38th National Nutrient Databank Conference

Towards Measuring the Food Quality of Grocery Purchases: an Estimation Model of the Healthy Eating Index-2010 Using only Food Item Counts

Le-Thuy T. Tran[a], Philip J. Brewster[a], Valliammai Chidambaram[a], John F. Hurdle[a],*

[a]*Department of Biomedical Informatics, University of Utah School of Medicine, Salt Lake City 84108, Utah, USA*

Abstract

Measuring the quality of food consumed by individuals or groups in the U.S. is essential to informed public health surveillance efforts and sound nutrition policymaking. For example, the Healthy Eating Index-2010 (HEI) is an ideal metric to assess the food quality of households, but the traditional methods of collecting the data required to calculate the HEI are expensive and burdensome. We evaluated an alternative source: rather than measuring the quality of the foods consumers eat, we want to estimate the quality of the foods consumers buy. To accomplish that we need a way of estimating the HEI based solely on the count of food items. We developed an estimation model of the HEI, using an augmented set of the What We Eat In America (WWEIA) food categories. Then we mapped ~92,000 grocery food items to it. The model uses an inverse Cumulative Distribution Function sampling technique. Here we describe the model and report reliability metrics based on NHANES data from 2003-2010.

1. Background and objective

There is sound literature-based evidence for the correlation between dietary habits and disease.[1-7] This has led to a heightened interest in the monitoring and assessment of the nutritional status of patients with diet-related disease, as well as an interest in public health nutrition assessment. The common methods of nutritional monitoring

---

 * Corresponding author: John F. Hurdle, MD, PhD. Tel.: +1-801-213-3232; fax: +1-801-581-4297.
   E-mail address: john.hurdle@utah.edu

```
Nomenclature

CDF              Cumulative Distribution Function
DGA              Dietary Guidelines for Americans
FNDDS            Food and Nutrient Database for Dietary Studies
FPED             USDA Food Patterns Equivalents Database
HEI or HEI-2010  USDA/NCI Healthy Eating Index (2010)
MPED             USDA MyPyramid Equivalents Database
SoFAAS           Solid Fat, Alcohol, Added Sugars
USDA             United States Department of Agriculture
WWEIA            USDA What We Eat In America food categories
```

and assessment listed in Brinkerhoff et al.[8] including dietary recall, food diaries, food frequency questionnaires, household inventories, and surveys. However, these approaches are costly, time-consuming, burdensome to consumers and subject to self-reporting errors and recall bias – especially in longitudinal work.

We want to evaluate an alternative approach: rather than measuring the quality of the foods consumers eat, we want to estimate the quality of the foods consumers buy. We plan to start with household grocery purchases, and our working hypothesis is that grocery data will be a strong predictor of overall dietary quality. Because grocery retailers collect sales data for internal purposes, such data promise to be inexpensive for research use. It is also free from self-report errors/bias and collecting it poses no burden on consumers. To accomplish this, we need a validated food quality metric paired with a way to apply that metric based solely on the counts of foods purchased

1.1. The Healthy Eating Index 2010 (HEI-2010)

The Healthy Eating Index (HEI) as mentioned in Guenther et al.[9, 10] is a measure of overall diet quality that encapsulates the United States Dietary Guidelines for Americans (DGA). The HEI-2010 is the latest HEI iteration, and we use "HEI" here to refer to this version. One of the great strengths of the HEI is that it can measure quality at various levels of the food stream, e.g., the national food supply, the community food environment, restaurants, and individual food intakes as discussed in[11-13]. This makes the HEI an excellent candidate as a scoring metric for our work.

The HEI assigns scores to 12 food components in a way that increases the overall quality score based on 8 food components that consumers should eat to maintain an adequate diet (e.g., total fruit, whole grains, seafood and plant proteins), while decreasing the overall quality score based on 4 components that consumers should choose in moderation (refined grains, saturated fats, sodium, and empty calories from solid fats, alcohol, and added sugars). The USDA encoded nutrition information from the ~8,000 foods in the Food and Nutrient Database for Dietary Studies (FNDDS) into the MyPyramid Equivalents Database (MPED), which summarizes the contribution of these foods to the components used by the HEI, based on standardized portions (100-gram ounce or cup equivalents). Using the MPED, which is now called the Food Patterns Equivalents Database (FPED), the HEI normalizes the amount of the foods eaten, per 100 grams, to each 1,000 kcal of total energy, and then scores the MPED-based food components. The HEI components and their corresponding scoring ranges are shown in Table 1.

1.2. The MyPyramid Equivalents Database (MPED) and the Food Patterns Equivalents Database (FPED)

MPED 2.0 provides the number of MyPyramid equivalents of the food groups and subgroups that are present in 100 grams of the 6940 food codes and 811 food modification codes reported in the WWEIA survey from 2003-2004. The Center for Nutrition Policy and Promotion (CNPP) addendum to the MPED 2.0 supplies data for 942 additional food items found in the WWEIA survey from 2005-2006. In addition, the CNPP provided supplemental data for whole fruit and fruit juices. Later replacing the MPED 2.0, the FPED provides the conversions for food codes and food modification codes reported in the WWEIA 2009-2010 surveys. There are a total of 8190 food codes