the probability of consumption on a given day times the average amount consumed on a "consumption day". The methods used for dietary components that are consumed nearly every day by nearly everyone differ slightly from those used for dietary components that are episodically consumed. In general, the former category (ubiquitously consumed or consumed daily by almost everyone) includes most nutrients whereas the latter category (episodically consumed) includes most foods, though there are exceptions. An excellent resource for learning about this methodology is the “Measurement Error Webinar Series” (http://riskfactor.cancer.gov/measurementerror/).

See HCHS/SOL Dietary Overview, Methods and Guidelines V1.1 (June 2013) for a description on how it was specified in HCHS/SOL.

4 RESOURCES

NDSR (Nutrition Data System for Research)
http://www.ncc.umn.edu/index.html)

Usual Dietary Intake
http://riskfactor.cancer.gov/diet/usualintakes/

Measurement error in dietary data (NCI Webinar Series)
http://riskfactor.cancer.gov/measurementerror/

The goals of the Webinar Series are to understand:

- the sources and magnitudes of dietary measurement errors;
- how measurement error may affect estimates of usual dietary intake distributions;
- how measurement error may affect analyses of diet-health relationships;
- how the effects of measurement error may be mitigated.

It is organized by collaborators from NCI, Office of Dietary Supplements, USDA, Gertner Institute, Texas A&M University, and Wake Forest University. The series is intended for nutritionists, epidemiologists, statisticians, graduate students, and others with an interest in measurement error in dietary intake data. Archived webinars are available at: http://riskfactor.cancer.gov/measurementerror/. An intermediate level of familiarity with statistics and dietary assessment is recommended. Webinars of particular interest for section 5 are:

- Webinar #2 “Estimating usual intake distributions for dietary components consumed daily by nearly all persons” describes statistical modeling techniques and data requirements for estimating usual intake.

- Webinar #6 “The problem of measurement error when examining diet-health relationships” explains types and magnitude of measurement error that occur in dietary data, statistical models for evaluating diet-health relationships (including energy adjustment models), and the qualitative and quantitative impact of measurement error on studies of diet-health relationships.

Advanced Dietary Analyses (NHANES)
http://www.cdc.gov/nchs/tutorials/dietary/advanced/index.htm

Other methods to estimate or predict usual intake include:

Iowa State University Method for Estimation of Usual Intake (ISU and ISUF)
http://streaming.stat.iastate.edu/cssm/index.php?option=com_content&view=article&id=38&Itemid=73

The Multiple Source Method (MSM)
https://msm.dife.de/

5 REFERENCES

Dekkers ALM, Verkaik J, Van Rossum CTM, Slob W, Ocké MC.*SPADE: Statistical Program To Assess Dietary Exposure—User's Manual*. National Institute for Public Health and Environment: Bilthoven (in preparation).

Dodd KW, Guenther PM, Freedman LS, Subar AF, Kipnis V, Midthune D, Tooze JA, Krebs-Smith SM. Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. *J Am Diet Assoc* 2006 Oct;106(10):1640-50.

Kipnis V, Midthune D, Buckman DW, Dodd KW, Guenther PM, Krebs-Smith SM, Subar AF, Tooze JA, Carroll RJ, Freedman LS. Modeling data with excess zeros and measurement error: application to evaluating relationships between episodically consumed foods and health outcomes. *Biometrics* 2009 Dec;65(4):1003-10.

Korn EL and Graubard BI (1999). Analysis of Health Surveys. Wiley Inter-Science.

Subar AF, Dodd KW, Guenther PM, Kipnis V, Midthune D, McDowell M, Tooze JA, Freedman LS, Krebs-Smith SM. The food propensity questionnaire: concept, development, and validation for use as a covariate in a model to estimate usual food intake. *J Am Diet Assoc* 2006 Oct;106(10):1556-63.

Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, Thompson FE, Potischman N, Guenther PM, Tarasuk V, Reedy J, Krebs-Smith SM. Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. J Nutr. 2015 Dec;145(12):2639-45.

Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, Guenther PM, Carroll RJ, Kipnis V. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *J Am Diet Assoc* 2006 Oct;106(10):1575-87.

Tooze JA, Kipnis V, Buckman DW, Carroll RJ, Freedman LS, Guenther PM, Krebs-Smith SM, Subar AF, Dodd KW. A mixed-effects model approach for estimating the distribution of usual intake of nutrients: the NCI method. *Stat Med* 2010 Nov 30;29(27):2857-68.

Parson R, Munuo SS, Buckman DW, Tooze JA, and Dodd KW, 2009. User's Guide of Analysis of Usual Intake accessed from http://riskfactor.cancer.gov/diet/usualintakes/macros_single.html

NOTE: For more references, see 'Recommended Resources' at NCI Measurement Error Webinar Series

APPENDIX A. Food Group Derived Variables (dietary recall level)

For information on NDSR Food Group Serving Count System and serving sizes see Appendix 10 of the NDSR 2013 Manual.

Table A. Food Group Derived Variables from 24hr Dietary Recalls

Food Group (Variable name)	Label	HCHS/SOL source variable name	Source Variable Description
FRUIT_CIT	Fruit – Citrus, servings/day	Y09A4	FRU0100 Citrus Juice
		Y09A6	FRU0300 Citrus Fruit
FRUIT_OTH	Fruit – Others, servings/day	Y09A5	FRU0200 Fruit Juice excluding Citrus Juice
		Y09A7	FRU0400 Fruit excluding Citrus Fruit
		Y09A8	FRU0500 Avocado and Similar
		Y09A9	FRU0600 Fried Fruits
		Y09A10	FRU0700 Fruit-based Savory Snack
FRUIT_ALL	Fruit – Others, servings/day	FRUIT_CIT	Fruit – Citrus, servings/day
		FRUIT_OTH	Fruit – Others, servings/day
FRUIT_ALL_WO JUICE	Fruit – Overall, without fruit juice, servings/day	Y09A6	FRU0300 Citrus Fruit
		Y09A7	FRU0400 Fruit excluding Citrus Fruit
		Y09A8	FRU0500 Avocado and Similar
VEG_DARK	Vegetable – Dark-green, servings/day	Y09A11	VEG0100 Dark-green Vegetables
VEG_ORAN	Vegetable – Orange, servings/day	Y09A12	VEG0200 Deep-yellow Vegetables
VEG_TOMA	Vegetable – Tomato, servings/day	Y09A13	VEG0300 Tomato
VEG_WPOT	Vegetable – White Potatoes, servings/day	Y09A14	VEG0400 White Potatoes
VEG_STAR	Vegetable – Starchy, servings/day	Y09A15	VEG0800 Fried Potatoes
		Y09A16	VEG0450 Other Starchy Vegetables
VEG_BEAN	Vegetable – Beans, servings/day	Y09A17	VEG0700 Legumes (cooked dried beans)
VEG_OTH	Vegetable - Other, servings/day	Y09A18	VEG0600 Other Vegetables
		Y09A19	VEG0900 Fried Vegetables
		Y09A20	VEG0500 Vegetable Juice
		Y09A164	MSC0500 Pickled Foods
VEG_ALL	Vegetable – Overall, servings/day	VEG_DARK	Vegetable – Dark-green, servings/day
		VEG_ORAN	Vegetable – Orange, servings/day
		VEG_TOMA	Vegetable – Tomato, servings/day
		VEG_WPOT	Vegetable – White Potatoes, servings/day
		VEG_STAR	Vegetable – Starchy, servings/day
		VEG_BEAN	Vegetable – Beans, servings/day
		VEG_OTH	Vegetable - Other, servings/day

Food Group (Variable name)	Label	HCHS/SOL source variable name	Source Variable Description
VEG_ALL_WOPOTATO	Vegetable – Overall, without potato , servings/day	Y09A11	VEG0100 Dark-green Vegetables
		Y09A12	VEG0200 Deep-yellow Vegetables
		Y09A13	VEG0300 Tomato
		Y09A16	VEG0450 Other Starchy Vegetables
		Y09A18	VEG0600 Other Vegetables
		Y09A20	VEG0500 Vegetable Juice
VEG_ALL_WOFPOTATO	Vegetable – Overall, without fried potato, servings/day	Y09A11	VEG0100 Dark-green Vegetables
		Y09A12	VEG0200 Deep-yellow Vegetables
		Y09A13	VEG0300 Tomato
		Y09A14	VEG0400 White Potatoes
		Y09A16	VEG0450 Other Starchy Vegetables
		Y09A17	VEG0700 Legumes (cooked dried beans)
		Y09A18	VEG0600 Other Vegetables
		Y09A19	VEG0900 Fried Vegetables
		Y09A20	VEG0500 Vegetable Juice
		Y09A164	MSC0500 Pickled Foods
FRUITVEG_WOFPOTATO	Fruit and vegetables without fried potato (servings/day)	FRUIT_ALL	Fruit – Others, servings/day
		Y09A15	VEG0800 Fried Potatoes
		VEG_ALL	Vegetable – Overall, servings/day
GRAIN_REF	Grain – Refined Grain, servings/day	Y09A23	GRS0100 Grains, Flour and Dry Mixes - Some Whole Grain
		Y09A24	GRR0100 Grains, Flour and Dry Mixes - Refined Grain
		Y09A26	GRS0200 Loaf-type Bread and Plain Rolls - Some Whole Grain
		Y09A27	GRR0200 Loaf-type Bread and Plain Rolls - Refined Grain
		Y09A29	GRS0300 Other Breads (quick breads, corn muffins, tortillas) - Some Whole Grain
		Y09A30	GRR0300 Other Breads (quick breads, corn muffins, tortillas) - Refined Grain
		Y09A32	GRS0400 Crackers - Some Whole Grain
		Y09A33	GRR0400 Crackers - Refined Grain
		Y09A35	GRS0500 Pasta - Some Whole Grain
		Y09A36	GRR0500 Pasta - Refined Grain
		Y09A38	GRS0600 Ready-to-eat Cereal (not presweetened) - Some Whole Grain
		Y09A39	GRR0600 Ready-to-eat Cereal (not presweetened) - Refined Grain
		Y09A41	GRS0700 Ready-to-eat Cereal (presweetened) - Some Whole Grain
		Y09A42	GRR0700 Ready-to-eat Cereal (presweetened) - Refined Grain
		Y09A44	GRS0800 Cakes, Cookies, Pies, Pastries, Danish, Doughnuts and Cobblers - Some Whole Grain
		Y09A45	GRR0800 Cakes, Cookies, Pies, Pastries, Danish, Doughnuts and Cobblers - Refined Grain
		Y09A47	GRS1000 Snack Bars - Some Whole Grain
		Y09A48	GRR1000 Snack Bars - Refined Grain
		Y09A50	GRS0900 Snack Chips - Some Whole Grain

Food Group (Variable name)	Label	HCHS/SOL source variable name	Source Variable Description
		Y09A51	GRR0900 Snack Chips - Refined Grain
		Y09A53	GRW1200 Flavored Popcorn
		Y09A54	GRO0100 Baby Food Grain Mixtures
GRAIN_WHL	Grain – Whole Grain, servings/day	Y09A22	GRW0100 Grains, Flour and Dry Mixes - Whole Grain
		Y09A25	GRW0200 Loaf-type Bread and Plain Rolls - Whole Grain
		Y09A28	GRW0300 Other Breads (quick breads, corn muffins, tortillas) - Whole Grain
		Y09A31	GRW0400 Crackers - Whole Grain
		Y09A34	GRW0500 Pasta - Whole Grain
		Y09A37	GRW0600 Ready-to-eat Cereal (not presweetened) - Whole Grain
		Y09A40	GRW0700 Ready-to-eat Cereal (presweetened) - Whole Grain
		Y09A43	GRS0700 Ready-to-eat Cereal (presweetened) - Some Whole Grain
		Y09A46	GRW1000 Snack Bars - Whole Grain
		Y09A49	GRW0900 Snack Chips - Whole Grain
		Y09A52	GRW1100 Popcorn
GRAIN_ALL	Grain – Overall, servings/day	GRAIN_WHL	Grain – Whole Grain, servings/day
		GRAIN_REF	Grain – Refined Grain, servings/day
MEAT_RED	Meat – Red Meat, servings/day	Y09A55	MRF0100 Beef
		Y09A56	MRL0100 Lean Beef
		Y09A57	MRF0200 Veal
		Y09A58	MRL0200 Lean Veal
		Y09A59	MRF0300 Lamb
		Y09A60	MRL0300 Lean Lamb
		Y09A61	MRF0400 Fresh Pork
		Y09A62	MRL0400 Lean Fresh Pork
		Y09A65	MRF0500 Game
MEAT_LUNCH	Meat – Luncheon, servings/day	Y09A63	MCF0200 Cured Pork
		Y09A64	MCL0200 Lean Cured Pork
		Y09A74	MCF0100 Cold Cuts and Sausage
		Y09A75	MCL0100 Lean Cold Cuts and Sausage
MEAT_POUL	Meat – Poultry, servings/day	Y09A66	MPF0100 Poultry
		Y09A67	MPL0100 Lean Poultry
		Y09A68	MPF0200 Fried Chicken - Commercial Entrée and Fast Food
MEAT_FISH	Meat – Fish, servings/day	Y09A69	MFF0100 Fish - Fresh and Smoked
		Y09A70	MFL0100 Lean Fish - Fresh and Smoked
		Y09A71	MFF0200 Fried Fish - Commercial Entrée and Fast Food
		Y09A72	MSL0100 Shellfish
		Y09A73	MSF0100 Fried Shellfish - Commercial Entrée and Fast Food
MEAT_ORG	Meat – Organ Meats, servings/day	Y09A76	MOF0100 Organ Meats

Food Group (Variable name)	Label	HCHS/SOL source variable name	Source Variable Description
MEAT_EGG	Meat – Eggs, servings/day	Y09A79	MOF0300 Eggs
		Y09A80	MOF0400 Egg Substitute
MEAT_NUT	Meat – Nuts, servings/day	Y09A81	MOF0500 Nuts and Seeds
		Y09A82	MOF0600 Nut and Seed Butters
MEAT_SOY	Meat – Soy, servings/day	Y09A83	MOF0700 Meat Alternatives
		Y09A117	DOT0800 Infant Formula - Nondairy
MEAT_ALL	Meat – Overall (servings/day)	MEAT_RED	Meat – Red Meat, servings/day
		MEAT_LUNCH	Meat – Luncheon, servings/day
		MEAT_POUL	Meat – Poultry, servings/day
		MEAT_FISH	Meat – Fish, servings/day
		MEAT_ORG	Meat – Organ Meats, servings/day
		MEAT_EGG	Meat – Eggs, servings/day
		MEAT_NUT	Meat – Nuts, servings/day
		MEAT_SOY	Meat – Soy, servings/day
NUT_LEGUMES	Nut and Legumes, servings/day	Y09A17	VEG0700 Legumes (cooked dried beans)
		Y09A81	MOF0500 Nuts and Seeds
		Y09A82	MOF0600 Nut and Seed Butters
MILK_MILK	Milk – Milk, servings/day	Y09A84	DMF0100 Milk - Whole
		Y09A85	DMR0100 Milk - Reduced Fat
		Y09A86	DML0100 Milk - Low Fat and Fat Free
		Y09A87	DMN0100 Milk - Nondairy
		Y09A88	DMF0200 Ready-to-drink Flavored Milk - Whole
		Y09A89	DMR0200 Ready-to-drink Flavored Milk - Reduced Fat
		Y09A90	DML0200 Ready-to-drink Flavored Milk - Low Fat and Fat Free
		Y09A91	DML0300 Sweetened Flavored Milk Beverage Powder with Non-fat Dry Milk
		Y09A92	DML0400 Artificially Sweetened Flavored Milk Beverage Powder with Non-fat Dry Milk
		Y09A93	SWT0600 Sweetened Flavored Milk Beverage Powder without Non-fat Dry Milk
		Y09A94	MSC1100 Artificially Sweetened Flavored Milk Beverage Powder without Non-fat Dry Milk
		Y09A114	DOT0500 Dairy-based Sweetened Meal Replacement/Supplement
		Y09A115	DOT0600 Dairy-based Artificially Sweetened Meal Replacement/Supplement
		Y09A116	DOT0700 Infant Formula
MILK_CHES	Milk – Cheese, servings/day	Y09A95	DCF0100 Cheese - Full Fat
		Y09A96	DCR0100 Cheese - Reduced Fat
		Y09A97	DCL0100 Cheese - Low Fat and Fat Free
		Y09A98	DCN0100 Cheese - Nondairy
MILK_YOGU	Milk – Yogurt, servings/day	Y09A99	DYF0100 Yogurt - Sweetened Whole Milk
		Y09A100	DYR0100 Yogurt - Sweetened Low Fat
		Y09A101	DYL0100 Yogurt - Sweetened Fat Free

Food Group (Variable name)	Label	HCHS/SOL source variable name	Source Variable Description
		Y09A102	DYF0200 Yogurt - Artificially Sweetened Whole Milk
		Y09A103	DYR0200 Yogurt - Artificially Sweetened Low Fat
		Y09A104	DYL0200 Yogurt - Artificially Sweetened Fat Free
		Y09A105	DYN0100 Yogurt - Nondairy
MILK_DESR	Milk – Dessert, servings/day	Y09A106	DOT0100 Frozen Dairy Dessert
		Y09A108	DOT0300 Pudding and Other Dairy Dessert
		Y09A109	DOT0400 Artificially Sweetened Pudding and Other Dairy Dessert
MILK_ALL	Milk – Overall, servings/day	MILK_MILK	Milk – Milk, servings/day
		MILK_CHES	Milk – Cheese, servings/day
		MILK_YOGU	Milk – Yogurt, servings/day
		MILK_DESR	Milk – Dessert, servings/day
FAT_DISC	Fat – Discretionary Fat, servings/day	Y09A110	FCF0100 Cream
		Y09A111	FCR0100 Cream - Reduced Fat
		Y09A112	FCL0100 Cream - Low Fat and Fat Free
		Y09A113	FCN0100 Cream - Nondairy
FAT_OIL	Fat – Oil, servings/day	Y09A118	FMF0100 Margarine - Regular
		Y09A119	FMR0100 Margarine - Reduced Fat
		Y09A120	FOF0100 Oil
		Y09A121	FSF0100 Shortening
		Y09A122	FAF0100 Butter and Other Animal Fats - Regular
		Y09A123	FAR0100 Butter and Other Animal Fats - Reduced Fat
		Y09A124	FDF0100 Salad Dressing - Regular
		Y09A125	FDR0100 Salad Dressing - Reduced Fat/Reduced Calorie/Fat Free
		Y09A160	MSC0100 Gravy - Regular
		Y09A161	MSC0200 Gravy - Reduced Fat/Fat Free
		Y09A162	MSC0300 Sauces and Condiments - Regular
		Y09A163	MSC0400 Sauces and Condiments - Reduced Fat
FAT_ALL	Fat – Overall, servings/day	FAT_DISC	Fat – Discretionary Fat, servings/day
		FAT_OIL	Fat – Oil, servings/day
SUGAR_SGR	Sugar – Sugar, servings/day	Y09A126	SWT0400 Sugar
		Y09A127	MSC1200 Sugar Substitute
		Y09A128	SWT0500 Syrup, Honey, Jam, Jelly, Preserves
		Y09A129	SWT0700 Sauces, Sweet - Regular
		Y09A130	SWT0800 Sauces, Sweet - Reduced Fat/Reduced Calorie/Fat Free
SUGAR_DESR	Sugar – Dessert, servings/day	Y09A131	SWT0100 Chocolate Candy
		Y09A132	SWT0200 Non-chocolate Candy
		Y09A133	SWT0300 Frosting or Glaze
		Y09A165	MSC0600 Miscellaneous Dessert

Food Group (Variable name)	Label	HCHS/SOL source variable name	Source Variable Description
SUGAR_SWTB	Sugar – Sugar Sweetened Beverage, servings/day	Y09A134	BVS0400 Sweetened Soft Drinks
		Y09A137	BVS0300 Sweetened Fruit Drinks
		Y09A139	BVS0500 Sweetened Tea
		Y09A142	BVS0100 Sweetened Coffee
		Y09A145	BVS0200 Sweetened Coffee Substitutes
		Y09A148	BVS0600 Sweetened Water
		Y09A151	BVS0700 Nondairy-based Sweetened Meal Replacement/Supplement
		Y09A154	BVO0100 Non-alcoholic Beer
		Y09A155	BVO0200 Non-alcoholic Light Beer
SUGAR_SWTB_WJUICE	Sugar Sweetened Beverage, include fruit juice, servings/day	Y09A4	FRU0100 Citrus Juice
		Y09A5	FRU0200 Fruit Juice excluding Citrus Juice
		Y09A134	BVS0400 Sweetened Soft Drinks
		Y09A137	BVS0300 Sweetened Fruit Drinks
		Y09A139	BVS0500 Sweetened Tea
		Y09A142	BVS0100 Sweetened Coffee
		Y09A145	BVS0200 Sweetened Coffee Substitutes
		Y09A148	BVS0600 Sweetened Water
		Y09A151	BVS0700 Nondairy-based Sweetened Meal Replacement/Supplement
		Y09A154	BVO0100 Non-alcoholic Beer
		Y09A155	BVO0200 Non-alcoholic Light Beer
SUGAR_DIETB	Sugar – Diet Beverage, servings/day	Y09A135	BVA0400 Artificially Sweetened Soft Drinks
		Y09A136	BVU0300 Unsweetened Soft Drinks
		Y09A138	BVA0300 Artificially Sweetened Fruit Drinks
		Y09A140	BVA0500 Artificially Sweetened Tea
		Y09A141	BVU0400 Unsweetened Tea
		Y09A143	BVA0100 Artificially Sweetened Coffee
		Y09A144	BVU0100 Unsweetened Coffee
		Y09A146	BVA0200 Artificially Sweetened Coffee Substitutes
		Y09A147	BVU0200 Unsweetened Coffee Substitutes
		Y09A149	BVA0600 Artificially Sweetened Water
		Y09A152	BVA0700 Nondairy-based Artificially Sweetened Meal Replacement/Supplement
		Y09A153	BVU0600 Nondairy-based Unsweetened Meal Replacement/Supplement
SUGAR_ALL	Sugar – Overall, servings/day	SUGAR_SGR	Sugar – Sugar, servings/day
		SUGAR_DESR	Sugar – Dessert, servings/day
		SUGAR_SWTB	Sugar – Sugar Sweetened Beverage, servings/day
		SUGAR_DIETB	Sugar – Diet Beverage, servings/day
WATER	Water, servings/day	Y09A150	BVU0500 Unsweetened Water

Food Group (Variable name)	Label	HCHS/SOL source variable name	Source Variable Description
ALCOHOL	Alcohol, servings/day	Y09A156	BVE0100 Beer and Ale
		Y09A157	BVE0400 Cordial and Liqueur
		Y09A158	BVE0300 Distilled Liquor
		Y09A159	BVE0200 Wine
SNACK_SWT	Snack – Sweet Snack, servings/day	Y09A43	GRW0800 Cakes, Cookies, Pies, Pastries, Danish, Doughnuts and Cobblers - Whole Grain
		Y09A44	GRS0800 Cakes, Cookies, Pies, Pastries, Danish, Doughnuts and Cobblers - Some Whole Grain
		Y09A45	GRR0800 Cakes, Cookies, Pies, Pastries, Danish, Doughnuts and Cobblers - Refined Grain
		Y09A46	GRW1000 Snack Bars - Whole Grain
		Y09A47	GRS1000 Snack Bars - Some Whole Grain
		Y09A48	GRR1000 Snack Bars - Refined Grain
SNACK_SALT	Snack – Salty Snack, servings/day	Y09A49	GRW0900 Snack Chips - Whole Grain
		Y09A50	GRS0900 Snack Chips - Some Whole Grain
		Y09A51	GRR0900 Snack Chips - Refined Grain
		Y09A52	GRW1100 Popcorn
		Y09A53	GRW1200 Flavored Popcorn
SNACK_NUT	Snack – Nuts, servings/day	Y09A81	MOF0500 Nuts and Seeds
		Y09A82	MOF0600 Nut and Seed Butters
SNACK_CRACK	Snack – Cracker, servings/day	Y09A31	GRW0400 Crackers - Whole Grain
		Y09A32	GRS0400 Crackers - Some Whole Grain
		Y09A33	GRR0400 Crackers - Refined Grain
SNACK_VEGFR	Snack – Fruit or Vegetable Savory(servings/day)	Y09A10	FRU0700 Fruit-based Savory Snack
		Y09A21	FMC0100 Vegetable-based Savory Snack
SNACK_ALL	Snack – Overall, servings/day	SNACK_SWT	Snack – Sweet Snack, servings/day
		SNACK_SALT	Snack – Salty Snack, servings/day
		SNACK_NUT	Snack – Nuts, servings/day
		SNACK_CRACK	Snack – Cracker, servings/day
		SNACK_VEGFR	Snack – Fruit or Vegetable Savory(servings/day)

APPENDIX B. Data cleaning for 24-hr recalls based on daily energy intake

There is not a single best way to clean diet data for extreme values. Data cleaning methods depend on the diet instrument used to assess dietary intake, the specific nutrient or food of interest, and the study population, among other things. We suggest the investigator or data analyst consult with someone who has experience with analyzing dietary data. In particular, for 24hr dietary recalls it is recommended that data cleaning is done separately for each dietary recall since intake from 1st recall is known to be higher than intake from the 2nd recall. We have found this to be the case with in HCHS/SOL as well.

Also, consider using self-reported intake amount (Y04A15) and NDSR "Notes from the Trailer tab" or "Food detail Window Notes" (e.g. variables Y04A154 and Y02A141) to understand, determine and clean extreme low or high values. Examples in HCHS/SOL include two participants with zero energy intake due to fasting (identified by DTIA154), and extreme values confirmed (e.g. "14 chicken wings eaten. Confirmed by participant" identified by DIEA141), etc.

Among the 1,466 children in SOL Youth, 96% (n=1,428) have the 1st dietary recall (in-person at clinic visit) and 93% (n=1384) have the 2nd recall (most conducted by telephone). NDSR files have data from all these recalls.

To create FOOD_GROUP_DERV dataset and variables for the HEI-2010 we excluded 54 dietary recalls because the daily energy intake (variable Y04A20) was considered too extreme (defined as below the sequence-gender specific 1st percentile or above the 99th percentile; table B). As a consequence, 1,400 children have the 1st dietary recall, 1,384 have the 2nd recall, 1,305 (90%) have both recalls and 13 have none.

Table B. Energy intake (kcal/day) percentiles 1 and 99 by dietary recall and sex, SOL Youth

Recall	Sex	N	Percentile 1	Percentile 99
1st	Girl	727	480.50	3558.54
2nd	Girl	700	441.85	3257.04
1st	Boy	701	519.36	4287.35
2nd	Boy	684	512.34	3764.39
Total		2,812		