



# USDA's National Household Food Acquisition and Purchase Survey: Methodology for Imputing Missing Quantities To Calculate Healthy Eating Index-2010 Scores and Sort Foods Into ERS Food Groups

Lisa Mancino, Jessica E. Todd, and Benjamin Scharadin

## What Is the Issue?

USDA's National Household Food Purchase and Acquisition Survey (FoodAPS) collected detailed information about the types of food households acquire over the course of 1 week, between April 2012 and January 2013. One objective of the survey was to evaluate the nutritional quality of household food acquisitions as measured against the 2010 Healthy Eating Index (HEI-2010) and whether economic and sociodemographic factors affect the nutritional quality of these acquisitions. Total grams or calories can be used to normalize expenditures for comparing food quality across households or venues; however, complete quantity information is needed for research questions not related to nutritional quality. Unfortunately, information on the quantities acquired is missing for some items. Rather than have individual researchers grapple with how to handle these missing quantities, ERS has added imputed quantities to the FoodAPS data. Also, the ability to study specific categories of food is useful for research on policies that affect those categories, such as subsidies for fruits and vegetables or taxes on certain beverages. Therefore, ERS developed a food classification scheme for the data that groups foods in relation to their main ingredient, quality, and likely price premiums for convenience and processing. Researchers now have more than one choice when aggregating the FoodAPS items into groups for analysis.

## What Did the Study Find?

The HEI-2010 is calculated by considering the proportions in which foods are consumed or are included in a shopping basket or pantry. Thus, knowing the quantity of each item acquired is essential. When quantity information is missing, the researcher must either drop the item from HEI-2010 calculations or impute a quantity so it can be included. ERS researchers developed a methodology to impute missing quantities based on all available information about each item, including the item description, the type of store where it was purchased, the geographic location of the purchasing household, the amount paid for the item (when available), and the size of the household. A dataset with imputed quantities has been posted to the FoodAPS public use files on the ERS website. This allows researchers to include all FoodAPS items with nutrient information in calculations of HEI-2010 scores and component densities and will facilitate comparisons across various studies.

ERS is a primary source of economic research and analysis from the U.S. Department of Agriculture, providing timely information on economic and policy issues related to agriculture, food, the environment, and rural America.

We compute and compare household-level HEI-2010 scores over the 7 days of FoodAPS to individual-level HEI-2010 scores from 2 days of dietary recall data from the 2011-12 National Health and Nutrition Examination Survey (NHANES). In both cases, we use the simple algorithm to generate household-level (or respondent-level, in the case of NHANES) HEI-2010 component densities and then estimate means. The comparison of means indicates that while there are some slight differences in HEI-2010 component densities, both surveys show that Americans acquire too few fruits, vegetables, and whole grains, too many refined grains and empty calories, and too much salt. The Stata programs used to calculate HEI-2010 scores for FoodAPS acquisitions and data with imputed quantities are provided to users on the ERS website.

Also in this study, ERS developed a food classification system for FoodAPS, based on the 2010 *Dietary Guidelines for Americans* as well as on other aspects of food items, like the convenience and processing that consumers consider when purchasing food. This classification system differs from others, such as the USDA's *What We Eat in America* food categories, because it allows data users to separate foods not only by main ingredient and nutritional characteristics, but also by level of convenience and form, such as fresh, canned, or frozen. Each food item was classified into 1 of 82 ERS food groups using all available information about the item.

## **How Was the Study Conducted?**

Information about food items, the stores from which they were obtained, and household characteristics available in the FoodAPS data were used to impute quantities when missing. The SAS programs developed by the U.S. Department of Health and Human Services, National Cancer Institute for HEI-2010 scores were adapted to FoodAPS using Stata. Detailed item-level information was used to classify foods into 82 ERS food groups. Stata 12.2 was used for all quantity imputations. Sorting food items into food groups was performed using a series of programs in both SAS 9.0 and Stata.